

気象状況に応じた屋外画像生成

恒岡 知生[†] 新田 直子[†] 中村 和晃[†] 馬場口 登[†]

[†] 大阪大学大学院工学研究科 〒565-0871 大阪府吹田市山田丘 2-1

E-mail: †tsuneoka@nanase.comm.eng.osaka-u.ac.jp,

††{naoko,k-nakamura,babaguchi}@comm.eng.osaka-u.ac.jp

あらまし 現在、世界各地に気象センサが設置されており、任意地点の気象状況を表す情報として、温度や湿度などの信号値を容易に知ることができる。一方、より人間にとって理解しやすい情報として、空の画像など、各地のユーザからの投稿を提供するサービスも出現している。しかし現状では、提供される画像は非常に限定された地点、時間のものとなっている。そこで本研究では、任意地点の画像を、現在の気象状況に応じて変換することにより、任意地点、時間における屋外画像の生成を目指す。提案手法では、時間及び緯度経度情報付きの画像が多数投稿されるマイクロブログから屋外画像、気象センサから対応する時間、位置の信号値を対として収集する。収集したデータを用いて、顔画像に対する年齢など、与えられた画像に対し属性を変更する Invertible Conditional GAN を学習することにより、気象状況に応じた屋外画像を生成する。

キーワード マイクロブログ、画像生成、気象状況、Invertible conditional GAN

1. はじめに

近年、世界各地に設置された気象センサによって、世界中の気象状況の観測が容易になった。観測された気象状況は、Openweathermap [1] 等のインターネット上のサービスにより公開され、任意地点、時間の気象状況を温度や湿度等の信号値で取得できる。しかし、気象センサによって観測される情報は信号値であるため、ユーザが知りたい地点、時間の気象状況は直感的には理解しにくい。一方、ウェザーリポート [2] などでは、人間にとってより理解しやすい情報として、各地のユーザから投稿された屋外画像等を提供している。しかしこのような屋外画像は、ユーザが自発的に投稿するため、非常に限定された地点、時間の情報に留まる。

このように、気象センサのような機械センサの他に、各地の人間を実世界を観測するセンサとみなし、それぞれの利点、欠点を補完するよう利用する研究が進められている。この際、ウェザーリポートのように多数のユーザが利用するソーシャルネットワークサービス（SNS）が人間による実世界観測情報の情報源となることが多いため、各ユーザはソーシャルセンサと呼ばれる。特に、実世界の多様な状況について人間が観測した結果を、ジオタグと呼ばれる投稿位置の緯度・経度情報と共にリアルタイムに投稿すると考えられるマイクロブログを利用した研究が多い。例えば、Wang ら [3] は、機械センサである監視カメラにより特定エリアを常時観測することにより、異常状態を検出し、カメラの設置地点、時間付近におけるマイクロブログへの投稿から、パレード等、発生した状況に関する意味内容を抽出した。Giridhar ら [4] は、高速道路に設置された速度計を用いて車の速度を常時観測することにより、渋滞等の異常状態を検出し、速度計の設置地点、時間に投稿された「traffic」というキーワードを含むマイクロブログへの投稿から、異常状態の発生原因を特定した。これらの研究では、特定のエリアを

常時観測する機械センサは異常状態の検出に有用であること、また、マイクロブログには実世界で異常状態が発生した際、影響を受けた多数のユーザから異常状態に関する観測情報が投稿されやすいことを利用している。

一方、任意地点、任意時間の気象状況を伝える屋外画像のような通常状態における観測情報は、Twitter [5] のような非常に多数のユーザにより利用されるマイクロブログにも投稿されないため、取得できない。そこで本研究では、ソーシャルセンサから実際に得られる屋外画像を事例とし、任意地点、任意時間の気象状況を表す屋外画像の生成を試みる。近年、画像生成の分野において、多数の実画像を学習データとして用いることにより、実画像と見分けがつかないような偽画像を生成する敵対的生成ネットワーク (Generative Adversarial Networks:GAN) [6] が注目されている。特に、GAN を拡張した可逆的条件付き GAN(ICGAN) [7] では、学習データとして、例えば顔画像と年齢など、実画像と対となる条件ベクトルを大量に用いることにより、入力された実画像に対して条件ベクトルに応じた変更を行った偽画像を生成することができる。

よって本研究では、マイクロブログから得られる屋外画像に対し、同地点、同時刻における気象センサにより観測された信号値を条件ベクトルとして大量に収集する。これを ICGANs の学習データに用いることにより、任意地点の過去の屋外画像を、該当地点の現在の気象センサから得られる信号値に応じて変更し、任意地点、時間における気象状況に応じた屋外画像の生成を実現する。

2. 提案手法

気象センサ i ($i = 1, 2, \dots$) は、観測地点 pg_i において、時間 pt_i^j ($i = 1, 2, \dots, j = 1, 2, \dots$) の温度、雨量等、 F 種類の信号値 $M_i^j = (m_{i,1}^j, \dots, m_{i,F}^j)$ を取得する。本研究では、任意の地点における気象センサから得られる信号値が与えられたとき、観測

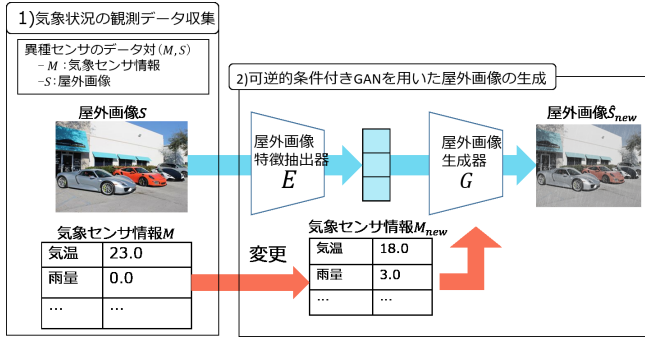


図1 提案手法の概要

された気象状況，地点に対して適切な屋外画像を得ることを目的とする。

人々が自身で観測した現在の情報をテキストや画像で気軽に投稿することができるマイクロブログには，各地の人々が観測した屋外状況を示す画像が，ジオタグと共に投稿される．そこでまず，マイクロブログへ投稿された屋外画像 S_n に対して，投稿地点 $sg_n (n = 1, 2, \dots)$ ，投稿時間 st_n に最も近接する気象センサから取得された信号値 M_i^j を収集し，信号値に応じた屋外画像生成モデルを学習する．学習した屋外画像生成モデルに画像を入力することで，入力画像を現在の気象センサからの信号値 M_{new} に応じて変更した屋外画像 $\hat{S}_{new}$ の生成が可能となる．

提案手法は図1に示すように，以下の2ステップにより構成される．

Step1) 気象状況の観測データ収集

マイクロブログへのジオタグ付き投稿に付与された画像から，人間が気象状況を理解しやすい情報として，屋外画像を収集する．さらに，収集した屋外画像の投稿地点，時間に最も近い地点，時間で観測した気象センサが取得した気象状況の信号値を収集する．

Step2) 可逆的条件付き GAN を用いた屋外画像生成

Step1) において収集した気象状況の観測データを用いて，可逆的条件付き GAN(ICGAN) [7] を用いた屋外画像生成モデルを構成する屋外画像生成器と屋外画像特徴抽出器を学習する．

Step2-1) 屋外画像生成器の学習

条件ベクトルとなる気象状況の信号値と乱数を入力して信号値に応じた画像を生成するよう屋外画像生成器を学習する．

Step2-2) 屋外画像特徴抽出器の学習

Step2-1) で学習した生成器に気象状況の信号値と乱数を入力して画像を生成し，入力した乱数を推定するよう特徴抽出器を学習する．

Step2-3) 気象状況を考慮した屋外画像生成モデルの構築

学習した屋外画像生成器と屋外画像特徴抽出器から，ICGAN を用いた屋外画像生成モデルを構築する．構築したモデルに画像を入力することで，入力画像を現在の気象センサからの信号値 M_{new} に対応した屋外画像 $\hat{S}_{new}$ の生成を実現する．

2.1 気象状況の観測データ収集

気象状況を表す屋外画像とそれに対応する気象状況の信号値を収集する．マイクロブログに投稿された画像から屋外画像を

表1 Twitter への画像付き投稿例

投稿時間	投稿位置	テキスト	投稿画像
2016/03/05 15:28:42	44.6, -124.5	YES! It is raining! I LOVE THE RAIN! #ElNino #luviophile #SoCal https://t.co/cf1xNrDmN7	

抽出し，それに対応する気象状況の信号値をインターネット上のサービスから取得する．

2.1.1 屋外画像の収集

代表的なマイクロブログの1つである Twitter [5] には，3億人を超えるユーザから1日5億以上のテキスト形式の投稿が公開されており，投稿時間の他に投稿位置の情報や，画像が付与されたものも存在する．本研究ではこのうち，人間が気象状況を理解しやすい情報として屋外画像 S_n を収集する．

Twitter への投稿に画像が付与されている場合，本文中に画像や外部サイトへのリンクが付与される．このようなリンク先の画像は多種多様な内容であるため，気象状況を確認できる屋外画像の投稿のみを収集する必要がある．表1に Twitter への画像付き投稿の例を示す．ここで，投稿された画像に対して，Bolei らの手法 [8] を用いて屋外画像を収集する．Bolei らは，places365 という 365 種類のシーンラベルが付与された画像データベースを用い，画像のシーン推定器を学習した．365 種類のシーンラベルはそれぞれ屋内，屋外のクラスに分類することができるため，屋内のシーンラベルを 0，屋外のシーンラベルを 1 として，シーンラベルの k 位の推定結果 $Pre(S_n)_k = l, l \in \{0, 1\}$ に対して，上位 K 位の平均 $\frac{1}{K} \sum_{k=1}^K Pre(S_n)_k > Th_{outdoor}$ を満たす画像 S_n を屋外画像として抽出する．Bolei らの手法による屋外画像推定例を図2に示す．

屋外画像として抽出された画像には，風景を中心とする画像と被写体が存在する画像がある．被写体が存在する画像は，被写体を中心に画像が構成されているため，気象状況が判断しやすい風景部分が写りにくい傾向がある．そのため，より気象状況が判断しやすい風景を中心とした屋外画像を抽出する必要がある．マイクロブログに投稿される画像には，ユーザが友人等と写真を撮影した被写体が人間の画像が多い点に着目し，Redmon らの物体検出手法 [9] を用いて，屋外画像から人間が含まれない画像を抽出する．Redmon らの手法は，全層が畳み込み層で構築されたネットワークを用いて，画像を分割し，分割した各領域に対して，物体の存在確率を導出し，物体が存在すると判定された領域に対して，物体のクラスを判定する物体検出器を学習した．収集した屋外画像 S_n に対して，物体の存在確率 $P_O(S_n) > Th_{object}$ を満たす領域が含まれ，物体のクラスが person と判定された画像を屋外画像 S_n から取り除く．Redmon らの手法を用いた人物検出例を図3に示す．

2.1.2 気象センサ情報の収集

Openweathermap [1] には，世界中の 20 万地点に設置された気象センサによって数時間おきに観測された情報が公開されており，表2に示すような気象センサ情報を取得できる．抽出した画像 S_n の投稿位置 $sg_n (n = 1, 2, \dots)$ ，時間 st_n と最も近

表 2 気象センサ情報の例

観測時間	観測地点	温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
2016/03/06 01:20:49	37.1951 -93.2862	18.3	1014	67	1.6	270	1	0	0	Clear
2016/03/06 02:11:53	37.775 -122.418	-4.2	1014	93	3.6	170	90	1.78	0	Rain

OUTDOOR



Top-1 : 0.112, mountain
Top-2 : 0.100, canyon
Top-3 : 0.088, desert
Top-4 : 0.057, rainforest
Top-5 : 0.057, valley

INDOOR



Top-1 : 0.536, closet
Top-2 : 0.150, fabric_store
Top-3 : 0.118, storage_room
Top-4 : 0.059, clothing_store
Top-5 : 0.020, shoe_shop

図 2 Bolei らの手法による屋外画像推定例

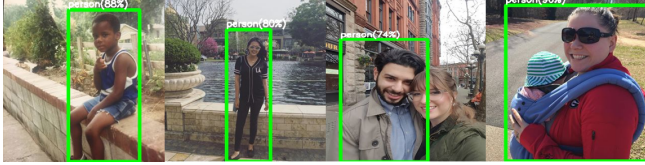


図 3 Redmon らの手法による人物検出例

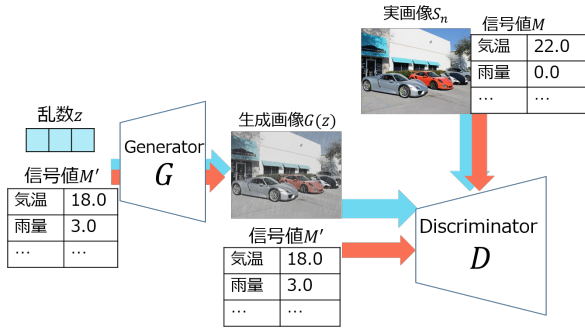


図 4 屋外画像生成器の学習概要

接した観測地点 $pg_i (i = 1, 2, \dots)$, 時間 $pt_i^j (j = 1, 2, \dots)$ において得られた気象センサによる信号値 $M_i^j = (m_{i,1}^j, \dots, m_{i,F}^j)$ を, 同じ気象状況に対する観測情報として収集する.

2.2 可逆的条件付き GAN を用いた屋外画像の生成

2.1 節で収集した気象状況の観測データを用い, ICGAN を用いた任意地点の気象状況を考慮した屋外画像の生成モデルを構築する. ICGAN は, 生成器への入力を乱数ではなく, 入力画像から屋外画像特徴抽出器によって取得した特徴量と条件ベクトルを屋外画像生成器に入力し, 入力画像の特徴を残したまま, 条件ベクトルに応じた変更を行った画像を生成することができる.

2.2.1 気象状況を考慮した屋外画像生成器の学習

2.1 節で収集した気象状況の観測データを用い, 条件ベクトルに応じた画像を生成する条件付き GAN (CGAN) [10] を用いて, 気象状況を考慮した屋外画像生成器を学習する. 図 4 に屋外画像生成器の学習概要図を示す. CGAN は, 生成器 G に乱数 z と条件ベクトル $\hat{M}$ を入力することで, 条件ベクトルを考慮した生成画像 $G(z, \hat{M})$ を出力する. 生成画像 $G(z, \hat{M})$ と信号値 $\hat{M}$, もしくは実画像 S_n と信号値 M を識別器 D に入力し, 入力した画像と信号値に対する識別結果 $D(S_n, M)$, $D(G(z, \hat{M}), \hat{M})$ を出力する. 識別器は実画像と生成画像を正しく識別するよう

表 3 生成器の構成

演算	フィルタサイズ	ストライド	フィルタ数	バッチ正規化	活性化関数
結合	生成器への入力と条件ベクトルを結合				
全畳み込み	4×4	2×2	512	行う	ReLU
全畳み込み	4×4	2×2	256	行う	ReLU
全畳み込み	4×4	2×2	128	行う	ReLU
全畳み込み	4×4	2×2	64	行う	ReLU
全畳み込み	4×4	2×2	3	行わない	Tanh

表 4 識別器の構成

演算	フィルタサイズ	ストライド	フィルタ数	レイヤ正規化	活性化関数
畳み込み	4×4	2×2	64	行う	Leaky ReLU
結合	1 層目の出力と条件ベクトルを結合				
畳み込み	4×4	2×2	128	行う	Leaky ReLU
畳み込み	4×4	2×2	256	行う	Leaky ReLU
畳み込み	4×4	2×2	512	行う	Leaky ReLU
畳み込み	4×4	1×1	1	行わない	Sigmoid

に, 生成器は識別器が実画像と誤判別する画像を生成するよう, 以下の目的関数を最適化して, 生成器 G と識別器 D を逐次的に学習する.

$$\min_g \max_d v(\theta_g, \theta_d) = \mathbb{E}_{S_n, M \sim p_{data}} [\log D(S_n, M)] + \mathbb{E}_{z \sim p_z, \hat{M} \sim p_M} [\log(1 - D(G(z, \hat{M}), \hat{M}))] \quad (1)$$

ただし θ_g, θ_d は生成器 G , 識別器 D のパラメータ, p_M は未知の信号値の分布, p_{data} は実際の信号値の分布, z は一様分布や正規分布など既知の分布 p_z からサンプルされた乱数とする. 識別器が上手く実画像と生成画像を識別できるようになると, $D(S_n, M)$ が 1 に近づき, $D(G(z, \hat{M}), \hat{M})$ が 0 に近づくため, $v(\theta_g, \theta_d)$ が大きくなる. 一方, 生成器が実画像に近い画像を生成できるようになると, 識別器が生成画像を識別できなくなるため, $D(G(z, \hat{M}), \hat{M})$ が 1 に近づき, $v(\theta_g, \theta_d)$ が小さくなる.

GAN は, 生成器が未知の分布を実画像の分布に近づけることで, 生成結果を向上させるが, 分布間の距離の定義によって学習の収束性, 安定性に差異がある. GAN の目的関数は, 近似的に Jensen-Shannon (JS) 距離を近似的に最小化することに対応するが, JS 距離による最小化では, 多くの場合において, 勾配消失問題によって学習が安定しない. これは, 未知の分布と実画像の分布が連続ではなかった場合に, 生成画像と実画像を完全に分類出来てしまうため, 誤差, 勾配が 0 に近づいてしまうためである. そのため, 他の距離指標を用いて, 生成器を最適化する必要がある. Arjovsky ら [11] は, 各分布を荷物量とみなし, 荷物を他の分布に移し替えるコストを示す Wasserstein 距離を用いることで, 目的関数を連続的に定義できることを示した.

そこで本研究では, 生成器の出力分布と実画像の分布の Wasserstein 距離を示す目的関数を最適化する. 以下に, Wasserstein 距離を用いた目的関数を示す.

表 5 屋外画像特徴抽出器の構成

演算	フィルタサイズ	ストライド	フィルタ数	バッチ正規化	活性化関数
畳み込み	5 × 5	2 × 2	32	行う	ReLU
畳み込み	5 × 5	2 × 2	64	行う	ReLU
畳み込み	5 × 5	2 × 2	128	行う	ReLU
畳み込み	5 × 5	2 × 2	256	行う	ReLU
全結合	-	-	4096	行う	ReLU
全結合	-	-	100	行わない	なし

$$\min_g \max_d v(\theta_g, \theta_d) = \mathbb{E}_{S_n, M \sim p_{data}} [D(S_n, M)] - \quad (2)$$

$$\mathbb{E}_{z \sim p_z, \hat{M} \sim p_M} [D(G(z, \hat{M}), \hat{M})]$$

ただし、式 (2) によって、生成器と識別器が最適化されたとき、識別器は生成器の出力分布と実画像の分布下のほぼ全ての点において単位勾配ノルムを持つことが示されている [11]。そのため、識別器が生成器の出力分布と実画像の分布下のほぼ全ての点においてノルム 1 の勾配を持つよう、ペナルティ項を目的関数に導入する WGAN-gp が提唱されている [12]。ペナルティ項を導入した以下の目的関数を最適化するように、生成器 G と識別器 D を逐次的に学習する。

$$\hat{S}_n = \epsilon S_n + (1 - \epsilon) G(z, \hat{M}), \epsilon \sim U[0, 1] \quad (3)$$

$$\begin{aligned} \min_g \max_d v(\theta_g, \theta_d) = & \mathbb{E}_{S_n, M \sim p_{data}} [D(S_n, M)] \\ & + \mathbb{E}_{z \sim p_z, \hat{M} \sim p_M} [D(G(z, \hat{M}), \hat{M})] \quad (4) \\ & + \lambda \mathbb{E}_{\hat{S}_n \sim \mathbb{P}_{\hat{S}_n}} [(\|\nabla_{\hat{S}_n} D(\hat{S}_n)\|_2 - 1)^2] \end{aligned}$$

λ は、ペナルティ係数である。

また、識別器は実画像と生成画像の判別だけではなく、画像に対する条件ベクトルが一致しているかに対しても頑強である必要がある [13]。そこで、実画像と間違った条件ベクトルのペアを用いて、条件ベクトルの一致に対する学習も行う。間違った条件ベクトルを fM と定義すると、 $D(G(z, \hat{M}), \hat{M})$ は以下のように変形される。

$$D(G(z, \hat{M}), \hat{M}) = 0.5 * D(G(z, \hat{M}), \hat{M}) + 0.5 * D(S_n, fM) \quad (5)$$

表 3 に生成器の構造、表 4 に識別器のネットワーク構造を示す。生成器において、画像のアップサンプリングを行う全畳み込み層では、出力画像にチェッカー状の模様が生じる問題がある。そこで、全畳み込みの代わりに、最近傍法によって画像を拡大した後、畳み込みをすることで、アップサンプリングする [14]。

2.2.2 屋外画像特徴抽出器の学習

2.2.1 節で学習した CGAN を用いて、屋外画像特徴抽出器を学習する。通常の CGAN では、入力が乱数であるため、ある屋外画像に対して、屋外画像の特徴を残したまま、画像を生成することができない。そこで、実画像に対する潜在変数となる z' を推定する特徴抽出器を学習し、CGAN と統合することにより、実画像に対して条件ベクトルとなる気象状況の信号値 M に応じた変更を加える生成器を構築する。

特徴抽出器 E は、図 5 に示すように、2.2.1 節で学習した生



図 5 屋外画像特徴抽出器の学習概要

成器 G に乱数 z と信号値 M' を入力し生成した画像 $G(z, M')$ を用いて、入力した乱数 z を推定するよう学習する。つまり、入力した乱数 z と生成した画像の推定乱数 $E(G(z, M'))$ の差を小さくする以下の損失関数の最小化を目的とする。

$$\mathcal{L}_{ez} = \mathbb{E}_{z \sim p_z, M' \sim p_y} \|z - E(G(z, M'))\|_2^2 \quad (6)$$

表 5 に特徴抽出器のネットワーク構造を示す。

2.2.3 任意地点の気象状況を考慮した屋外画像生成モデルの構築

学習した屋外画像生成器と屋外画像特徴抽出器を用いて、ICGAN を構築する。入力画像を屋外画像特徴抽出器に入力し、入力画像に対する潜在表現となる z' を抽出する。潜在表現 z' と条件ベクトルとなる気象状況の信号値 M を屋外画像生成器に入力し、気象状況の信号値 M に応じた屋外画像 $G(z', M)$ を生成する。

3. 実 験

本章では、任意地点の気象状況を考慮した屋外画像の生成実験の概要を述べた後、提案手法によって収集した気象状況を観測したデータ、収集したデータを用いて学習した屋外画像の生成モデルについて評価及び考察を行い、今後の課題について言及する。

3.1 気象状況観測データの妥当性

収集した画像のうち、2016 年の 3 月 3 日から 2016 年 3 月 9 日の 1 週間分の画像 204,844 枚に対して、提案手法によって屋外画像を抽出した。提案手法は、Bolei らの手法による屋外画像抽出と Redmon らの手法による物体検出の 2 種類の手法を用いて、適切な屋外画像を抽出する。

まず、Bolei らの手法を用いて、屋外画像を抽出した。パラメータは、 K と $Th_{outdoor}$ がある。 $K = 10$ に固定し、 $Th_{outdoor} = 0.5, 0.7, 0.8, 0.9, 1.0$ と変更したときの画像枚数と抽出画像から、抽出結果を評価した。抽出された画像枚数と画像例を図 7 に示す。 $Th_{outdoor}$ が小さいほど、抽出された画像に気象状況を確認できない画像が含まれやすい。図 6 に、これらの誤検出された画像の例を示す。 $Th_{outdoor}$ が大きくなるほど、画像の誤検出は減少し、 $Th_{outdoor} = 0.5$ のときには誤検出が見られた被写体との距離が近い画像や屋内画像は、 $Th_{outdoor} = 0.9$ まで大きくすることで、ほぼ誤検出されなくなった。また、 $Th_{outdoor} = 1.0$ まで大きくしたとき、誤検出画像に対する抽出結果にはほぼ変化が見られなかったが、総抽



図 6 誤抽出された画像例

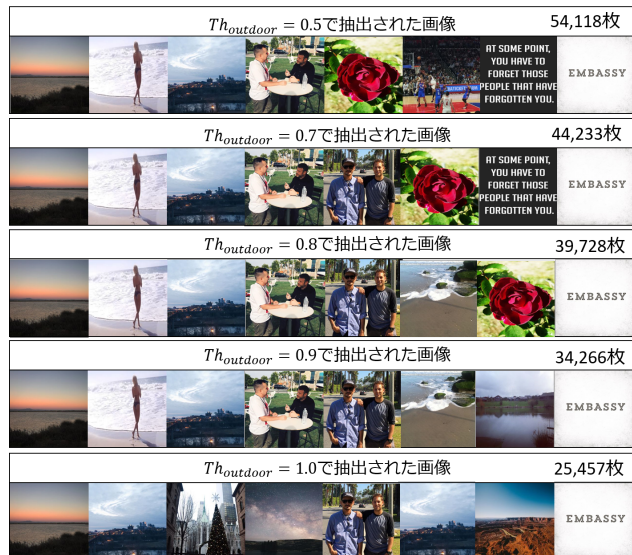


図 7 Bolei らの手法による屋外画像抽出結果

出枚数が大きく減少した。これは、 $Th_{outdoor}$ が一定以上に大きくなると、誤検出画像が減少するわけではなく、多くの屋外画像が見抽出抽出となるためと考えられる。

Bolei らの手法を用いて、屋外画像を抽出したが、抽出された屋外画像の中には、人間が画像の中心となって、気象状況が判断しにくい画像が存在する。そこで、Redmon らの手法を用いて、Bolei らの手法によって抽出した屋外画像 34,266 枚から、さらに風景を中心とした屋外画像のみを抽出した。パラメータとして、 Th_{object} があり、 $Th_{object} = 0.4, 0.5, 0.6, 0.7$ と変更したときの抽出画像枚数と画像例から、抽出結果を評価した。 Th_{object} が大きいほど、物体存在確率が大きいものしか含まれないため、人物が写っていても誤抽出される画像が多かったが、 $Th_{object} = 0.5$ まで小さくすると、顔の一部しか写っていない画像や人間がかなり小さい画像もほぼ見られなくなった。 $Th_{object} = 0.4$ まで小さくすると、 $Th_{object} = 0.5$ までには検出されなかった人物に対しても、検出可能になったが、 $Th_{object} = 0.5$ までの検出結果に比べて、明らかに誤検出が多くなった。

以上より、今回の実験では、 $Th_{outdoor} = 0.9$ 、 $Th_{object} = 0.5$ が適切であると考えられる。また、2 週間分の投稿画像 539,160 枚から抽出された屋外画像 52,209 枚を用いて、Openweathermap から、屋外画像が投稿された地点、時間に最も近い気象センサの信号値を収集した。表 7 に対となる屋外画像と気象センサの信号値の例を示す。Openweathermap に公開されたデータには気象センサによる信号値と共に、「Clear」、「Rain」といっ

表 6 データセットに含まれる気象状況

天候状況	データ数
Clouds	20,610
Clear	18,098
Rain	4,798
Snow	4,575
Haze	3,856
Drizzle	213
Dust	43
Thunderstorm	16
合計	52,209



図 8 Redmon らの手法による屋外画像の抽出結果

た天候を表す単語が含まれる。この単語を踏まえると、青空が撮影された画像に対し「Clear」、曇り空の画像には「Clouds」、雪景色の画像には「Snow」といったように、屋外画像に対して適切な信号値が対として得られていることが分かる。一方、収集した信号値のうち特に雨量、雪量は取得が失敗したとき「NULL」となる場合がある。このため、雨の画像に対し、正しく「Rain」の信号値が収集された場合も、雨を表す雨量に対して適切な信号値が得られない例も存在した。また、屋外画像の撮影時間と Twitter への投稿時間にタイムラグがあるため、屋外画像に対して対応する信号値が正しく得られない場合もあった。

また、表 6 に示すように、気象センサの信号値に付与された天候を表す単語によると、収集された環境の気象状況には偏りが生じる。これは屋外画像が撮影されるときは天候が晴れもしくは曇りの場合が多いためと考えられる。

3.2 屋外画像生成器の評価

抽出した 52,209 件の屋外画像と気象センサからの信号値の対を用い、条件ベクトルに応じた画像を生成する CGAN を用いて、信号値 M に応じた屋外画像を生成する生成器を学習した。屋外画像生成器は、

- ・CGAN を用いた屋外画像生成器
- ・WGAN-gp を用いた屋外画像生成器
- ・CGAN を用いて使用するデータを変更した屋外画像生成器の 3 パターンを用いて、目的関数や学習データの変化による影響を評価した。学習の前処理として、屋外画像のサイズを全て 64×64 に変更し、気象状況の信号値 M_i^j に含まれる NULL は全て 0 に変えた後、信号値 M_i^j を標準化した。

表 7 対となる屋外画像と気象センサの信号値の例

屋外画像	温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
	26.57	1015	10	2.1	0	1	0	0	Clear
	6.42	1021	45	2.1	110	90	0	0	Clouds
	-0.79	1021	89	3.6	80	90	0	2	Snow
	-0.04	1018	92	5.1	40	90	NULL	NULL	Rain
	9.05	1016	71	4.7	240	1	0	0	Clear

表 8 CGAN を用いた屋外画像生成器による生成例

生成画像	温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
(a) 	13.62	1031	58	2.6	360	1	0	0	Clear
(b) 	10.1	1012	66	2.4	230	90	0	0	Cloud
(c) 	3.1	1002	98	3.6	310	75	0.3	0	Rain
(d) 	12.26	1015	73	2.2	120	90	NULL(0)	NULL(0)	Rain

CGAN を用いた屋外画像生成器では、 p_z は正規分布とし、最適化アルゴリズムを Adam ($\beta_1 = 0.5$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$), 学習率を 0.0002, バッチサイズを 64, 学習回数を 300 回とした。 p_z から得られる乱数と気象センサから得られた信号値 M を入力し、表 8 に同じ乱数に対して、異なる信号値 M を与えた生成結果を示す。

同じ乱数に対して異なる信号値 M を与えたとき、表 8(a), (b) のように類似した風景ではあるが、信号値の雲量に応じて空の様子が異なる画像が生成された。特に学習データに多く含まれた晴れと曇りの信号値を用いた場合、このように適切に気象状況に応じた画像が生成された。一方、学習データにおける数が少ない天候が付与された信号値を与えた場合や、信号値に「NULL」を含む信号値を与えた場合は、表 8(c), (d) のように見た目に気象状況が分かりにくい画像が生成された。そのため、天候の偏りの解消や、信号値に「NULL」が観測された際に、0 のような固定値を入れるのではなく、近辺に設置された他の気象センサから情報を補完するなどして、データセットの質を向上させる必要がある。

次に、WGAN-gp を用いて、生成器を学習した。 p_z は正規分布とし、最適化アルゴリズムを Adam ($\beta_1 = 0.5$, $\beta_2 = 0.9$, $\epsilon = 10^{-8}$), 学習率を 0.0001, バッチサイズを 64, 学習回数を 300 回とした。また、1 回生成器を学習する前に、識別器を

表 9 WGAN-gp を用いた成器の生成結果例

生成画像	温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
(a) 	13.62	1031	58	2.6	360	1	0	0	Clear
(b) 	10.1	1012	66	2.4	230	90	0	0	Cloud
(c) 	3.1	1002	98	3.6	310	75	0.3	0	Rain
(d) 	12.26	1015	73	2.2	120	90	NULL(0)	NULL(0)	Rain

5 回学習した。表 8 と同様に、同じ乱数に対して、異なる信号値 M を与えた生成結果を表 9 に示す。また、生成に用いた信号値 M は、表 8 の生成例を生成した際の信号値を用いた。

表 8(a)~(d) と同様に、学習データに多く含まれた晴れと曇りの信号値に対しては、表 9(a), (b) のように気象状況に応じた画像を生成するが、学習データにあまり含まれなかった天候の信号値や、「NULL」が含まれた信号値を与えた場合は、表 9 の表 9 は適切な画像が生成されなかった。しかし、表 8 に比べると表 9 は全体が鮮明な画像が生成された。また表 8(c), (d) に比べると表 9(c), (d) の画像は信号値に適切ではないが、空が曇ったような気象状況が判断できる画像が生成された。これは、目的関数を変更したことによって、GAN の性能が向上したため、出力結果に影響を及ぼしたと考えられる。

最後に、CGAN を用い、別のデータセットを用いて、生成器を学習した。GAN は、顔画像データ等、データセットに含まれる画像が全体的に似通っている場合、とくに良好な結果を出す。そこで気象状況の変化は、雲や太陽等、空に多く表れることに着目し、Bolei らの手法を用いて屋外画像を抽出する際に、シーンラベルの推定結果上位 K 位に「sky」が含まれる画像を屋外画像として抽出した。抽出した「sky」がシーンラベル推定結果に含まれる屋外画像 8,749 枚を「sky」データセットとして、屋外画像生成器を学習した。パラメータ、信号値はと同様のものを用いた。

生成結果を表 10 に示す。これまでの生成結果と同様、表 10(a),(b) では、信号値の雲量に対して生成画像の雲量が適切に変化し、表 10(c), (d) では、信号値に対して適切な変化が行えなかった。また、表 10(a), (b), (c), (d) の生成画像が、これまでの生成結果に比べて全体が類似した風景で出力されており、同一の乱数による生成結果として、より適切であると思われる。これは、データセットを「sky」に絞ったことで、学習に用いた画像の傾向が似ていたためであると考えられる。ただし、表 9(a)~(d) と比較すると、生成された画像全体がぼやけた状態になった。これは、「sky」のシーンラベルが推定された画像が 8,749 枚と非常に少ない枚数になったため、十分な学習が行えなかったためであると考えられる。

表 10 「sky」データセットで学習した屋外画像生成器の生成結果例

生成画像	温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
(a)	13.62	1031	58	2.6	360	1	0	0	Clear
(b)	10.1	1012	66	2.4	230	90	0	0	Cloud
(c)	3.1	1002	98	3.6	310	75	0.3	0	Rain
(d)	12.26	1015	73	2.2	120	90	NULL(0)	NULL(0)	Rain

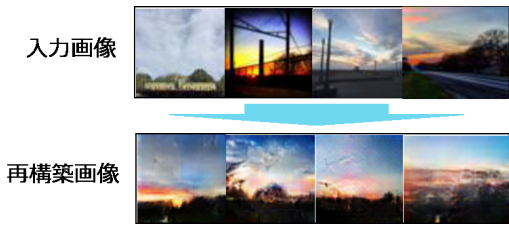


図 9 CGAN を用いた屋外画像生成器による学習データを用いた再構築例

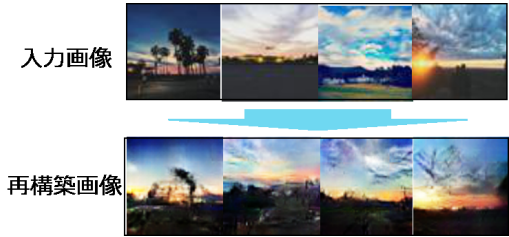


図 10 WGAN-gp を用いた屋外画像生成器による学習データを用いた再構築例

3.3 屋外画像特徴抽出器の評価

前節で学習した 3 種類の屋外画像生成器を用いて、正規分布に従う乱数と気象センサから得られた信号値を入力し、それぞれ 100,000 枚の画像を生成した。これを学習データとして、屋外画像特徴抽出器を学習した。最適化アルゴリズムやパラメータは屋外画像生成器と同じものとし、学習回数は 300 回とした。

入力画像から抽出した特徴量と信号値を用いて、屋外画像生成器による入力画像の再構築によって、屋外画像特徴抽出器の性能を評価した。

CGAN を用いた屋外画像生成器によって学習した屋外画像特徴抽出器で、画像の再構築を行った例を図 9 に示す。再構築した画像は、入力画像の概形を部分的に捉えることもあるが、全く異なる画像を生成する場合もあり、全体的にばげた画像となった。これは、学習に用いた画像は、前節で学習した屋外画像生成器から生成されるため、屋外画像生成器が十分に学習できていないと、特徴抽出器の学習が上手く行えないためであると考えられる。

次に、WGAN-gp を用いた屋外画像生成器によって学習した

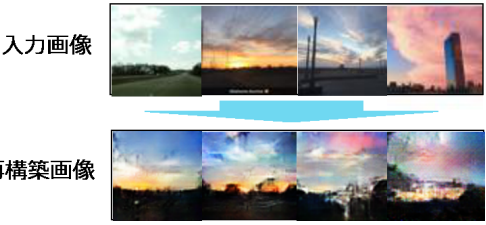


図 11 「sky」データセットを用いた屋外画像生成器による学習データを用いた再構築例

表 11 変更した気象状況の信号値 M'

温度	気圧	湿度	風速	風向	雲量	雨量	雪量	天候
23.2	1014	67	1.6	270	1	0	0	Clear

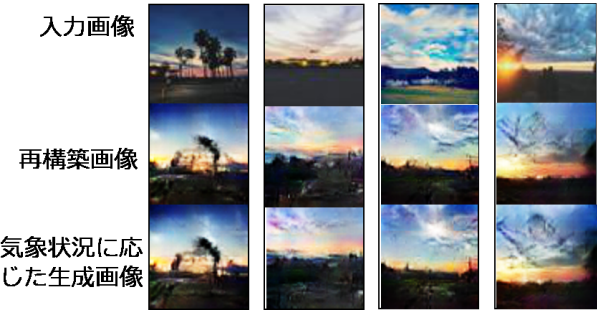


図 12 気象状況に応じた屋外画像生成例

屋外画像特徴抽出器で、画像の再構築を行った例を図 10 に示す。図 9 に比べて、入力画像の特徴を捉え、画像を再構築しているのが分かる。これは、目的関数を変更したことで、生成器の性能が向上したことで、屋外画像特徴抽出器の学習データの質が向上したためであると考えられる。

次に、別のデータセットを用いた屋外画像生成器によって学習した屋外画像特徴抽出器で、画像の再構築を行った例を図 11 に示す。再構築の精度が低く、また再構築画像がどの画像も同じ形を取っている。これは、学習データが「sky」のシーンラベルが推定された 8,749 枚に変わったため、屋外画像生成器によって生成された学習データ内に、近い見た目をした画像が多く存在したためと考えられる。

3.4 任意地点の気象状況を考慮した屋外画像の生成結果

最後に、3.2 節、3.3 節で学習した屋外画像生成器と屋外画像特徴抽出器を結合して、入力した屋外画像を気象状況に応じて変更した画像を生成した。その際、3.3 節で最も適切な画像の再構築を行った WGAN-gp による屋外画像生成器と屋外画像特徴抽出器を用いた。図 12 に、任意地点の気象状況を考慮した屋外画像の生成結果を示す。また、対応する信号値 M' を表 11 に示す。

空の一部が青く「Clear」に近づいている部分もあるが、3.2 節における屋外画像生成器の生成例と比べると、気象状況の信号値による画像の変化が小さい。これは、再構築した画像は、屋外画像生成器のみによる生成例よりも全体的にばげているため、気象状況の信号値による変化が上手くいかなかったと考えられる。

4. ま と め

本研究では、ICGAN を用いた気象状況を考慮した屋外画像生成モデルによって、気象状況を考慮した屋外画像を生成する手法を提案した。

アメリカ本土に投稿された2週間分のジオタグ付き投稿から収集した画像539,160枚を用いた実験では、適切なパラメータの調節によって、誤検出、未検出がほぼない適切な屋外画像52,209枚の画像を抽出した。また、屋外画像に対して対応付けた気象センサの信号値は、ある程度の整合性を持って、屋外画像に適切な信号値が付与された。

収集した屋外画像によって学習された屋外画像生成器は、晴れや曇り等、学習に用いたデータ内に十分な数が存在する気象状況に対して、対応する屋外画像を生成した。屋外画像特徴抽出器によって、ある程度入力画像の概形を捉えた再構築画像を生成した。さらに、屋外画像生成器と屋外画像特徴抽出器によって、任意地点の気象状況を考慮した屋外画像の生成モデルを構築し、元画像の特徴を残した画像を信号値によって変化させた。

しかし、屋外画像特徴抽出器による再構築画像にぼやけが生じると、信号値によって適切に変化せず、また屋外画像特徴抽出器は学習の際に、屋外画像生成器の性能によって影響を受けることが示された。今後の課題としては、屋外画像特徴抽出器の抽出精度の向上が挙げられる。また、抽出精度は屋外画像生成器による影響を受けるため、屋外画像生成器の性能を向上させる手法の検討が必要である。

謝 辞

本研究の一部は科学研究費補助金基盤研究(S)16H06302による。

文 献

- [1] “OpenWeatherMap”, <https://openweathermap.org/>
- [2] “ウェザーリポート”, <https://weathernews.jp>
- [3] Y. Wang and M. S. Kankanhalli, “Tweeting Cameras for Event Detection”, International Conference on World Wide Web, pp. 1231–1241, 2015.
- [4] P. Giridhar, M. T. Amin, T. Abdelzaher, L. Kaplan, J. George, and R. Ganti, “Clarisense: Clarifying Sensor Anomalies Using Social Network Feeds”, International Conference on Pervasive Computing and Communication, pp. 395–400, 2014.
- [5] “Twitter”, <https://twitter.com/>
- [6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative Adversarial Nets”, Advances in Neural Information Processing Systems, pp. 2672–2680, 2014.
- [7] G. Peranau, J. van de Weijer, and B. Raducanu, “Invertible Conditional GANs for Image Editing”, NIPS Workshop on Adversarial Training, 9 pages, 2016.
- [8] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, “Places: A 10 Million Image Database for Scene Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 pages, 2016.
- [9] J. Redmon, and A. Farhadi, “YOLO9000: Better, Faster, Stronger”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 9 pages, 2016.

- [10] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets”, Computing Research Repository, vol. abs/1411.1784, 7 pages, 2014.
- [11] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein gan”, arXiv:1701.07875, 32 pages, 2017.
- [12] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Chintala, “Improved Training of Wasserstein GANs”, Advances in Neural Information Processing Systems, pp. 5769–5779, 2017.
- [13] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, “Generative Adversarial Text to Image Synthesis”, arXiv:1605.05396, 10 pages, 2016.
- [14] A. Odena, V. Dumoulin, and C. Olah, “Deconvolution and Checkerboard Artifacts”, <http://distill.pub/2016/deconv-checkerboard/>, 2016.