

半教師あり学習による気象条件に応じた屋外画像の生成

松崎 大輔[†] 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: [†]matsuzaki@nanase.comm.eng.osaka-u.ac.jp,

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

あらまし 近年、多種多様なセンサが実世界を観測しており、特に気象センサは時空間的に密に世界各地の気象状況を観測している。しかし、気象センサから得られる温度や湿度といった信号値のみから人間が直感的に気象状況を把握することは困難である。一方、人間がより理解しやすい形式である画像で各地を観測できるカメラは設置場所が限定される。そこで本研究は、任意の場所、時間の気象状況を表す屋外画像を、気象センサ値を条件とする画像変換により生成することを目的とする。従来の条件付き画像変換は、学習データとして、条件と画像の対を大量に用意する必要がある。提案手法では、気象センサから得られる信号値に対する対として、収集が容易なソーシャルメディアに投稿された屋外画像にノイズが多いことに着目し、特に整合性の高い少量の対のみを用いた半教師あり学習による条件付き画像変換により、気象条件に応じた屋外画像生成を実現する。

キーワード 気象センサ、画像変換

1 はじめに

近年、多種多様なセンサが実世界を観測しており、特に気象センサは時空間的に密に世界各地の気象状況を観測している。例えばOpenWeatherMap [1] では、世界各地の約 20 万地点に設置された気象センサから数時間おきの観測値を公開している。しかし、気象センサから得られる温度や湿度といった信号値のみから人間が直感的に気象状況を把握することは困難である。一方、ウェザーリポート [2] などでは、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。

画像生成の分野では、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) [3] が注目されている。例えば、条件付き GAN [4,5] では、学習データとして、顔画像と性別や年代など、実画像と条件の対を大量に用いることにより、乱数を入力とし、条件に応じて実際には存在しない偽の顔画像を生成することができる。また入力された画像の条件を変更する画像変換も実現されている。pix2pix [6] や CycleGAN [7] では、白黒とカラー画像、昼と夜の画像、線画と実写画像など、同じ対象、もしくは異なる対象に対して 2 つの異なる条件の画像対を学習データとして用いることにより、一方の条件のもと撮影された実画像から、他方の条件の偽画像を生成することができる。また、条件付き GAN と同様に、実画像と条件の対を学習データとし、入力画像を多様な条件に応じて変更することにより偽画像を生成するネットワークも提案されている [7,8]。

そこで本研究では、このような画像変換技術を用い、任意地点の屋外画像に対し、現在の気象状況を条件として与えること

により、該当地点の現在の気象条件を表す屋外画像の生成を目指す。従来の条件付き画像変換手法は、学習データとして条件と画像の対を大量に用意する必要があるため、各地の気象センサの信号値（以降、気象信号値と呼ぶ）と共に、同じ地点、時間の気象状況を捉える屋外画像を学習データとして大量に収集する必要がある。多様な屋外画像の収集には、ソーシャルネットワーキングサービス (SNS) を利用することが望ましいが、画像の撮影時間と投稿時間のずれや、画像の撮影デバイスの時間設定などに依存する撮影時間情報のずれなどの影響により、気象信号値と正確に対応のとれた屋外画像の収集は困難である。提案手法ではまず、時空間的な対応付けのみでなく、意味的な整合性を考慮することにより、できるだけ正確な対応関係を持つ気象信号値と屋外画像の対を収集する。また、少数の学習データ対のみを用いて画像変換を実現するため、収集した気象信号値と屋外画像を用いて、画像に対する気象信号値の推定器を学習する。この推定器を用いることにより、学習データとして大量の屋外画像のみを収集し、推定器により推定された気象信号値と対として用いることにより、気象条件に応じた画像変換器の半教師あり学習を実現する。

2 提案手法

本研究では、任意の位置における気象センサから得られる任意の時間の気象信号値が与えられたとき、観測された気象状況を適切に表す屋外画像を生成することを目的とする。

SNS には、各地の人々が観測した屋外状況を示す画像が、位置、時間情報と共に投稿され得る。そこでまず画像変換器の学習データに用いるため、SNS へ投稿された屋外画像を収集する。次に、各屋外画像の気象状況を表す気象信号値として、投稿位置、時間に最も近接する気象センサから取得された信号値を収集する。さらに、これらの屋外画像と気象信号値は同じ気

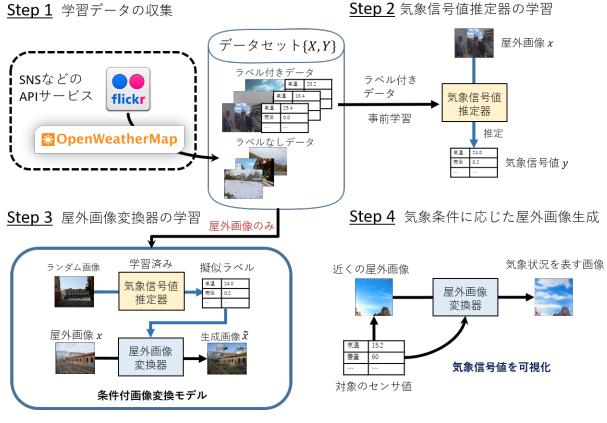


図 1 提案手法の概要

象状況を観測したものであるという前提のもと、画像と気象信号値それぞれから推定される気象状況の整合性が高い場合のみ、画像を気象信号値と対とする。

このように収集した少數の信頼できる学習データ対のみを用いて画像変換を実現するため、提案手法は画像変換器の半教師あり学習を行う。まず、対となる気象信号値と屋外画像を用いて、画像に対する気象信号値の推定器を学習する。この推定器を用いることにより、SNS から収集されたすべての屋外画像に対して、推定器により推定された気象信号値と対として、気象条件に応じた画像変換器を学習する。

提案手法は図 1 に示すように、以下の 4 ステップにより構成される。

Step 1) 学習データ収集

SNS から屋外画像を収集し、そのうち、位置、時間が近接する気象信号値が、屋外画像が示す気象状況と整合性が高い場合、対とする。

Step 2) 気象信号値推定器の学習

Step 1)において対となった屋外画像と気象信号値を用いて、画像から気象信号値を推定する気象信号値推定器を学習する。

Step 3) 気象条件に応じた屋外画像変換器の学習

Step 1)で収集したすべての屋外画像に対して、Step 2)で学習した気象信号値推定器を用いた条件付き画像変換モデルにより、画像データのみを入力とし、気象条件に応じた屋外画像変換器を学習する。

Step 4) 気象条件に応じた屋外画像生成

与えられた気象信号値に対し、近接する位置で撮影された任意の時間の屋外画像を取得し、Step 3)で学習した屋外画像変換器に入力することにより、気象状況を表す屋外画像を生成する。

以降の節より、それぞれのステップに関する詳細を述べる。

2.1 学習データ収集

まず、SNS に投稿された画像 S から、各地の気象状況を観測したものとして、屋外で撮影された画像を収集する。各投稿画像に対し、Zhou らの手法 [9] である 365 種類のシーンラベル付き画像データベースを用いて学習されたシーン推定器を用いて屋外判定を行う。365 種類のシーンラベルはそれぞれ屋内、

	温度	湿度	気圧	雲量	雨量	風速
303.71	25	1008	0	0	0.96	
x	$y \in \mathbb{R}^N (N = 6)$					

図 2 データ対の例

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 $Pre(S)_k = l$, $l \in \{0, 1\}$ に対して、上位 K 位の平均 $\frac{1}{K} \sum_{k=1}^K Pre(S)_k > Th_{outdoor}$ を満たす画像 S を屋外画像 x として抽出する。

次に、各屋外画像の投稿位置に最も近い位置に設置された気象センサにより、投稿時間に最も近い時間に観測された気象信号値を取得する。ただし、画像の撮影時間と投稿時間のずれなどの影響により、位置と時間が近接した屋外画像と気象信号値は必ずしも同じ環境を観測したものとは限らない。そこで、同じ環境を観測した屋外画像と気象信号値は同じ気象状況を表すという前提のもと、画像と気象信号値それぞれから推定される気象状況の整合性が高い場合のみ、画像を気象信号値と対とする。

ここでは、晴れ、雨などの異なる天候クラスを表す屋外画像のデータセットが存在すること、また、気象センサの観測値に気象信号値と共に、同様の天候クラスを表すラベルが含まれることに着目し、これらを学習データとして、画像、及び気象信号値からそれぞれ天候クラスを推定する天候推定器を学習する。屋外画像 x 及び気象信号値 y それぞれに対する天候推定器による推定結果 \hat{c}_x , \hat{c}_y の交差エントロピーと、 y と共に観測値に含まれる天候クラスを表すラベル c_y と \hat{c}_x の交差エントロピーの和を x と y の乖離度 L とする。ここで、乖離度 L を、

$$L = -\lambda \hat{c}_y \log \hat{c}_x - (1 - \lambda) c_y \log \hat{c}_x \quad (1)$$

と定義する。乖離度 L の第一項では、屋外画像 x 、気象信号値 y それぞれに対する天候推定結果の類似度を計算し、第二項では、屋外画像 x に対する天候推定の精度を計算する。 L が小さいほど、気象信号値 y が、屋外画像 x 及び y 共に気象センサから得られる天候クラスと整合性が高いと考えられる。 $L < Th$ をみたす、屋外画像と信号値を整合性の高いデータ対として収集する。また、屋外画像 x と気象信号値 y のデータ対の例を図 2 に示す。ここで、気象信号値 y のベクトルの次元数を N とする。

2.2 気象信号値推定器の学習

2.1 節にて収集された屋外画像 x と気象信号値 y の対を用いて、屋外画像から気象信号値を推定する気象信号値推定器 W を学習する。学習の際に用いる損失関数 \mathcal{L}_{est} は以下を用いる。

$$\mathcal{L}_{est} = \mathbb{E}_{x,y \sim p_{data}} [|W(x) - y|^2] \quad (2)$$

気象信号値 y は一般に温度、湿度など複数の要素を持つベクトルであるため、推定器 W により推定した信号値 $W(x)$ と

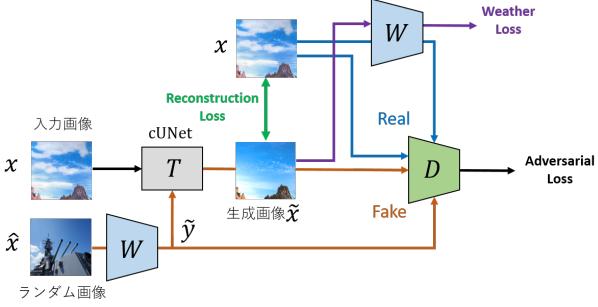


図 3 ネットワーク構成

L2 ノルムをとったものを誤差とする。またこのとき、正則化やデータ拡張などの過学習を防ぐ手法を併用し、次節で述べる画像変換器 T の性能を向上させる。

2.3 気象条件に応じた屋外画像変換器の学習

収集したすべての屋外画像を用いて、気象条件に応じた屋外画像変換器 T を学習する。図 3 にネットワーク構成を示す。図 3 のネットワークは、主に入力画像 x を条件 y に応じて変換する画像変換器 T 、偽画像 \hat{x} と実画像 x を判別する判別器 D 、画像 x に対して条件 y を推定する推定器 W の 3 つのユニットで構成される。ただし判別器 D は画像と条件の対を入力とし、 W は 2.2 節で学習した気象信号値推定器を用いる。

変換器 T の損失関数 \mathcal{L}_T は判別器 D の判別誤差 (Adversarial Loss) \mathcal{L}_{adv} と、生成画像 \hat{x} の再構成誤差 (Reconstruction Loss) \mathcal{L}_{con} 、生成画像 \hat{x} に対する気象信号値の推定誤差 (Weather Loss) \mathcal{L}_w の 3 つの要素からなり、各要素のバランスをとる係数である λ_{con} 、 λ_w を用いて以下のように表される。

$$\begin{aligned}\mathcal{L}_T &= \mathcal{L}_{adv} + \lambda_{con}\mathcal{L}_{con} + \lambda_w\mathcal{L}_w \\ \mathcal{L}_{adv} &= \mathbb{E}_{x, \hat{x} \sim p_{data}} [-D(T(x, W(\hat{x})), W(\hat{x}))] \\ &= \mathbb{E}_{x, \hat{x} \sim p_{data}} [-D(\tilde{x}, W(\hat{x}))] \\ \mathcal{L}_{con} &= \mathbb{E}_{x, \hat{x} \sim p_{data}} \left[\frac{\frac{1}{P}|x - T(x, W(\hat{x}))|}{\frac{1}{N}|W(x) - W(T(x, W(\hat{x})))|} \right] \quad (3) \\ &= \mathbb{E}_{x, \hat{x} \sim p_{data}} \left[\frac{\frac{1}{P}|x - \tilde{x}|}{\frac{1}{N}|W(x) - W(\tilde{x})|} \right] \\ \mathcal{L}_w &= \mathbb{E}_{x, \hat{x} \sim p_{data}} [|W(\hat{x}) - W(T(x, W(\hat{x})))|^2] \\ &= \mathbb{E}_{x, \hat{x} \sim p_{data}} [|W(\hat{x}) - W(\tilde{x})|^2]\end{aligned}$$

ここで、画像の全画素数を P とし、 $|X|$ は X の L1 ノルムを表す。 \mathcal{L}_{adv} はヒンジ損失関数であり、実画像と見分けがつかないような画像を生成するための制約となる。 \mathcal{L}_{con} の分子は再構成誤差であり、画像の内容を大きく変化させないための制約であるが、分母に変換前後の条件変化を用いることにより、条件の変化が大きい場合に制約を弱める。 \mathcal{L}_w は条件 $W(\hat{x})$ で生成した画像 \tilde{x} に対して、推定される気象信号値と $W(\hat{x})$ の誤差であり、条件に対して適切な画像を生成するための制約である。 \mathcal{L}_T を最小化することにより、変換器 T は、実画像と見分けがつかないように、入力画像を条件に応じて変換するよう学習される。

判別器 D の損失関数 \mathcal{L}_D を次式に示す。

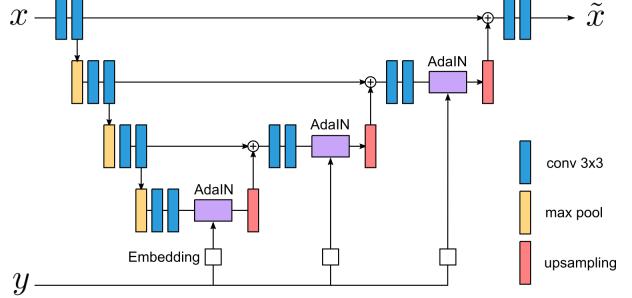


図 4 conditional U-Net の構造

$$\begin{aligned}\mathcal{L}_D = & \mathbb{E}_{x, \hat{x} \sim p_{data}} [\text{Relu}(1 - D(x, W(x))) \\ & + \text{Relu}(1 + D(T(x, W(\hat{x})), W(\hat{x})))]\end{aligned}\quad (4)$$

\mathcal{L}_D は、判別器に対して用いられるヒンジ損失関数であり、これを最小化することにより、判別器 D が、実画像 x が推定される気象信号値 $W(x)$ と共に与えられたとき 1、条件 $W(\hat{x})$ に對して生成した画像 \tilde{x} が $W(\hat{x})$ と共に与えられたとき 0 を出力する、つまり正しく実画像と生成画像を判定するよう学習される。また、これら全ての損失関数は x と \hat{x} のみを入力とするため、条件 y を必要としない。

学習時には、2.1 節で収集したすべての屋外画像を入力画像とする。各入力画像 x に対し、収集した屋外画像からランダムに 1 枚画像 \hat{x} を選択し、2.2 節で学習した気象信号値推定器 W により、疑似条件 $\hat{y} = W(\hat{x})$ を推定する。入力画像 x を擬似条件 \hat{y} によって変換し、生成画像 $\tilde{x} = T(x, \hat{y})$ を得る。判別器 D に実画像と条件の対 $x, W(x)$ 、生成画像と条件の対 \tilde{x}, \hat{y} を、推定器 W に生成画像 \tilde{x} をそれぞれ入力し、 \mathcal{L}_{adv} 、 \mathcal{L}_{con} 、 \mathcal{L}_w を算出し、これらの和を最小化するよう学習する。このように、学習に用いる条件を単純な乱数とするのではなく、収集した屋外画像の気象条件の分布に沿った条件とすることにより学習が安定する。

従来の GAN を用いた画像変換の手法の多くは、入力画像をエンコーダなどによって潜在空間に写像し、その潜在表現を条件に応じて変換した後、アップサンプリングを適用し出力画像とするものである。気象状況を表す変化では、オブジェクトの形状自体には大きな変化はなく、空模様やライティングなどの色変化が多い。よって、潜在表現に圧縮する手法では余分な再構成を学習しなければならない。そこで、変換器 T として、入出力の差分を学習する U-Net [10] のネットワークをベースとし、条件付きに対応した conditional U-Net を提案する。conditional U-Net の構造を図 4 に示す。条件の挿入方法は、近年では StyleGAN [5] などの高解像度で複雑なスタイルを含む手法にも用いられている StyleTransfer [11] で提案された Adaptive Instance Normalization(AdaIN) を採用する。まず、任意の次数を持つ条件 y に対して Linear Embedding を適応し固定次数を確保する。次に各レベルの層のアップサンプリングを行う直前に、AdaIN を用いてチャンネルごとに中間特徴量の分散と平均を制御する。AdaIN の直後は Dropout することで確率的な影響を考慮する。

3 実験

3.1 学習データ収集

Flickr [12] から、2016 年にアメリカにて撮影された画像約 110 万枚から $K = 10$, $Th_{outdoor} = 1.0$ として屋外画像約 32 万枚を抽出した。これらの屋外画像のうち、Flickr から撮影位置及び、Exif データに含まれる撮影時間が収集できるものに対し、Openweathermap から、撮影位置に最も近い気象センサの、撮影時間に最も近い気象信号値を収集した。この中から、屋外画像の撮影位置と気象センサの観測位置の誤差が 10km 以下、且つ、それぞれの撮影・観測時間の誤差が 60 分以内のデータ対、合計 181,757 セットを抽出した。さらに、収集した屋外画像、気象信号値のそれぞれに対する天候推定器を学習し、天候推定結果 \hat{c}_x , \hat{c}_y を用いることにより、整合性の高い気象信号値のみを屋外画像と対応付ける。

Openweathermap に公開された気象センサの観測値には気温、湿度、気圧、風速、雲量などの気象信号値と共に、「Clear」、「Rain」といった 11 種類の天候クラスの情報が含まれる。そこで、気象信号値に対する天候推定器は、Openweathermap からランダムに選出した観測値を用い、信号値から天候クラスを推定するよう学習した。また、晴れ、曇、霧、雨、雪の 5 種類の天候クラスに対する屋外画像データセット [13] を用い、同様に屋外画像から天候クラスを推定する天候推定器を学習した。画像に対して事前学習された VGG16 [14] を転移学習し、気象信号値に対しては 3 層ペーセptron を用いた。天候クラスは屋外画像データセットに合わせ、いずれも晴れ、曇、霧、雨、雪の 5 種類とした。画像を入力とする天候推定器の適合率と再現率は、各天候クラスの平均で、それぞれ 81.0%, 83.5% となった。一方で、気象信号値を入力とする天候推定器の適合率と再現率は、それぞれ 65.7%, 70.3% となり、画像からの天候推定器よりも劣る結果となった。これらの推定器を用いて、対応付けた屋外画像 x と気象信号値 y の整合性を評価した。ただし、乖離度 L を計算する際、特に気象信号値に対する天候クラスによる推定精度のばらつきを考慮し、気象信号値の天候推定の適合率 P を用いて、係数 λ を $\lambda_p P$ とした。このように、適合率 P を用いることにより、気象信号値からの天候推定精度が低い場合は、屋外画像に対する天候推定精度を重視する。

図 5 に、2.1 節の天候推定の結果、整合性が高い、つまり L が小さいものから上位 5 セットの気象信号値と対応付けられた屋外画像例を示す。各画像に対応付けられた気象信号値と共に気象センサから得られた天候ラベルも示しており、気象センサから得られる観測値と整合性のとれた屋外画像が収集されることが分かる。

3.2 天候クラスに応じた屋外画像生成

まず、天候クラスによる変換が学習可能か検討する。2.2 節で学習する気象信号値推定器 W を天候クラス分類器とし、3.1 節の天候推定に用いた屋外画像データセット [13] で、天候クラスに応じた屋外画像の変換器 T を学習する。ただし、2:1:1 の

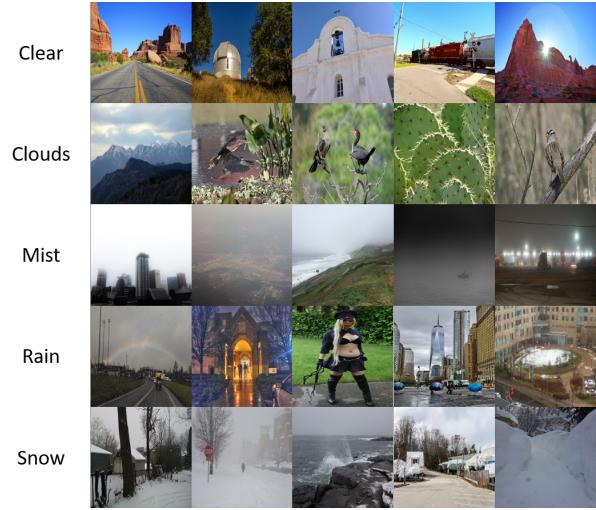


図 5 気象信号値と対応付けられた屋外画像例

入力 x 晴れ 曇り 雨 雪 霧



図 6 天候クラスを用いた変換結果(左から、入力画像 x 、条件 y をそれぞれ変更した時の生成画像 \tilde{x})

割合で学習用データ、テスト用データ、評価用データにそれぞれ分割した。

学習用データ 91,894 セットのうち、3 割の 27,568 セットを用いて、天候クラス分類器 W を学習した。このとき、条件 y は天候クラスの One-hot ベクトルとし、ネットワークには Resnet101 [15] を用いた。図 7 に学習した天候クラス分類器の分類結果を示す。

また、学習した天候クラス分類器と学習用データの屋外画

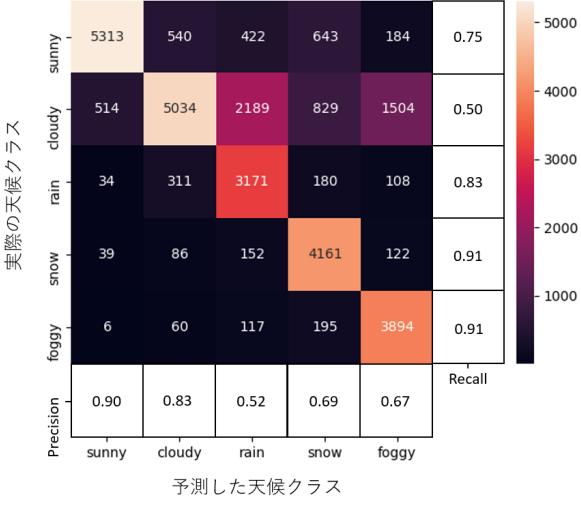


図 7 天候クラス分類器の分類結果

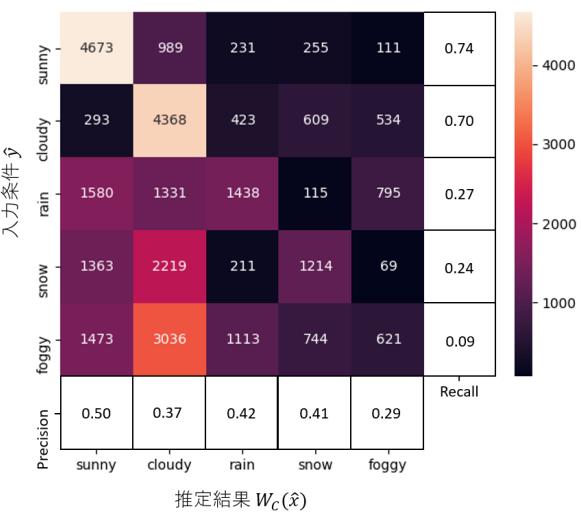


図 8 天候クラスに応じた屋外画像変換の精度評価

像を用いて、天候クラスに応じた屋外画像変換器 T を学習した。また、屋外画像変換器 T の判別器 D には SNGAN [16] の Projection Discriminator を用いた。変換器 T と判別器 D を用いて敵対的学習を行う際に用いたパラメータは、最適化手法は Adam で $\beta_1 = 0, \beta_2 = 0.999, wd = 5e - 6$ 、学習率は $lr = 1e - 4$ とした。また、 $\lambda_{con} = 1, \lambda_w = 1$ とした。さらにバッチサイズは 16 で、学習エポック数は 20 エポックとした。

図 6 に、異なる入力画像 x に対し、条件 y として晴れ、曇、雨、雪、霧の 5 つの天候クラスを与えたときの生成結果を示す。条件が晴れの場合は晴天、雪の場合は地面に雪が積もるというように、各天候クラスに対して視覚的に対応の取れた画像が生成された。

また、これらの気象状況に応じた屋外画像変換の定量的評価を行うために、評価用データにより学習した天候クラス分類器 W_C を用いる。テスト用データからランダムに選出したデータ対 $\{\hat{x}, \hat{y}\}$ の条件 \hat{y} を用いて、屋外画像変換を行う。変換後の出力である \tilde{x} から天候クラス分類を行い、変換の条件 \hat{y} と分類

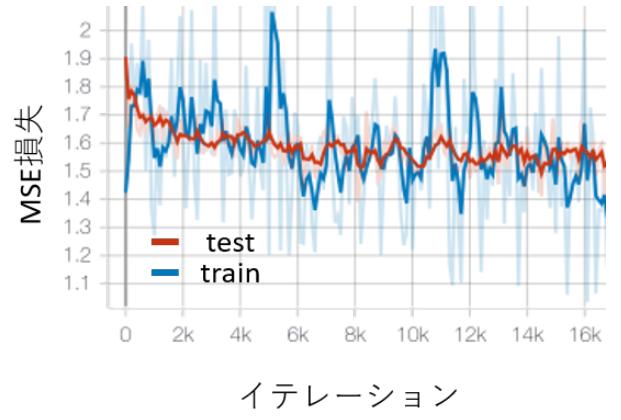


図 9 気象信号値推定器の MSE 損失

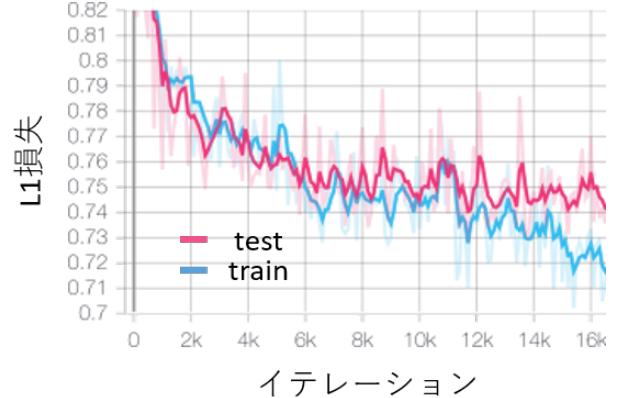


図 10 気象信号値推定器の L1 損失

結果 $W_C(\hat{x})$ から混合行列を作成し、各クラスに対する適合率 (Precision) と再現率 (Recall) を計算することによって、妥当性の評価を行う。評価用データを用いて、天候クラスに応じた屋外画像変換の精度を、図 8 に示す。

図 8 では、再現率をみると晴れ (sunny) と曇り (clouds) の天候ラベルによる変換精度が最も高いことがわかる。これは、晴れ、曇りのクラスの学習データ数が多いため、変換の学習が十分に行えたと考えられる。一方で、霧 (foggy) や雪の変換結果は、曇りと判定されることが多い。これは、雪、霧の変換結果が、曇りと視覚的に近い天候であるため、天候のクラス分類に失敗した可能性がある。

3.3 気象信号値に応じた屋外画像生成

3.1 節で収集したデータセットの中から、乖離度 L が $Th = 2.5$ 未満のデータ (102,893 セット) をラベル付きデータとし、それ以外をラベルなしデータとする。ラベル付きデータのうち、半分である 51,446 セットを学習データとし、残りの 51,447 セットをテストデータとして、気象信号値推定器 W を学習した。推定器 W の学習時、イテレーションによる平均二乗誤差 (Mean Squared Error; MSE) と L1 ノルム誤差を、それぞれ図 9 と図 10 に示す。

図 9 は、青線 (train) が学習時、赤線 (test) がテスト時の MSE 損失 L_{est} である。学習時の L_{est} は、データ拡張による影

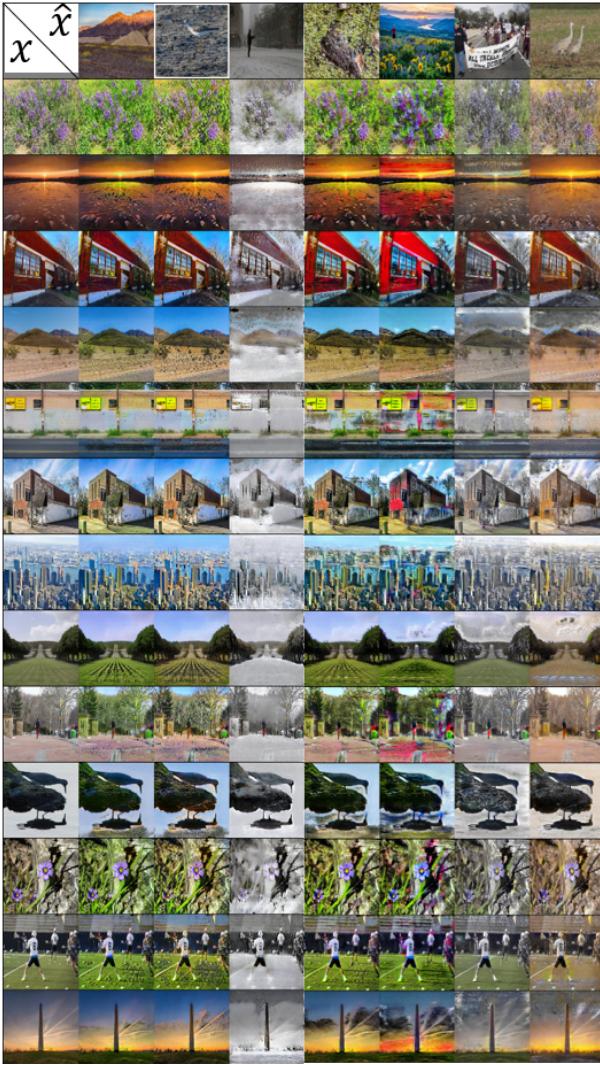


図 11 気象センサ値を用いた変換結果(左端:入力画像 x , 最上段: 任意地点の気象条件 \hat{y} と対応した屋外画像 \hat{x})

響で分散が大きくなっているが、テスト時の L_{est} は追従するよう損失が小さくなっているため、汎化しているといえる。しかし、 L_{est} は二次関数であるため、教師データ y と出力 $W(x)$ の差が 1 未満の場合では、変化が小さい。一方で、一次関数である図 10 の L1 ノルム誤差は、12,000 イテレーションを超える学習で、学習時とテスト時の誤差が大きくなっているのがわかる。このことから、10 エポック(約 160,000 イテレーション)で、十分汎化したとみなし、学習を終了した。

次に、学習用データと乖離度 L が $L \geq 2.5$ のデータのうち、屋外画像のみを用いて、屋外画像変換器 T を学習した。各パラメータは、上述の天候クラスによる変換器と同様の値を用いた。図 11 に、異なる入力画像 x を、ランダムに選択した別の画像 \hat{x} に対し気象信号値推定器で推定した $W(\hat{x})$ を条件 y として変換した結果を示す。それぞれオブジェクトの形状を維持しつつ、多様な条件 y に視覚的に対応した出力画像が得られた。

また、3.2 節と同様に、天候クラス分類器 W_C を用いて、生成された画像の妥当性の評価を行った。テスト用データからランダムに選出したデータ対 $\{\hat{x}, \hat{y}\}$ の条件 \hat{y} を用いて、屋外画像

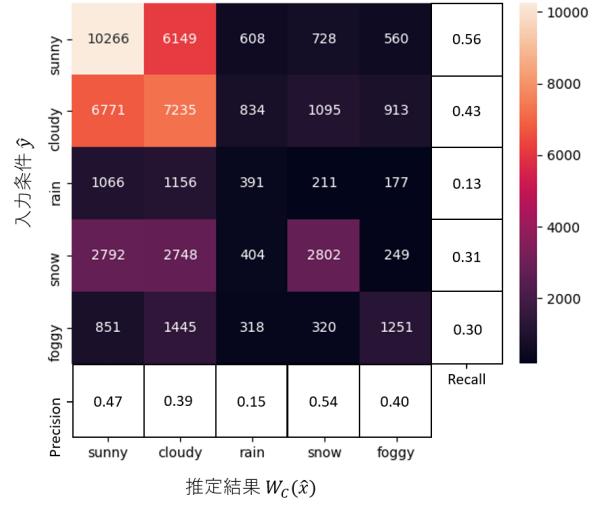


図 12 気象信号値に応じた屋外画像変換の精度評価

表 1 気象信号値に応じた屋外画像の変換誤差

気象センサ	雲量	気温	湿度	気圧	降雨量	風速	平均
L_{val}	0.270	0.356	0.360	0.153	0.407	0.432	0.330
$\sigma(L_{val})$	0.217	0.260	0.316	0.138	0.488	0.428	0.308
$L_{val}\sigma(Y_i)$	10.0	3.15	8.10	1.37	0.472	1.03	-

変換し、変換後の出力である \hat{x} から天候クラス分類を行い、変換の条件 \hat{y} と分類結果 $W_C(\hat{x})$ から混合行列を作成し、適合率(Precision)と再現率(Recall)を計算した。評価用データを用いて、気象信号値に応じた屋外画像変換の評価結果を図 12 に示す。図 12 では、天候クラスを入力条件に用いた場合と比べ、混合行列全体の相関が低い。しかし、雪と霧の変換では、適合率・再現率ともに天候クラスの精度を上回っている。また、晴れと曇りを互いに間違って分類しているため、それぞれの精度は天候クラスのものに及んでいない。これは、変換に雲量センサの信号値を用いているため、晴れと曇りの中間となる変換結果が多く存在するためと考えられる。この結果から、入力条件に天候クラスではなく、連続値である気象信号値を用いた場合でも、天候クラスを用いた変換と同程度の変換精度を得られたと考えられる。

さらに、評価用データにより学習した気象信号値推定器 W_E を用いて、 \hat{y} を入力条件としたときの、変換後の出力である \hat{x} から信号値推定を行うことで誤差を計算し、変換性能の評価を行った。誤差関数は、

$$\mathcal{L}_{val} = \mathbb{E}_{\hat{x} \sim p_g, y_c \sim \hat{p}_{data}} [|W_E(\hat{x}) - \hat{y}|] \quad (5)$$

で与えられる。テスト用データを用いて、式 5 を各センサごとに計算した結果を表 3.3 に示す。ここで、 $\sigma(Y_i)$ は学習用データに含まれる各気象信号値のデータ分散を表す。表 3.3 の L_{val} を各センサで比較すると、気圧の信号値がこの指標において最も変換誤差が低く、分散も少ない。一方で、風速に応じた屋外画像の変換誤差は高いことがわかる。これは、他の信号値と比べて、気圧の変動による視覚的な気象状況の変化は、高い相関性があると考えられる。その次に精度が高い雲量に関しても、

気温や湿度、降雨量の信号値に比べ、高い相関性があるといえる。また L_{val} は、各センサごとのデータ分散 $\sigma(Y_i)$ と積をとることによって、本来の信号値のスケールに戻すことができる。これらは、任意の屋外画像が各信号値によって変換された時の視覚的誤差を示しており、それぞれ雲量では $\pm 10[\%]$ 、気温は $\pm 3.15[^{\circ}C]$ 、湿度は $\pm 8.1[\%]$ 、気圧は $\pm 1.37[\text{hPa}]$ 、降雨量は $\pm 0.47[\text{mm/h}]$ 、風速は $\pm 1.03[\text{m/s}]$ の誤差範囲で変換できることがわかる。

4 まとめ

本研究は、任意の場所、時間の気象状況を表す屋外画像生成を目的に、気象センサ値を条件とする画像変換手法を提案した。同じ環境を観測した屋外画像と気象センサの信号値を、観測した気象状況の整合性も考慮して収集し、気象信号値推定器の事前学習に用いることにより、SNS から収集した大量の屋外画像のみを用いた画像変換器の半教師あり学習を実現した。実際に SNS から収集した多様な屋外画像に対し、任意の気象信号値という粒度の細かい条件に応じて適切な画像が生成されることを確認した。また、変換結果を定量的に評価することによって、天候クラスに応じた屋外画像変換は、晴れや雨、雪の天候クラスに対して高い精度の変換ができることがわかった。また、気象信号値に応じた屋外画像変換では、天候クラスの変換に比べ、入力条件に天候クラスを用いていないため、天候クラスに対する相関は低いが、雪や霧に対する変換精度は、天候クラスの精度と同程度の変換精度が得られた。さらに、信号値推定器を用いて、気象信号値に応じた屋外画像変換の誤差を計算することにより、本来の信号値のスケールで誤差範囲を求めることができた。

今後の課題として、教師あり学習を行った場合など、学習データや損失関数の変更した場合との比較により、提案する半教師あり学習のアプローチの有効性を検証する必要がある。

本研究の一部は、科学研究費補助金（基盤（C）19K12019、基盤（S）16H06302）の助成による。

文 献

- [1] OpenWeatherMap. <https://openweathermap.org/>.
- [2] ウェザーリポート. <https://weathernews.jp>.
- [3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.
- [4] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.
- [5] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [6] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*, 2016.
- [7] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [8] Guim Perarnau, Joost Van De Weijer, Bogdan Raducanu, and Jose M. Álvarez. Invertible conditional gans for image editing. *arXiv preprint arXiv:1611.06355*, 2016.
- [9] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 6, pp. 1452–1464, 2017.
- [10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *arXiv preprint arXiv:1505.04597*, 2015.
- [11] Xun Huang and Serge J. Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. *Proceedings of the IEEE International Conference on Computer Vision*, 2017.
- [12] Flickr. <https://www.flickr.com/>.
- [13] Wei-Ta Chu, Xiang-You Zheng, and Ding-Shiuan Ding. Image2weather: A large-scale image dataset for weather property estimation. *Proceedings of the IEEE Second International Conference on Multimedia Big Data*, pp. 137–144, 2016.
- [14] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.
- [15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *arXiv preprint arXiv:1512.03385*, 2015.
- [16] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018.