

来遊者に興味を与える地域特性を発見するための
Computer Vision APIに基づく SNS 投稿画像分析手法の検討
Analytical Method of SNS's Images Based on Computer Vision API
to Discover Geographical Characteristics

橋本幸二郎[†] 三代沢 正[†] 宮部真衣[†] 土屋健[†] 尾崎健[†] 広瀬啓雄[†]
Kohjiro Hashimoto Tadashi Miyosawa Mai Miyabe Takeshi Tsuchiya Ozaki Takeshi Hiroo Hirose

1. 背景

人口減少問題を抱える日本では、特に地域において産業・雇用の創出などの様々な課題に直面している。この課題の解決には地域の活性化が必要であり、地域活性化を考える上で、重要な要素になりうるのが観光である。政府は、2018 年観光立国推進戦略会議にて、2020 年外国人旅行者数 4000 万人を目指す目標を掲げた[1]。図 1 に外国旅行者の訪日者数の推移を示す[2]。訪日者数は年々伸びており、2017 年の時点で 3000 万人に近い。一方、図 2 に長野県茅野市の観光客数の推移を示す[3]。茅野市は白樺湖や車山高原等の観光地が多数存在するにも関わらず、観光客数は伸び悩んでいることが確認できる。

この問題に対して、市は観光客数増加のための対策を講じようとする。しかし、講じる対策で利用対象とする地域の特徴と実際の来遊者が興味を持つ地域の特徴にミスマッチが起きる場合には、その対策が効果的に働かない。このようなミスマッチを回避するためにも、また、対策の改善を促すためにも、来遊者目線での地域特性の把握が重要となる。

一方、ソーシャルネットワーク(SNS)上には、来遊者が旅行や出張時に来遊した場所の印象や体験が言葉や画像といった形式で掲載されている。それ故、SNS 上の情報は、来遊者の行動分析や観光スポット分析に効果的である。SNS 上の情報を分析する事例としては、テキストマイニングのオープンソースを用いてブログサイトから出現頻度の高い語を抽出し、旅行先に関する特徴を分析する例[4]、写真共有サイト flicker に投稿された画像群に紐付いた位置(緯度・経度)情報を基に、興味を持って撮影した場所 POI(Point of Interest)のヒートマップを作成する例[5]、Wikipedia や Yahoo 知恵袋等の様々な WEB からの情報源を用いて観光地の特徴ベクトルを生成する例[6]、SNS 上の画像群をカテゴリ分類することでユーザの嗜好を可視化する例[7]等がある。位置情報の利用は、観光客の行動分析や観光経路分析に用いられ、交通事情やバリアフリー改善に繋がる[8][9]。一方、投稿画像には撮影者の興味対象が映っていると捉えることができる。このとき、地域毎に投稿画像を収集し、撮影者の興味対象を分析することにより、来遊者に興味を与える地域特性を分析することが可能となる。

[†] 公立諏訪東京理科大学 工学部情報応用工学科,
Department of Applied Information Engineering, Suwa
University of Science

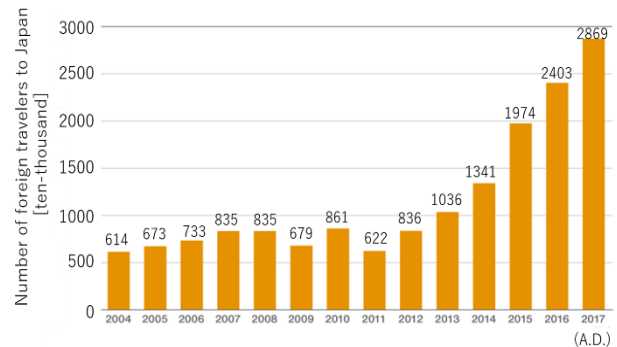


図1 訪日外国人観光客数の推移[2]

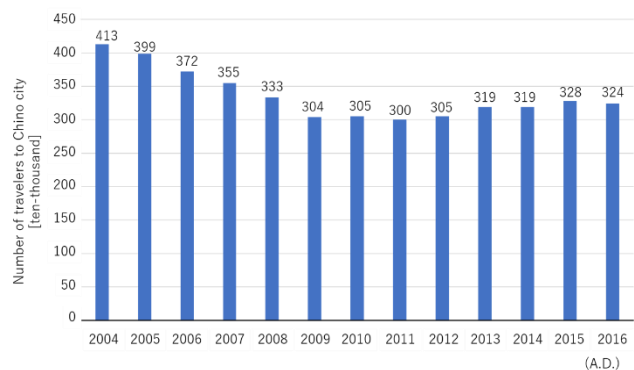


図2 茅野市への観光客数の推移[3]

SNS 上に投稿された画像情報に対して、既存の画像認識技術を駆使し、分析する例がある[10]。しかし、従来の画像解析技術は解析したい対象の画像特徴量を設計者が定める必要がある。興味対象候補は膨大に考えられ、その特徴量を事前に設計することは困難であり、分析結果も設計者の基準に依存してしまうことなど、膨大な SNS 投稿画像を解析する上での問題が存在する。一方、Google や Microsoft などは 2016 年にクラウド上の画像解析サービスを公開した。このサービスは深層学習技術を用いており、人工知能により自動的に決定される特徴量に基づく画像解析技術となっている。それ故、一般物体認識の汎用性は高く、上記に掲げた問題を解決しうる画像分析方法と考えられる。中でも、本研究では、Microsoft が提供する Computer Vision API(以降、CVA と記述)を利用した画像

自動タグ付け機能に着目する。画像自動タグ付け機能とは、入力された画像に対して領域分割を行い、各領域に対する名称をタグとして自動付加する機能である。すなわち、画像に何が映っているかをタグレベルで言語化し抽出してくれる機能と言える。このとき、任意の地域の SNS 投稿画像を収集、タグ化し、そのタグ情報を統計分析することで、来遊者に興味を与える地域特性をタグ情報として抽出できると考えられる。

本論文では、SNS 投稿画像群を CVA に基づきタグ化し、そのタグ情報を分析する手法を検討する。具体的には、タグの出現頻度に基づき来遊者の興味度合を定量的評価する頻度ベースの分析手法と他地域の出現タグの差異に基づき来遊者の興味度合を定量的評価する比較ベースの分析手法を提案し、その有効性を検証する。

2. 写真共有サイト flicker

flicker は写真共有サイトであり、世界中で利用されている。投稿写真には、写真 id、撮影日、撮影された緯度経度情報、投稿者 id、タイトル等の情報が紐づけられており、flicker API を利用したプログラムを作成することにより、上記のメタ情報を含む投稿画像を収集することができる。本論文では、この flicker に投稿された画像群を利用する。具体的には、分析対象地域を、長野県の諏訪市、松本市、茅野市の 3 つに限定し、投稿画像の緯度経度情報に基づき各地域の投稿画像群を収集する。こうして収集した画像群に対して次章で述べる CVA の画像解析機能を適用することで、画像をタグ情報に変換する。

3. CVA に基づく画像自動タグ付け機能

Microsoft が提供する CVA を利用し、プログラムを作成することで、図 3 に示すような画像解析が可能となる。図 3(a)は入力画像であり、図 3(b)が分析結果である。ここでは、画像自動タグ付け機能を実装しており、入力画像中に写っている対象がタグ情報として自動付加される。図 3(b)において、“name”の後方に表記される“grass”や“outdoor”がタグに相当する。なお、CVA が付加できるタグの総数は不明であり、事前検証により少なくとも 2000 個以上のタグが存在することを確認している。一方、“confidence”の後方に表記される値はタグ付加に対する確信度を表しており、[0,1]の範囲で表記される。

図 3 に示すように、CVA は画像情報をタグ情報という言葉レベルに変換する機能と言える。本研究では、各地域の画像群をこの機能によりタグ化し、次章で述べる分析手法に基づき来遊者に興味を与える地域特性を分析する。



(a) 入力画像

```
タグ [ { "name": "grass", "confidence": 0.9999968 }, { "name": "outdoor", "confidence": 0.9998648 }, { "name": "sky", "confidence": 0.9996803 }, { "name": "path", "confidence": 0.649573743 }, { "name": "dirt", "confidence": 0.475375861 }, { "name": "traveling", "confidence": 0.3048256 }, { "name": "railroad", "confidence": 0.213923573 } ]
```

(b) 解析結果(タグ情報+確信度)

図 3 CVA プログラムの実行例

4. 興味対象分析手法

4.1 頻度ベース分析手法

本手法では、前節で述べた CVA に基づき付加される各タグに対し、来遊者の興味度合を定量的に評価する。ここでは、タグに対する興味度合を、出現タグの頻度で定義する。ただし、CVA が付加するタグには不確実性が存在し、それが確信度(Confidence)として出力される。それ故、ここでは確信度の総和をタグに対する興味度合と定義する。具体的に定式化すると以下になる。flicker により得られた地域 s ($1 \leq s \leq S$) の画像群を $I_s = \{img_1, img_2, \dots, img_n, \dots, img_{N_s}\}$ ($1 \leq n \leq N_s$) とする。ここで、 S は分析する地域数であり、 N_s は地域 s から収集した画像数を表す。CVA が所有するタグが M 個するとし、画像 img_n に CVA を適用し得られるタグ m に対する確信度を C_{mn} としたとき、地域 s のタグ m に対する興味度合 E_{ms} を下式で定義する。

$$E_{ms} = \frac{\sum_{n=1}^{N_s} C_{mn}}{N_s} \quad \dots (1)$$

この興味度合は値が大きい程、タグ m は地域 s において来遊者に興味を与える対象と捉える。

4.2 比較ベース分析手法

前節で定義した興味度合は、頻度を基準としている。一方、他の地域では見ることができないような対象が存在すれば、その対象に対する興味度合は高いと考えられる。そこで本節では、他の地域が持つタグ情報との比較に基づく興味度合を定義する。この興味度合の定義には、ロジステ

イック回帰モデルを用いる。

ロジスティック回帰モデルは統計的回帰モデルの一種であり、2 種のクラスタリング問題及び入出力の相関性を評価する際に用いられる手法である。概要を以下に述べる。ある現象の発生する確率を P とする。その現象の発生を説明するために観測された変数群を $\mathbf{x}=(x_1, x_2, \dots, x_r)$ とする。この $\mathbf{x}$ は一般に説明変数と呼ばれる。このとき、 $\mathbf{x}=(x_1, x_2, \dots, x_r)$ の状態において、現象が発生する確率 $P(\mathbf{x})$ を下式で表す。

$$\begin{aligned} P(\mathbf{x}) &= \Pr\{\text{現象の発生} \mid x_1, x_2, \dots, x_r\} \\ &= F(x_1, x_2, \dots, x_r) \end{aligned} \quad \dots (2)$$

$\Pr$ は確率を表しており、 P は実質、説明変数 $\mathbf{x}=(x_1, x_2, \dots, x_r)$ の条件の下で現象が発生する条件付き確率を意味する。そして、この $\Pr$ を関数 $F(\cdot)$ とみなしてモデル化していく。ロジスティック回帰モデルでは、関数 $F(\cdot)$ として以下に示すロジスティック関数が用いられる。

$$p(\mathbf{x}) = F(Z) = \frac{\exp(Z)}{1 + \exp(Z)} \quad \dots (3)$$

ここで、 Z は下式で表す。

$$Z = \eta_0 + \eta_1 x_1 + \eta_2 x_2 + \dots + \eta_r x_r \quad \dots (4)$$

つまり、説明変数の重回帰式 Z と、出力範囲 $[0,1]$ の値を持つ発生確率 $p(\mathbf{x})$ をロジスティック関数でリンクされたモデルである。このモデルは、説明変数を入力として二値 (0, 1) を出力とする現象のモデル化に有効であり、主にクラスタリング手法として用いられる。一方、このモデルは予め学習データを用意し、説明変数の係数である $\boldsymbol{\eta}=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$ を学習させることで得られるが、この係数が出力と説明変数との相関性を表す指標となることから、データ分析時にもよく用いられる。そこで、本研究では、モデル学習によって得られる説明変数の係数を用いてタグに対する興味度合を定義する。具体的な本研究への適用方法を以下に述べる。

ロジスティック回帰モデルにおいて、学習データは説明変数 $\mathbf{x}=(x_1, x_2, \dots, x_r)$ と結果変数 $y \in \{0, 1\}$ のペアになる。ここで、説明変数を CVA により出力されるタグ情報に対する one-hot ベクトルの総和で表す。例えば、CVA が持つタグが "outdoor"、"car"、"sea"、"driver"、"train" の 5 種類 ($M=5$) であったとき、それぞれのタグを説明変数 $\mathbf{x}=(x_1, x_2, x_3, x_4, x_5)$ に対応させる。そして、ある画像を CVA に入力したときの出力タグが "car"、"driver" であった場合、説明変数を $\mathbf{x}=(0, 1, 0, 1, 0)$ とする。一方、結果変数については、分析したい地域の画像であれば $y=1$ 、そうでなければ $y=0$ とする。

例として、諏訪市を分析する場合を考える。このとき、各地域の画像群を収集し、学習データは以下のように集める。そして、学習することで諏訪市のロジスティック回帰モデルを生成する。

[学習データ]

画像 1(諏訪市): 結果変数 $y=1$ 、説明変数 $\mathbf{x}=(0, 1, 0, 1, 0)$
 画像 2(松本市): 結果変数 $y=0$ 、説明変数 $\mathbf{x}=(1, 1, 0, 0, 0)$
 画像 3(茅野市): 結果変数 $y=0$ 、説明変数 $\mathbf{x}=(0, 0, 0, 0, 1)$
 画像 4(諏訪市): 結果変数 $y=1$ 、説明変数 $\mathbf{x}=(1, 1, 0, 1, 0)$
 画像 5(茅野市): 結果変数 $y=0$ 、説明変数 $\mathbf{x}=(0, 0, 0, 1, 1)$

⋮

モデル学習において、説明変数の係数である $\boldsymbol{\eta}=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$ は、分析対象地域の画像の場合は結果変数を 1 に近づけるため値が大きくなり、分析対象外地域の画像の場合は結果変数を 0 に近づけるため値が小さくなる。すなわち、他の地域にはないようなタグが存在する場合には、そのタグの係数の値が大きくなる。そこで、本論文では、上記の要領で、諏訪市、松本市、茅野市のロジスティック回帰モデルを生成し、説明変数の係数である $\boldsymbol{\eta}=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$ を各タグに対する興味度合とみなす。なお、モデルの学習には最尤推定法を用いており、プログラム実装時には機械学習ライブラリの `skit-learn` を用いた。また、CVA におけるタグ数 M は不明であったため、諏訪市、松本市、茅野市の画像群を CVA に適用した際に出現したタグ数で代用した。

5. 評価実験

5.1 実験内容

今回分析する地域は、長野県の諏訪市、松本市、茅野市の 3 つに限定した。flicker より収集した各市の画像枚数は諏訪市 3,215 枚、松本市 10,283 枚、茅野市 3,638 枚であり、それぞれ 2010 年から 2017 年の画像データである。実験では、収集した画像群に CVA を適用し、タグ情報化した後、4 章で述べた頻度ベース及び比較ベース手法に基づき各タグに対する興味度合を算出した。算出した興味度合を地域毎で上順に並べ、各地域が持つ観光事情との照らし合わせにより、高い興味度合を持つタグの妥当性を評価した。なお、収集した画像群全てを CVA に適用した結果、得られるタグ数 M は 1821 個であった。

5.2 頻度ベース分析手法の分析結果

表 1、2、3 は頻度ベース手法に基づき算出された各タグに対する興味度合について、上位 5 位までを記したものである。表 1 は諏訪市、表 2 は松本市、表 3 は茅野市である。

結果より、どの市にも outdoor、sky、nature 等の同じタグが上位に来ていることがわかる。これは、CVA より付加されるタグに冗長なタグが存在するためである。例えば、outdoor タグは野外で撮影された画像に対しては必ず付加され、空が映っていれば必ず sky タグが付加されることを確認している。すなわち、興味のある対象物とは関係なく、必然的に移り込むような対象が存在するため、その

表 1. 頻度ベースに基づく興味度合算出結果(諏訪市)

順位	1	2	3	4	5
タグ	outdoor	sky	tree	grass	nature
興味度合	0.64	0.37	0.22	0.16	0.11

表 2. 頻度ベースに基づく興味度合算出結果(松本市)

順位	1	2	3	4	5
タグ	outdoor	sky	tree	nature	mountain
興味度合	0.71	0.35	0.32	0.24	0.20

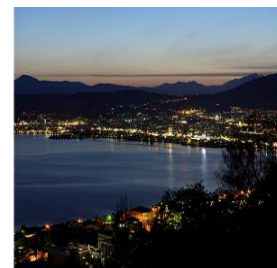
表 3. 頻度ベースに基づく興味度合算出結果(茅野市)

順位	1	2	3	4	5
タグ	outdoor	sky	nature	tree	mountain
興味度合	0.80	0.41	0.34	0.32	0.24

ようなタグは必然的に興味度合が高くなる。これは、CVA単体利用では、来遊者に興味を与える地域特性の分析に有効ではないことを意味しており、上記に挙げた冗長タグを評価しないような別処理が必要になる。

また、この頻度ベースの結果に対して、各市で代表的な観光対象物がどの順位に相当しているかを検証した。例えば、松本市では、松本城が有名であり、実際に収集した画像の中には図 4(a)に示すように松本城の写真が多く含まれていることを確認している。また、諏訪市では、諏訪湖が有名であり、実際に収集した画像の中には図 4(b)に示すように諏訪湖の写真が多く含まれていることを確認している。しかしながら、表 1、2、3 の結果から有名観光地に対応するタグは上位にない。この原因は CVA により付加されるタグに類似した意味のタグが含まれているためである。図 5、6 にその一例を示す。図 5、6 はともに松本城の画像であるが、付加されるタグは、castle、church、shrine、pagoda 等の類似した意味のタグが付加されていることがわかる。このように異なる名称のタグ付けが行われることにより、興味度合が分散され、上位に上がらなかったものと考えられる。

以上の結果を踏まえ、目視により冗長なタグを省き、かつ、手作業で類似タグ名称を統一し、再度頻度ベースにより興味度合を評価した。表 4、5、6 にその結果を示す。それぞれ上位 10 位までを示しているが、outdoor、indoor、sky のタグは冗長タグと判断し、省いている。また、castle、church、shrine、pagoda、temple を castle に統一し、water、lake、sea を lake に統一している。表 4 より、諏訪市の代表的観光地である諏訪湖を表すであろう lake が上位に存在することがわかる。一方、表 5 より、松本市の代表的観光地である松本城を表すであろう castle が上位に存在することがわかる。表 6 より、茅野市では snow、mountain、hill 等の山に関するタグが上位に存在



(a) 松本市画像群に含まれる松本城画像 (b) 諏訪市画像群に含まれる諏訪湖画像

図 4 各市を代表する有名観光地画像



(a) 入力画像

```
タグ[{"name":"tree","confidence":0.99975},
{"name":"outdoor","confidence":0.99800},
{"name":"castle","confidence":0.99800},
{"name":"temple","confidence":0.48304},
{"name":"city","confidence":0.48086},
{"name":"street","confidence":0.32968},
{"name":"building","confidence":0.29790},
{"name":"church","confidence":0.26432},
{"name":"autumn","confidence":0.249164}]
```

(b) 解析結果(タグ情報+確信度)

図 5 松本城画像を CVA に適用した際の出力結果(その 1)



(a) 入力画像

```
タグ[{"name":"outdoor","confidence":0.99516},
{"name":"sky","confidence":0.99171},
{"name":"red","confidence":0.97319},
{"name":"wooden","confidence":0.81731},
{"name":"temple","confidence":0.81731},
{"name":"shrine","confidence":0.12361},
{"name":"pagoda","confidence":0.05187}]
```

(b) 解析結果(タグ情報+確信度)

図 6 松本城画像を CVA に適用した際の出力結果(その 2)

することがわかる。茅野市は登山やスキーが有名であり、実際に茅野市の収集画像を確認した所、山や登山の画像が多く存在した。それ故、この結果は妥当な結果と言える。以上より、頻度ベース手法に基づき、地域の特徴を表すタグが抽出されることが確認できた。ただし、冗長なタグの除外処理や類似タグの統合処理が必要となる点が課題と言える。一方、CVA は一般物体認識を前提としていることから、付加されるタグが抽象的である。例えば、表 6 の person や plant 等のタグは抽象であり、その情報だけでは、どのような画像が集まっているのか、地域の特徴がイメージしにくい。このことから、一つのタグから、来遊者に興味を与える地域特性を分析することは難しいと考える。ただし、より具体的な情報としてタグが得られるのであれば効果的と考える。例えば、人が映っている画像に対し

表 4. 頻度ベースに基づく興味度合再算出結果（諏訪市）

順位	タグ	興味度合	順位	タグ	興味度合
1	tree	0.22	6	plant	0.10
2	grass	0.16	7	outdoorobject	0.09
3	nature	0.11	8	fireworks	0.09
4	lake	0.10	9	mountain	0.09
5	building	0.10	10	person	0.07

表 5. 頻度ベースに基づく興味度合再算出結果（松本市）

順位	タグ	興味度合	順位	タグ	興味度合
1	tree	0.32	6	grass	0.10
2	nature	0.24	7	plant	0.09
3	mountain	0.20	8	building	0.08
4	person	0.18	9	castle	0.07
5	water	0.11	10	ground	0.06

表 6. 頻度ベースに基づく興味度合再算出結果（茅野市）

順位	タグ	興味度合	順位	タグ	興味度合
1	nature	0.33	6	plant	0.14
2	tree	0.32	7	hill	0.12
3	mountain	0.24	8	person	0.12
4	snow	0.23	9	slope	0.08
5	grass	0.15	10	forest	0.07

て、“子供が諏訪湖で釣りをしている”といった対象と動作、環境情報を含む画像全体を表現したタグとして付加されるのであれば、分析する上で解釈が容易となり、効果的と考える。

5.3 比較ベース分析手法の分析結果

次にロジスティック回帰モデルに基づき算出される興味度合を検証した。表 7、8、9 に結果を示す。それぞれ算出された興味度合の上位 10 位のタグを示している。

表 7 より、fireworks が最上位に存在するが、諏訪市は諏訪湖祭湖上花火大会と呼ばれる約 50 万人が来場する有名な花火大会がある。図 7(a)に示すような花火の画像は諏訪市のみに存在したため、興味度合が高く算出されたと言える。他方、shrine が上位に存在するが、これは諏訪市にある諏訪大社の画像が影響している。諏訪市では、御柱祭りと呼ばれる 7 年毎に開かれる奇祭があり、諏訪大社で行われる。実験で用いた画像群は 2010 年から 2017 年のものであり、御柱祭りの画像(図 7(b))が含まれていた。それ故、諏訪大社を表すであろう shrine が上位に上がったものとする。

表 7. 比較ベースに基づく興味度合算出結果（諏訪市）

順位	タグ	興味度合	順位	タグ	興味度合
1	fireworks	0.33	6	Sea	0.14
2	automotive	0.32	7	Feet	0.12
3	shrine	0.24	8	election	0.12
4	can	0.23	9	steam	0.08
5	box	0.15	10	festival	0.07

表 8. 比較ベースに基づく興味度合算出結果（松本市）

順位	タグ	興味度合	順位	タグ	興味度合
1	pagoda	0.33	6	cycling	0.14
2	camping	0.32	7	preschool	0.12
3	hiking	0.24	8	mailbox	0.12
4	herb	0.23	9	clothing	0.08
5	trek	0.15	10	temple	0.07

表 9. 比較ベースに基づく興味度合算出結果（茅野市）

順位	興味度合	関心度	順位	タグ	興味度合
1	snow	0.33	6	bunch	0.14
2	little	0.32	7	ramp	0.12
3	pulling	0.24	8	highway	0.12
4	table	0.23	9	standing	0.08
5	platform	0.15	10	jacket	0.07



(a) 諏訪市画像群に含まれる花火大会画像



(b) 諏訪市画像群に含まれる御柱画像

図 7 諏訪市画像群に含まれていた画像例



(a) 松本市画像群に含まれるキャンプ画像



(b) 松本市画像群に含まれるサイクリング画像

図 8 松本市画像群に含まれていた画像例



(a) 茅野市画像群に含まれる雪山登山画像



(b) 茅野市画像群に含まれる子供の画像

図 9 茅野市画像群に含まれていた画像例

表 8 より、pagoda が最上位に存在し、これは松本城を表していると想定できる。その他、camping や hiking のようなアクティビティのタグが上位に存在することがわかる。収集画像を確認した所、図 8 に示すような、キャンプやサイクリングの画像が多数含まれていた。故に妥当な結果と言え、松本城のような有名観光地だけでなく、アクティビティとしての特徴も抽出できていると言える。

表 9 より、茅野市の最上位は snow となっている。これは図 9(a)に示すような雪山での登山画像が多く含まれていたためと考えられる。また、図 9(b)に示すように子供の画像も多く含まれており、2 番目の little はそれを表すタグと言える。このことから、家族連れで遊べる場所が茅野市には存在することがわかる。しかし、画像だけではその場所を特定することができないため、写真を撮影した位置情報も含め分析することにより、より詳細な地域特徴を見出すことができると考えられる。

以上から、比較ベースでは、他の地域にはない特徴的なイベントを表すタグが上位に来る傾向にあることがわかる。また、アクティビティに関するタグが浮上したことも興味深い。ただし、表 7 に示すような Can や Box 等の解釈し難いタグも上位に浮上していることがわかる。これは、CVA のタグの冗長性、抽象性に起因していると考えられる。

6. まとめ

本論文では、Google や Microsoft が提供する深層学習に基づくクラウド画像解析サービスに着目し、そのサービスを用いた SNS 投稿画像分析の応用性を検証した。具体的には、Microsoft の Computer Vision API を用いた画像自動タグ付け機能を利用し、SNS 投稿画像をタグ情報化、統計分析することにより、来訪者に興味を与えるタグを推定する手法を提案した。提案手法は、タグの出現頻度に基づきタグに対する興味度合を定量化する頻度ベース手法と、ロジスティック回帰モデルの学習に基づき得られる説明変数の係数を興味度合として定量化する比較ベース手法を提案した。後者は、他の地域の画像タグ情報を含めて評価されるため、他の地域にはないタグを地域特徴とみなして、高い値が算出される手法になっている。

実験では、flicker より収集した、諏訪市、松本市、茅野市の画像群を用いて検証した。頻度ベース手法の検証結果から、CVA は冗長なタグが含まれること、そして、類似す

る意味を持つタグの存在が頻度を分散させ、正確に頻度の高い対象が見出すことができないことを確認した。また、CVA のタグ表記が抽象的なため、得られたタグに対して地域の特徴をイメージすることが難しいことが分かった。

一方、比較ベース手法の検証結果から、諏訪市の花火大会や松本市のキャンプ、茅野市の雪山登山等、他の地域にはない特徴的なタグが得られることを確認した。しかし、頻度ベースと同様、抽象的なタグの存在が解釈し難いタグを上位に浮上させることがわかり、地域の特徴をイメージすることが難しい問題は同様に存在することを確認した。

今後の課題として大きく 3 つ検討している。一つは CVA が持つ冗長なタグを除外する処理、類似する意味を持つタグを統一する処理を自動で行う方法を検討することである。CVA は今後も更新、拡張されることが想定されるため、有効に利用するためにも後処理となる方法を確立することは重要と考える。二つ目は、より詳細なタグを付加する方法を検討する。例えば、“子供が諏訪湖で釣りをしている”といった対象と動作、環境情報を含む画像全体を表現したタグとして付加することができれば、分析する上で解釈が容易とる。幸い、深層学習技術には画像から文章に変換する img2txt と呼ばれる手法が提案されており、この手法を用いた分析手法を検討する。三つ目に、画像以外の情報も含め分析する手法を検討する。画像だけでは解釈に限界があると考え、位置情報やつぶやき情報などを統合的に評価し分析する手法が必要と考える。

参考文献

- [1] 国土交通省 観光庁、“観光ビジョン実現プログラム 2018”、URL:https://www.mlit.go.jp/kankochou/news02_000354.html。
- [2] JTB 総合研究所、“インバウンド 訪日外国人動向”、URL:<https://www.tourism.jp/tourism-database/stats/inbound/>。
- [3] 長野県庁、“平成 29 年観光地利用者統計調査結果”、URL:<https://www.pref.nagano.lg.jp/>。
- [4] 村上嘉代子、川村秀憲、“外国人から見た日本旅行-英語ブログからの観光イメージ分析-”、人工知能学会誌、Vol.26、No.3 pp.286-293、2011。
- [5] 大久保航太郎、望月崇廣、渡邊貴之、“複数の位置情報付きソーシャルメディアから得られた POI の比較”、第 14 回情報科学技術フォーラム FIT2015、O-025、pp.481-482、2015。
- [6] 上原尚、嶋田和孝、遠藤勉、“Web 上に混在する観光情報を活用した観光地推薦システム”、電子情報通信学会、技術研究報告、Vol.112、No.367、pp.13-18、2012。
- [7] 石野淳一、中田洋平、日吉久礎、“SNS 上の画像群からのユーザー嗜好の抽出と観光広告への応用”、情報処理学会第 76 回全国大会、6T-5、pp.587-588、2014。
- [8] 鈴木恵二、“観光動向データの取得と分析-はこだてまちナビとレンタカー観光の GPS ログ解析-インバウンド 訪日外国人動向”、人工知能学会誌、No.26、Vol.3、pp.248-255、2011。
- [9] 三代沢正、広瀬啓雄、土屋健、尾崎剛、倉田紀子、“大規模イベントにおける位置情報型 AR の活用実験とアクセスログによる利用者の複合的行動分析手法の提案”、情報社会学会誌、Vol.12、No.1、pp.19-32、2017。
- [10] 王佳な、野田雅文、高橋友和、出口大輔、井出一郎、村瀬洋、“Web 上の大量の写真に対する画像分類による観光マップの作成”、情報処理学会論文誌、Vol.52、No.12、pp.3588-3592、2011。