

画像オープンデータ化推進のための人工知能技術を活用した支援システムの構築と実践

屠 芸豪 (名古屋大学 大学院情報学研究科, tu.yunhao@a.mbox.nagoya-u.ac.jp)

河野 祐希 (元名古屋大学 大学院情報学研究科, river.u.878@gmail.com)

浦田 真由 (名古屋大学 大学院情報学研究科, mayu@i.nagoya-u.ac.jp)

遠藤 守 (名古屋大学 大学院情報学研究科, endo@i.nagoya-u.ac.jp)

安田 孝美 (名古屋大学 大学院情報学研究科, yasuda@i.nagoya-u.ac.jp)

Construction and practice of a support system utilizing artificial intelligence technology to promote image open data

Yunhao Tu (Graduate School of Informatics, Nagoya University)

Yuki Kawano (Ex-Graduate School of Informatics, Nagoya University)

Mayu Urata (Graduate School of Informatics, Nagoya University)

Mamoru Endo (Graduate School of Informatics, Nagoya University)

Takami Yasuda (Graduate School of Informatics, Nagoya University)

要約

近年、社会全体でデジタル化が進み、行政をはじめとする様々な分野において、データと情報通信技術との利活用が重要になってきている。行政分野においては、自治体における行政運営や社会課題に対応するために、情報通信技術を活用し、データに基づく政策立案が必要である。そのため、自治体におけるオープンデータをはじめとするデータの推進と利活用が重要となる。本研究では、自治体での画像オープンデータ化の実践に向け、産官学連携により人工知能技術を活用した支援システムを構築する。現状、全国の自治体では文章や表形式のオープンデータの推進は進んでいるものの、肖像権やプライバシー等の問題により、画像のオープンデータ化については普及が進んでいない。そこで本研究では、画像のオープンデータ化を推進するための支援システムを開発した。加えて、産官学連携により自治体における実証実験を通じて支援システムを検証した。その結果、自治体の保有する画像のオープンデータ化に対する実現可能性と有効性を確認することができた。これにより、画像オープンデータの利活用に繋げると考える。また、自治体における画像オープンデータに関連する業務負担を軽減し、業務効率の向上を図る。更に、本研究を通じて自治体業務に人工知能技術を活用した効果と意義を検証し、人工知能技術の活用に対する理解とイメージを深めることを目指す。

キーワード

画像オープンデータ, 産官学連携, 深層学習, 画像認識, 匿名化処理

1. はじめに

近年、様々な分野において人工知能技術をはじめとする情報通信技術（以下 ICT）を活用した取り組みが活発に行われている。社会全体でデジタル化が進み、ビッグデータと呼ばれるような大規模なデータが生み出されている。政府は「第6期科学技術基本計画」を策定し、科学技術イノベーション政策やビッグデータの活用を積極的に推進している（内閣府, 2021）。これらのことにより、地方自治体は行政業務に人工知能技術や官民データを活用し、業務の効率化と市民サービスの向上を図っている（総務省, 2019a）。現在、我が国では少子高齢化が加速し、自治体において限られた資源や財源の中で住民サービスの質を維持・向上させていくためには、自治体の保有するデータを利活用し、根拠に基づく政策立案が必要である。政府は「官民データ活用推進基本法」や「地方公共団体におけるデータ利活用ガイドブック」を策定し、個人情報

の保護に配慮しつつ、地方公共団体の保有するデータを政策立案や住民サービスの向上等に効果的に活用することを図っている（総務省, 2019b）。「官民データ活用推進基本法」において課題の解決や行政の高度化・効率化を推進するために、国および地方公共団体においてオープンデータへの取り組みが義務付けられている。現状、政府主導によりオープンデータへの取り組みが積極的に進められているが、都道府県での取り組みと比較し、市町村における取り組みの進捗は芳しくない。具体的には、都道府県のオープンデータ取組率は平成30年3月に100%を達成したのに対し（内閣官房情報通信技術（IT）総合戦略室, 2019）、令和3年4月12日時点で地方公共団体におけるオープンデータ取組率は約65%に留まっている（総務省, 2021）。また、オープンデータの質・種類・取り組み内容は自治体によって異なる。現在公開されているオープンデータの多くは、文章形式または表形式のデータである。これ以外に画像、音声および地理空間情報といったデータをオープンデータとして公開している自治体があるが、事例は少なく、表形式以外のデータのオープンデータ化は普及していない。特に直感的に分かりや

すい情報が多く含まれる画像データの公開が少ない。

そこで本研究では、産官学連携により自治体における画像のオープンデータ化に取り組むための方法や実装環境を整備し、人工知能技術である深層学習を活用した支援システムを開発する。加えて、愛知県の高浜市と日進市にて実証実験を行い、本システムの有効性と課題点を検証する。本取り組みにより、自治体において画像オープンデータの推進を支援し、業務負担を軽減して業務効率化の向上を目的とする。また、画像オープンデータから付加情報を抽出し、自治体におけるデータ利活用のグッドプラクティス事例に繋げることを目指す。本研究を通じて自治体における画像オープンデータの推進を実践し、深層学習を用いて支援することで、画像オープンデータ公開の促進を目指す。

2. オープンデータ

2.1 オープンデータの意義

オープンデータは営利目的の有無と関係なく二次利用可能であり、誰も無償で利用することができ、かつ機械判読に適したデータである（高度情報通信ネットワーク社会推進戦略本部，2019）。2012年7月に高度情報通信ネットワーク社会推進戦略本部が策定した「電子行政オープンデータ戦略」に基づき、行政業務や事業者が保有するデータの公開が進められてきた（高度情報通信ネットワーク社会推進戦略本部，2012）。官民データ活用推進基本法において国および地方公共団体が保有する官民データの容易な利用等が基本的施策として定められている。また、2017年5月にIT総合戦略本部が策定した「オープンデータ基本指針」では、公共データについてオープンデータを前提とした情報システムや業務プロセス全体の企画・整備および運用を行う「オープンデータ・バイ・デザイン」の考え方にに基づき、各省庁が保有するデータのオープンデータ化が原則として示されている。また、同指針は国と地方公共団体および事業者が保有する官民データを、容易に利用することも求めている。さらに、2018年に閣議決定された「世界最先端デジタル国家創造宣言・官民データ活用推進基本計画」において、2020年度までに地方公共団体のオープンデータ取組率を100%にすることが目標として掲げられてきた。

オープンデータとして公開されているデータは、様々な場面で活用されはじめている。例えば、新型コロナウイルス対策ダッシュボード（2021）では、都道府県別のPCR検査陽性者数、退院者数・死亡者数および感染症病床使用率等を示している。関連する機関が公開しているデータと連動し、リアルタイムに情報を伝えている。また、防災啓発マップ（2020）は、名古屋市近隣の7つの市町が公開している防災に関するオープンデータ等を活用し、避難時の避難場所や方法を啓発するためのものである。このようにオープンデータを利活用し、人々の生活を支える取り組みが進められている。

オープンデータは社会の変革に繋がる要因として政府より提唱されるSociety5.0において重要な位置づけとな

る（内閣府，2016）。より多くの人がオープンデータを自由に使うために、著作権に対応する使い方が重要である。そこで、クリエイティブ・コモンズ・ライセンス（CCライセンス）を使用し、データ提供者の権利を守ると同時に、利用側にとって容易にオープンデータを利用することができる。CCライセンスは画像のみならず、ウェブサイト上で存在するデータにも対応するライセンスである。CCライセンスは表示、非営利、改変禁止、継承といった4種の条件で構成される。これらの条件を組み合わせて6種類のライセンスを構成する。これにより、著者は著作権を保持しながら、データを自由に流通させることができる。また、データの利用側にとってCCライセンスで定まった条件を守れば公開データを自由に使うことができる（クリエイティブ・コモンズ・ジャパン，2021）。これらの取り組みにより、画像を含めるオープンデータを容易に利用することができ、新たな価値創造に繋がると考える。

2.2 オープンデータの現状

現在、公開されているデータセットの大部分は文章形式または表形式のデータである。自治体は総務省行政管理局が運営するデータカタログサイトにてオープンデータを公開している（総務省，2021）。同サイトにて掲載されているデータは国の府省庁によって公開されているものであるが、地方自治体においても同様である。11月1日時点で公開データに表形式のデータが多い（表1）。地方自治体が公開しているオープンデータの多くは、人口データやAED設置場所・施設情報等のデータである（鈴木他，2020）。これらのデータは政府が公開を推奨する「推奨データセット」と呼ばれるデータに該当する（内閣官房情報通信技術（IT）総合戦略室，2021）。表1によれば、PDF形式のデータが最も多く、JPEGという画像形式のデータが比較的に少ないことがわかる。

一方、文章や表形式以外のデータについて、いくつかの先進的な自治体にて公開の推進が進められている。例えば、石川県金沢市（2021）や愛知県東浦町（2021）にて地域の観光資源や歴史・文化に関する画像をサイトにて公開する取り組みが行われている。動画のオープンデー

表1：データカタログサイトより公開されるデータ形式ごとのデータ数

データ形式	データ数
PDF, pdf	13,586
HTML, html	7,460
XLS, xls	5,787
CSV, csv	2,808
XLSX,xlsx	1,577
ZIP, zip	892
JPEG, jpeg	412
XML	159
GIF	91

タとしては、福島県郡山市にて撮影した観光素材の映像を公開している（郡山市観光協会，2021）。地理空間情報のオープンデータ化について、北海道室蘭市（2021）や東京都世田谷区（2021）にてシェープファイルの公開のみならず、GISを活用してデータを地図上で直感的に確認する機能をカタログサイト上に実装する取組が行われている。しかし、現状ではこのような事例はまだ少ない。特に、画像のオープンデータ化が進まない理由として以下が考えられる。1つ目は、画像データの公開方法が、未だ確立されていないことが挙げられる。表形式のデータは推奨データセットとして公開すべきデータの種類、公開時のフォーマットとルールがまとめられている。一方、画像データをどのような形で公開すべきかが確立されていないことより、取り組むための方法が分からず、着手する難易度が高いと考える。2つ目は、肖像権と個人情報保護法を考慮して個人特定のできる写真の公開が難しい。例えば、イベントやお祭りの写真に人の顔が写っており、特定の人物だと判定できる場合がある。これらのことより、画像データの公開が難しくなり、公開には多くの労力とコストがかかる。そこで、画像の公開方法と個人情報保護を配慮する方法を確立すれば、より多くの価値ある画像データを公開することが可能になる。画像は表形式のデータに比べ直感的に分かりやすい情報であり、かつ画像の利用に特別な分析や知識を要さない場合がほとんどである。これらのデータを公開し誰もが利用可能とすることで、自治体業務を効率化して結果的に市民サービスの向上に繋がることを期待される。

2.3 画像オープンデータの意義

画像オープンデータとは、イラストデータ、写真データおよび動画データ等のデータをオープンデータ化したものである。本研究では主に写真データを対象としているが、他のデータにも適応可能である。自治体は画像オープンデータを推進するには多くの利点がある。1つ目は、観光地やイベント等の自治体の魅力を視覚的に表現することで、多種多様な主体からシティプロモーションの活性化に繋がると考える。また、市民や企業はポスターやパンフレットの素材として自由に使い、自治体の良さをアピールすることができる。データの二次利用を可能にすることで、自治体からの発信のみならず、市民や企業からの発信により地域の活性化に繋がると考える。2つ目は、自治体における画像提供の問い合わせへの対応にかかる労力を減らせることである。画像に限らず他の形式のデータについても、多くの自治体職員にとって情報公開請求への対応が負担となっている。特に画像の利用については各種メディアから画像提供の問い合わせが来る。加えて、自治体内において他の部門に対して画像の提供を求める場合がある。そこで、事前に公開可能な画像をあらかじめ公開しておくことで、問い合わせ等の行政業務をスムーズにすることが期待される。例えば、大阪市では地域資料の保存と情報利活用のために、大阪市立図書館がデジタルアーカイブを運営している。従来、この

アーカイブに寄せられる二次利用の申請が多くあり、事務作業が膨大であった。そこで、デジタルアーカイブの中から著作権が消滅した地域資料をオープンデータとして公開したところ、二次利用申請数は公開から2年間で約6割減少し、オープンデータ化に伴う新聞報道とも相まってアーカイブへのアクセス数は2年で2.8倍に伸びた。また、オープンデータとして公開されたコンテンツは、イベントの広報や記念品、バスのラッピングデザインおよびレトルト食品のパッケージ等で使用されている（総務省，2019c）。これにより、シティプロモーションにも繋がっていく。3つ目は、画像は言葉の壁を超える情報形式なので、コミュニケーションに活用しやすいことである。画像は言葉では伝わらないことを分かりやすく伝えることができ、日本語がわからない外国人にも情報を提供することができる。そのため、画像形式のイラストや写真による情報公開は、インバウンド対策への活用にも役立つと考える。

2.4 画像オープンデータ推進に関する先行研究

画像のオープンデータ化には多くのメリットがあるが、公開方法が確立されていないこと、そして被写体に関する肖像権やプライバシー等の問題があるため、画像オープンデータの推進が進んでいないことを述べてきた。また、画像オープンデータに向けた処理基盤が少なく、画像データの管理と公開を行うには、応分のコストと労力がかかる。鈴木らは画像オープンデータの管理と公開に向け、ウェブガイドラインを作成して公開している（画像オープンデータガイドライン，2021）。彼らは、画像オープンデータに関する基礎知識や画像公開の権利に関する注意事項や対処方法をガイドラインとしてまとめ、ウェブページに公開した。また、画像オープンデータに関する理解度の確認クイズ機能、質問およびノウハウを投稿する機能等も用意されている（鈴木他，2019）。更に、画像を公開する際の標準フォーマットも検討しており、画像を公開する際のメタデータの処理方法やフォーマットを提案した。加えて、すでに公開されている他の事例と比較し、自治体の画像保有状況および画像公開のニーズに対するヒアリングを行った。その結果、オープンデータとして画像を公開するサイトを作成し、実験的に画像オープンデータの公開を行っている。

2.5 本研究の提案と実践

先行研究の取り組みにより、自治体が画像オープンデータに取り組む際の環境が整備されつつある。しかし、依然として画像オープンデータへの取り組みには労力のかかる部分が残されているので、更なる改善の余地がある。そこで本研究では、自治体が画像オープンデータに取り組む際の標準的な方法を整備し、画像の匿名化処理を考慮した支援システムを構築する。具体的には、深層学習による画像認識を活用し、顔検出により画像における人の顔を検出し、モザイクをかける匿名化処理を行う。その上で、自治体の保有するさまざまな画像からオープン

データに適したデータを抽出し、画像オープンデータ推進の業務を支援するためのシステムを開発する。また、自治体における実証実験を通じて本システムの有効性を検証し、課題点等を明らかにする。これらの実践により、自治体における画像オープンデータ推進のハードルを下げ、業務の効率化を目指す。加えて、自治体業務における ICT を活用した実践例を示すことで、オープンデータの推進と利活用を活性化させていくことを目指す。

3. 画像オープンデータ推進支援システム

3.1 人工知能技術の発展と活用

人工知能研究は 1950 年代から続いており、ブームと冬の時代があった。現在は 2000 年代から続いてきた 3 度目のブームとなる（総務省，2016）。1 回目のブームでは、人工知能技術が複雑な問題に適応できなかったことから、冬の時代になった。2 回目のブームでは、人間は必要な情報をすべてコンピューターの処理できる形で記述し、コンピューターに入れる必要があった。そのため、労力とコストがかかり、活用できる領域が限定的であった。そこで、3 度目のブームでは、深層学習をはじめとする人工知能技術の発展により、データから特徴となる知識を自動的に取得することができるようになった。これにより

画像認識や自然言語処理など幅広い分野において、深層学習を用いた実用化が進み、以前に比べ格段の進歩を遂げつつある（LeCun et al., 2015）。本研究では深層学習による画像認識を活用し、自治体の保有する各種画像からオープンデータに相応しい画像を効率的に抽出することを目指す。これにより、自治体における業務負担を減らすと同時に、画像オープンデータの活用にも繋げる。本研究では画像認識技術のうち、顔検出とシーン認識を用いる。1 つ目は、顔検出を活用して画像中の人の顔の有無を判定する。2 つ目は、検出された顔にモザイク処理をかける画像の匿名化処理を行う。これにより、オープンデータとして使用可能なデータにする。3 つ目は、シーン認識を活用して 1 枚の画像に存在する複数のオブジェクトを認識し、画像の自動ラベリングを行う。これにより、公開画像に付加価値をつけることを検証する。

3.2 システムの構成

本研究で構築した自治体における画像オープンデータ推進のための支援システムの概念を図 1 に示す。本システムは「画像選定機能」、「画像匿名化機能」、「公開画像への自動ラベリング機能」といった 3 ステップの機能より構成される。「画像選定機能」では、自治体の所有する

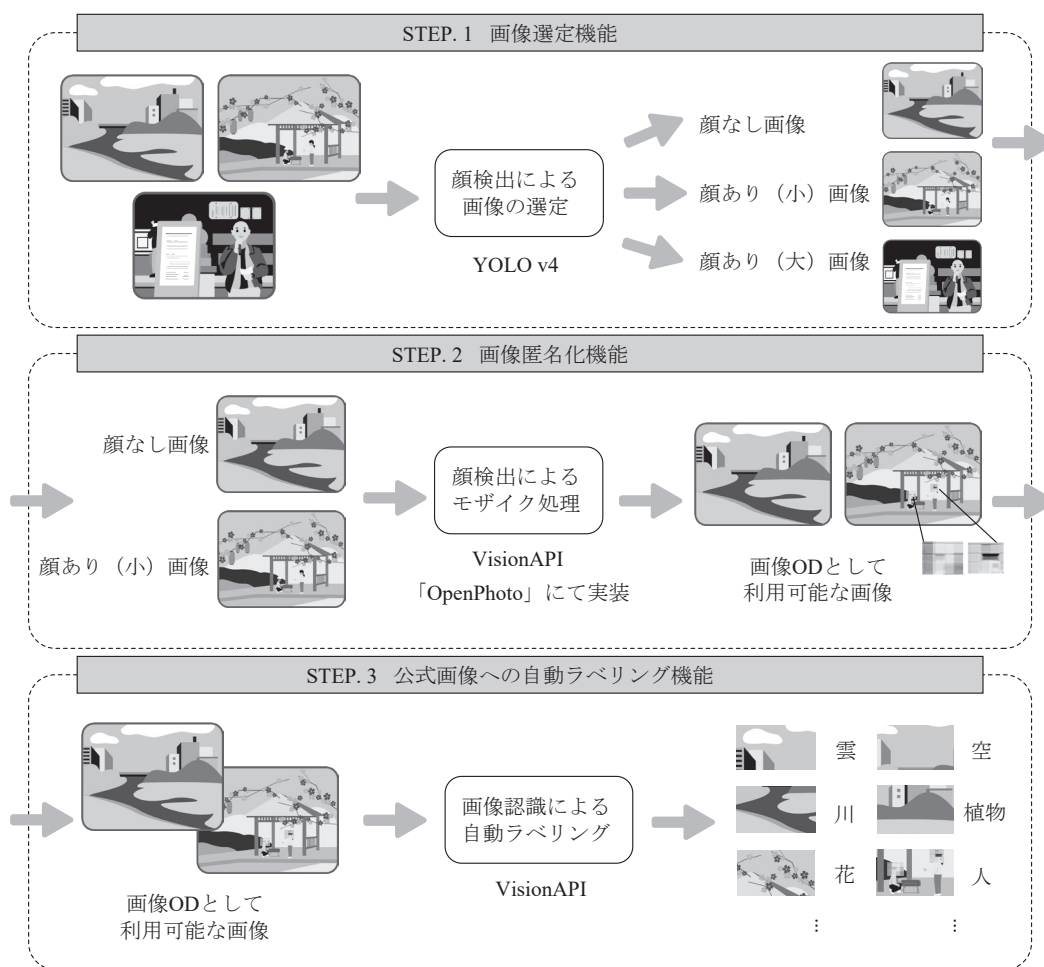


図 1：本システムの概念図

画像データを対象とし、画像に存在する顔の写り具合によって画像を分類する。このステップにて顔が大きく写っている画像を取り除く。顔が小さく写っている画像と顔の正面が写っていない画像を保存して次のステップに送る。「画像匿名化機能」では、選定された画像の中から肖像権への配慮が必要な写真に対し、顔にモザイクをかける匿名化処理により、画像オープンデータとして利用可能な画像とする。「公開画像への自動ラベリング機能」では、公開画像の活用幅を広げるために、シーン認識技術を活用して画像から付加的な情報を抽出することで、画像オープンデータの利活用に繋げる。

3.3 画像選定機能

本機能では、自治体の保有する大量な画像から公開に適した画像を選定することで、オープンデータの業務を支援する。先行調査によると、愛知県名古屋市近郊の7つの市町において、観光に携わる課の6割以上が50001枚以上の画像を保有している（鈴木他, 2019）。これら大量の画像から、オープンデータとして公開できそうな画像を探し出すには多くの時間とコストがかかる。そこで



図2：顔検出により選定された画像の例

注：出典：平成大橋開通（東浦町画像オープンデータ：クリエイティブ・コモンズ・ライセンス表示 4.0 国際）を改変して筆者作成。

画像選定機能では深層学習による顔検出を活用して画像中に顔が写っているかどうかを確認する。顔の写り具合により、下記の3つのカテゴリーに画像を分ける。1つ目は、顔が写っていない「顔なし画像」を抽出する。2つ目は、顔が写っていたとしても十分に小さいと判定される「顔あり（小）画像」を抽出する。顔が十分に小さい判定基準は顔検出の結果にて各顔に付与されたバウンディングボックスの面積が画像全体面積の0.15%未満とする。顔なし画像と顔あり（小）画像は公開できそうな画像として用いる。図2に顔の写り込みが十分小さいと判定された例を示す。愛知県東浦町より公開されている画像オープンデータを用いて本機能の検出結果を説明する。画像右側の男性が画像内で最大面積の顔の写り込みとなり、占有率は0.12%である。この画像に対して後述する画像匿名機能を行えば画像の価値が大きく損なわれないと判断し、オープンデータとして抽出することができる。3つ目は、画像にある顔の割合が大きい「顔あり（大）画像」を抽出する。この画像に対して画像匿名機能を用いると元画像の情報が損なわれる可能性があるため、オープンデータとして相応しくないと判定する。本機能を実装するために、比較的高速に計算する物体検出アルゴリズムYOLOv4を用いた（Bochkovskiy et al., 2020）。YOLOv4は物体検出アルゴリズムYOLOシリーズの1つであり、高速に計算することが特徴である。現状、各地方自治体は汎用的な計算機を用意するのが予算面で現実的ではない。そこで、自治体のローカル環境において軽量型で速く実行可能なプログラムが必要である。

3.4 画像匿名化機能

本機能では顔検出より検出された顔にモザイク処理を行う。図3に本機能の処理プロセスを示す。画像選定機能より選定された画像の中で、画像に写り込んでいる顔に対してモザイク処理を自動に行うことができれば、自治体職員の作業負担を軽減すると同時に、オープンデータ化する画像の選択肢を増やすことができる。例えば、イベントや地域のお祭りの写真画像では、主な対象物と



図3：顔検出による画像匿名化機能の使用例

注：出典：平成大橋開通（東浦町画像オープンデータ：クリエイティブ・コモンズ・ライセンス表示 4.0 国際）を改変して筆者作成。

は別に複数の人の顔が小さく写り込んでいる場合がある。そこで、これらの画像中のすべての顔にモザイクをかけ、個人が特定できないように匿名化処理を施す必要がある。しかし、公開するすべての画像に対して人手でモザイクをかけていくには大変な労力がかかり、コストが高い。そこで本研究では、顔検出より検出された顔に自動的にモザイクをかけることで、作業内容を軽減してより多くの画像データを公開することができるようにする。

本機能では今後のデータ公開を見据え、共同研究先であるインフォ・ラウンジ株式会社の協力により、OpenPhoto（2021）というサービスの画像登録ページにて顔検出による検出された顔にモザイク処理を行う機能を実装した。OpenPhotoは写真ギャラリーを作成するクラウドプラットフォームであり、画像ごとのライセンスの指定や画像へのメタデータの付与等をウェブサイト上で簡単に行うことができる。自治体の職員は画像をOpenPhotoに登録する際、ワンクリックで顔検出を行った結果に応じてモザイクをかけるかどうか選択することができる。本機能ではGoogle Cloud Vision API（2021）を導入して使用した。Vision APIは画像の認識や解析を行うオンラインサービスである。本研究のフィールドとなる高浜市と日進市など、多くの自治体はOpenPhotoを利用し、画像オープンデータを公開している。また、画像オープンデータの公開サイトとして使用するOpenPhotoはGoogle Cloud Platformにて実装されており、Vision APIと容易に連携を取ることことができる。更に、同じサーバーにて公開と管理の作業を実装して検出等の処理を行うことで、画像データ等を管理しやすくなる。本機能により画像の価値を大きく損なうことなく、公開に適した形に編集することが可能となった。

3.5 公開画像への自動ラベリング機能

本システムでは画像データの公開のみならず、公開後のデータ利活用にも取り組んでいる。そのため、画像から物体や風景といった情報をラベルとして抽出し、画像に付加して公開することにより、新たな利活用に繋げられる。これにより、様々なサービスを創出し、オープンデータを活用する取り組みの幅を広げることが期待される。例えば、ユーザーから公開した画像へのアクセス数を計測し、画像のラベルとの関係を分析することでニーズが大きい画像オープンデータを抽出することができる。また、画像のラベルをもとにコンセプトに近い他の画像と関連付けることにより、市町村の壁を超えて横断的に画像オープンデータを探することができる。加えて、機械判読に適した形で付加的な情報をまとめることでプログラムによる判別や加工が容易になる。例えば、もとの画像に雲や建物が存在するなら、画像から雲や建物といったラベルを抽出して保存する。これにより、画像の自動ラベリングができ、一緒に公開することができる。

本機能では深層学習によるシーン認識を用いて画像のラベルを抽出する。また、抽出された結果を画像に紐付けて保存する。シーン認識は難易度の高いタスクである

ため、最初からモデルを構築するには膨大なデータ、計算リソースおよび高度な専門知識が必要となる。そこで、画像匿名化機能と同じくGoogle Cloud Vision APIを用いて画像のラベルを抽出する。Vision APIでは1枚の画像から十数個程度のラベルを認識して抽出することができる。

4. 実証実験

4.1 実証実験の概要

実証実験では、愛知県高浜市と日進市より協力を受けた。なお、高浜市からは本支援システム全体に関する協力を受けた。また、日進市からは画像選出機能の実験まで協力を受けた。自治体職員が本システムのステップに従って画像オープンデータの作業を行い、本システムの性能や実用性について検証を行った。これにより、本システムの有効性、実現可能性および改善点等を明らかにする。

実証実験では自治体が提供可能な画像データ数が限定されるので、数千枚程度の画像を用いて実証実験を行った。まず、画像選出機能を用いて画像中の顔の有無を判定した。画像を選出する際、3.3節で述べたように顔の写り具合により画像を3つのカテゴリーに分けた。次に、顔が写っていない画像、顔が写っているが十分に小さい画像といった画像から、自治体職員がオープンデータ化したい画像を選出した。なお、顔が大きく写っている画像はオープンデータに適さないデータと判定されるので、公開しないことになる。こうして選ばれた画像をOpenPhotoにアップロードし、画像匿名化機能を用いてサイト上で顔に対するモザイク処理を実施した。さらに、処理後の画像を対象として自動ラベリングを行った。本システムに従って画像オープンデータに取り組んだ後、感想と今後の要望について愛知県高浜市の担当部署の職員1名にインタビューし、フィードバックを受けた。インタビューに先立ち、本システムに関するアンケートと検討事項を用意し、同職員から回答を得た。検討事項は大きく技術的な実現可能性、効果とコストおよび具体性といった3つの事項に分ける。表2にアンケートの内容と同職員からの回答を示す。なお、5段階評価の基準がここ書きで示されていないものは全て1が「そう思わない」であり、5が「そう思う」を示すものである。各機能についてアンケートとインタビューの結果をもとに実証実験を説明し、考察する。

4.2 画像選定機能

愛知県高浜市と日進市より画像データの提供を受けた。これらの画像に対して画像選出機能を行った。表3に実験結果を示した。本機能にてオープンデータ化できそうな画像として「顔なし画像」または「顔あり（小）画像」を抽出した。結果より、自治体職員が確認する必要がある画像数が日進市では3分の1近く、高浜市では4分の1近くに削減されたことがわかる。高浜市職員は「選出された画像の枚数（割合）は適切だったと思いますか」の質問には3（適切であった）と回答した。また「選出

表 2：アンケートの内容と回答結果

システム機能	アンケートの内容	回答
画像選出機能について	Q1. 選出された画像の枚数（割合）は適切だったと思いますか (1. 少なすぎる ← 3. 適切であった → 5. 多すぎる)	3
	Q2. 選出された画像から、オープンデータ化する画像を選ぶのは大変でしたか (1. 大変だった ← 5. 大変ではなかった)	5
	Q3. 選出された画像の多くは、オープンデータ化に適切だったと思いますか	3
画像匿名機能について	Q4. モザイク処理機能は使いやすかったですか(1. 使いにくかった ← 5. 使いやすかった)	2
	Q5. モザイク処理機能によってオープンデータ化の作業が楽になったと思いますか	2
	Q6. 顔検出の精度は良かったと思いますか	4
	Q7. モザイク処理機能によって、(量・種類ともに) より多くの写真がオープンデータ化可能になると思いますか	3
実証実験全体について	Q8. これらのシステムによって、画像オープンデータに取り組む上での作業が楽になったと思いましたか	4
	Q9. これらのシステムによって、より積極的に画像オープンデータに取り組みたいと思いましたか	4
	Q10. このシステムを、他の自治体にもおすすめしたいと思いますか	4
	Q11. このシステムを継続的・定期的に使って、画像オープンデータに取り組みたいと思いますか	4
取り組み全体について	Q12. この取り組みを通して、AI の活用についての理解が深まりましたか	4
	Q13. この取り組みを通して、自治体業務への AI 活用の意義・効果を実感できましたか	4
	Q14. この取り組みを通して、自治体の他の業務への AI 活用のイメージが湧きましたか	4

された画像から、オープンデータ化する画像を選ぶのは大変でしたか」の質問には5（大変ではなかった）と回答していることから、選出された画像は多すぎでも少なすぎでもなく適切な枚数であったと言える。検出の結果について目視で確認したところ、概ねクラスごとに画像を分けることができた。この中で後ろ向きや斜め後ろ向きの人物について検出対象から外すことができた。ただし、顔が明確に写りこんでいるにもかかわらず「顔なし画像」のカテゴリーに分けられた画像が散見された。なお、このような検出漏れについて本システムの目的からして、検出漏れの割合がそれほど多くない限りは問題ない。これに対し、顔が写っていないにもかかわらず選出対象から除外するという失敗を減らすことが重要である。検出結果ではこのような誤検出がほとんど見られなかった。誤検出の画像例として人形の顔を人の顔として誤認識したケースがあった。これ以外にわずかではあるが、鉄骨の繋ぎ目や打ち上げ花火を顔として誤検出したものがあった。

表 3：顔検出による画像選出の結果

	顔なし	顔あり (小)	顔あり (大)	合計
高浜市	365	588	1,728	2,681
日進市	377	356	1,975	2,708

4.3 画像匿名化機能

愛知県高浜市にて画像選出機能より「顔なし画像」、「顔あり (小) 画像」と判別した画像から 98 枚を OpenPhoto にアップロードした。この中で画像から検出された顔に

対してモザイクをかける画像匿名機能を利用した。実行結果について高浜市より以下のフィードバックを得た。モザイク処理機能の使いやすさ、オープンデータの作業への改善という質問に対して、やや使いにくかったという回答があった。その理由としては処理に時間がかかり、検出漏れに対応できないという点が挙げられた。処理に時間がかかることについて本機能はクラウド上で実行されており、パソコンのスペックに依存しないこと、そして個人のパソコンからサイトにアクセスして 5～10 秒程度で実行できたことから考えれば、市役所のネットワーク環境にてデータの送受信に時間がかかっていたと考えられる。そこで、事前に画像データをアップロードしておくなどの工夫で対応可能だと考える。また、検出漏れに対応できない点について、顔検出の精度は良かったと評価されたものの、検出漏れに対応する方法を用意しなかったことより、使いにくい要因になったと考えられる。そこで、顔検出のための閾値を下げるか、検出漏れがあった時にユーザーから直接モザイクをかけることで対応可能である。今回の実証実験を通じて、画像匿名化機能の使い勝手が向上すれば積極的に使いたいという意見を受けた。行政業務においてオープンデータ推進のみならず、モザイクによる匿名機能の需要が高い。例えば、高浜市で広報活動を行う際、名札や車のナンバープレートにモザイクをかける必要があり、本システムの画像匿名化機能をより広く使えるようにすれば、多くの業務に適用可能だという意見を受けた。

4.4 画像の自動ラベリング機能

自治体職員の協力により、画像からの自動ラベリング

表4：自動ラベリングの実験結果

ラベル付与枚数	ラベル名	ラベル付与枚数	ラベル名
7	Room	5	Lighting
	Ceiling		Table
	Workwear		Human body
6	People in nature		Furniture
	Soil		Community
	Wetland		Public space
5	Flooring	5	Publication
	Motor vehicle		

についても実験を行った。高浜市より選出された98枚画像を対象に自動ラベリングを行った。表4に自動ラベリングの実験結果を示す。ここでは5枚以上のラベルを示してある。各画像に付与されたラベルを多い順で並べた。抽出結果を見れば、RoomやCeilingといった一般的によく見られるラベルもあれば、画像特有の情報を表現しているラベルもあったことがわかる。例えば、People in nature, Soil, Wetlandなどのラベルが付与された画像では、田植え、芝刈りおよび畑仕事に関する農業系の写真が見られた。Communityのラベルが付与された画像はイベントやお祭りの写真が中心であった。実験では顔にモザイク処理を行った後の画像からラベリングを行ったにもかかわらず、人の存在に依存するCommunityを正しく抽出することができた。Public spaceのラベルが付与された画像はすべて公園の写真であった。Publicationのラベルが付与された画像は何らかの申請書類、チラシ、リーフレットなどを撮影したものである。今回の分析対象ではイベントの様子を撮影した写真が最も多かった。ラベルの付与結果から見れば、イベントの様子を比較的適切に捉えることができた。一方で、グッズや工芸品など個別性が高い物体のラベルを抽出することができなかった。

本機能を活用し、画像にラベルを付与することで、データのクラスタリングの自動化に繋がると考える。これにより、横断的に画像オープンデータを探すことが可能となり、自治体におけるデータ利活用を促進することが期待される。

4.5 提案システム全体

システム全体について高浜市より肯定的な評価を受けた。本システムを活用することにより、画像オープンデータ化の作業を効率化し、モチベーションの向上に繋がると考える。加えて、アンケートの間12～14の回答結果を見れば、本取り組みを通じて自治体における人工知能技術の更なる利活用にも繋がると考える。なお、提案システムにおいてユーザビリティ等の面では引き続き改善していく余地がある。ユーザビリティの観点から、システムの検出漏れに自治体職員が対応できる方法を実装する必要がある。また、市役所のパソコンから外部のクラウドサービスを利用するには時間がかかったため、事前

のデータ処理や自治体のネットワーク環境に即した実装が必要である。自治体の業務において本システムの各機能を活用することにより、自治体業務における業務効率化、用途ごとの画像保存およびデータ管理に役立つものとする。

5. おわりに

本研究では自治体における画像オープンデータ化に取り組むための支援システムを提案した。提案の有効性、実現可能性および課題点を明らかにすべく、自治体職員の協力のもとで実証実験を行った。本研究を通じて画像オープンデータの推進に地方自治体のデータとICTを活用した実践を示した。対象とする市町村が変わってもシステムに必要な要件や対象となるデータがほとんど変わらないので、本研究で考案した方法やシステムを他の自治体に横展開することが容易だと考える。従って、本システムを用いて自治体業務を支援することにより、画像オープンデータ推進に繋げることができたと考える。

全体として実現可能性はそれほど問題ないものの、ユーザビリティの向上と運用方法の検討が課題として残っている。今後、産官学連携のもとで本システムをよりよく改善し、自治体の現状に即した実装やシステムの推論ミスに対応する機能設定が必要である。運用の観点から考えれば、本システムの利用法、活用対象および使用のタイミングについて、より具体的にイメージする必要があると考える。加えて、画像オープンデータの取り組みにおいてデータの公開のみならず、活用したい画像の種類など市民や企業の要望も把握して反映することが重要である。

謝辞

本研究推進にあたり、インフォ・ラウンジ株式会社の肥田野様、伊藤様、下山様、小林様、そして愛知県高浜市の京極様、愛知県日進市の藤原様にお世話になりました。その他7市町オープンデータ会議でご協力頂いた職員の方々、および高浜市での実証実験に協力頂いた職員の方々にも感謝申し上げます。NECソリューションイノベータの深谷様、岡元様、篠田様には毎月の定例会議で有益なご助言を頂きました。関係者の皆様に感謝いたします。なお、本研究の一部はJSPS 科研費18H03493、公益財団法人 大幸財団の助成を受けたものです。

引用文献

- 愛知県東浦町 (2021). 東浦町画像オープンデータ. <http://open-imagedata.town.aichi-higashiura.lg.jp/>.
- 愛知県日進市. 日進市オープンデータポータルサイト. オープンデータ活用事例 (2020). 防災啓発マップ. http://www.city.nisshin.lg.jp/opendata/od_jirei.html.
- Bochkovski, A., Wang, C., and Liao, H. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934.
- 画像オープンデータガイドライン (2021). <https://>

ayane0321.github.io/image-opendata-guide/.

Google Cloud Vision API (2021). <https://cloud.google.com/vision>.

北海道室蘭市 (2021). むろらんオープンデータライブラリ. <https://library-muroran.opendata.arcgis.com/>.

石川県金沢市 (2021). 金沢市画像オープンデータ. <https://open-imagedata.city.kanazawa.ishikawa.jp/>.

郡山市観光協会 (2021). ムービーアーカイブサイト. <https://www.kanko-koriyama.gr.jp/moviearchives/>.

高度情報通信ネットワーク社会推進戦略本部 (2012). 電子行政オープンデータ戦略. https://www.kantei.go.jp/jp/singi/it2/pdf/120704_siryou2.pdf.

高度情報通信ネットワーク社会推進戦略本部 (2019). オープンデータ基本指針. https://cio.go.jp/sites/default/files/uploads/documents/data_shishin.pdf.

クリエイティブ・コモンズ・ジャパン (2021). クリエイティブ・コモンズ・ライセンスとは. <https://creativecommons.jp/licenses/>.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, Vol. 521 (7553), 436-444.

内閣府 (2016). 第5期科学技術基本計画. <https://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf>.

内閣府 (2021). 科学技術・イノベーション基本計画. <https://www8.cao.go.jp/cstp/kihonkeikaku/6honbun.pdf>.

内閣官房情報通信技術 (IT) 総合戦略室 (2019). 地方におけるオープンデータの取組状況について. https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/opendata_wg_dai9/siryou2-1.pdf.

内閣官房情報通信技術 (IT) 総合戦略室. 推奨データセット (2021). <https://cio.go.jp/policy-opendata>.

OpenPhoto オープンフォト (2021). <https://openphoto.jp/>.

新型コロナウイルス対策ダッシュボード (2021). <https://www.stopcovid19.jp/>.

総務省 (2016). 平成28年版情報通信白書, 人工知能 (AI) 研究の歴史. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/html/nc142120.html>.

総務省 (2019a). 地方自治体における新たな技術の活用状況について. https://www.soumu.go.jp/main_content/000601804.pdf.

総務省 (2019b). 地方公共団体におけるデータ利活用ガイドブック ver.2.0 概要. https://www.soumu.go.jp/main_content/000620311.pdf.

総務省 (2019c). オープンデータ化した地域資料の活用を通じて大阪の魅力を発信. https://www.soumu.go.jp/main_sosiki/joho_tsusin/top/local_support/ict/jirei/2019_005.html.

総務省 (2021). 令和3年版情報通信白書. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/pdf/01honpen.pdf>.

総務省 (2021). データカタログサイト. <https://www.data.go.jp/>.

鈴木彩音・浦田真由・遠藤守・安田孝美 (2019). 自治体

における観光振興のための画像公開方法の検討. 観光情報学会第20回研究発表会, 25-28.

鈴木彩音・浦田真由・遠藤守・安田孝美 (2020). 自治体における画像オープンデータ標準化に向けた取り組み.

第11回社会情報学会中部支部研究会論文集, 41-44.

東京都世田谷区 (2021). 世田谷区 GIS オープンデータ. <https://data-setagaya.opendata.arcgis.com/>.

Abstract

In recent years, digitalization has made key progresses in every corner of modern society, and the utilization of data and information and communication technology has become increasingly important in various fields including government. In the administrative field, it is critical to utilize information and communication technology and formulate policies based on data to respond to administrative management and social issues in local governments. Therefore, a prominent need rises for promoting the utilization of data especially open data in local governments. In this study, we build a support system that utilizes artificial intelligence technology through industry-government-academia collaboration toward the practice of promoting image open data in local governments. Currently, local governments nationwide are promoting open data in text and tabular format, but due to issues such as portrait right and privacy, it is not common to disclose images as open data in local governments. This makes it difficult to promote and utilize image open data. Therefore, in this study, we have developed a support system to promote the disclosure of images become open data. In addition, the support system was verified through demonstration experiments in local governments through industry-government-academia collaboration. As a result, we were able to confirm the feasibility and effectiveness of disclosure images owned by local governments into open data. We consider that this will lead to the utilization of image open data. Also, we reduce the work burden related to image open data in local governments and improve work efficiency. Furthermore, through this study, we aim to verify the effect and significance of using artificial intelligence technology for local government works, and to deepen the understanding and image of the use of artificial intelligence technology.

(受稿: 2021年11月11日 受理: 2021年12月27日)