

気象データから雲の形を予測するモデルの作成

水野 大運^{*}、中田 恭子^{*}、江越 航^{**}

概要

日常的な雲の形状を対象に、数値予報格子点資料 (GRIB2 データ) から機械学習を用いて全天カメラ画像を推論・検索する手法の確立を試みた。特徴量抽出器として EfficientNetV2-L を採用し、大阪市立科学館の全天画像群 (2019 年通年、計約 1,760 枚) から 1,280 次元の特徴量ベクトルを正解として抽出した。気象データからこの特徴量ベクトルを直接予測するネットワークを構築し、コサイン類似度と平均二乗誤差を融合したハイブリッド損失関数を導入してネットワークを学習させた。

実験では、画像範囲の絞り込みや太陽位置情報の付与、圧縮次元数や参照地点数の比較検討を行った。さらに、精度の向上を目指して、プーリング層の工夫や発展的なデータクレンジングを実施した。これによりアルゴリズムの工夫における限界とその要因について多角的な考察を行った。

1. はじめに

気象予報における数値シミュレーションは、降水や台風の進路など防災・産業を目的とした物理量の計算が主な対象であり、特定の地点から見上げる日常的な空の景観を予測する試みは極めて少ない。また、雲の発生プロセスを完全に再現することはいまだ困難である。そこで本研究では、GRIB2 形式の気象データからニューラルネットワークを用いてその時刻の全天画像の特徴量ベクトルを予測し、画像データベースから類似した景観を特定・提示する近傍検索型の景観予測推論システムを確立することを目的とした。

2. 研究手法およびデータセット

2-1. 使用データと初期クレンジング

入力データには、数値予報格子点資料の GRIB2 形式データ (これらのデータは京都大学生存圏研究所が運営する生存圏データベースによって収集・配布されたものである (<http://database.rish.kyoto-u.ac.jp>)) を使用した [1]。抽出要素は、下層・中層・上層雲量、地表面気圧、東西・南北風速、気温、相対湿度とした。風速、気温、相対湿度は、雲の発生高度に近い 500, 700, 850, 925 hPa の各気圧面の値を採用した。

画像データには、大阪市立科学館にて撮影された

全天カメラ画像 (2019 年通年、6:00~18:00、3 時間ごと) を用いた [2]。深夜画像や機材エラー画像などを除いた計約 1,760 枚からなるデータセットを構築した。

2-2. 画像特徴量の抽出と推論システム

画像情報を低次元の特徴量ベクトルへ圧縮する特徴量抽出器として、ImageNet を学習済の畳み込みニューラルネットワーク (CNN) である EfficientNetV2-L を採用した。分類層を排した状態で画像を入力し、平均プーリング層から 1,280 次元の特徴量ベクトル (正解ベクトル $\mathbf{b}$) を取得した。

一方、構築した予測推論システムに気象データを入力して出力される推論結果を予測ベクトル $\mathbf{a}$ とする。学習に用いていない気象データに対する推論時には、予測ベクトル $\mathbf{a}$ を用いて、画像データベース内の全正解ベクトル $\mathbf{b}$ との間で類似度を算出・比較し、最も類似度の高い画像を特定・提示する。本研究ではこのような近傍検索システムを構築した。

2-3. 損失関数

ネットワークの学習に用いる損失関数として、予測ベクトルの向きを揃えるコサイン類似度と、各要素の大きさを揃える平均二乗誤差 (MSE) を組み合わせたハイブリッド損失関数 (1) を定義した。

^{*}青山学院大学理工学部 2myod578@gmail.com

^{**}大阪市立科学館 egoshi@sci-museum.jp

$$Loss_{cos} = -\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = -\frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}}$$

$$Loss_{MSE} = \frac{1}{n} \sum_{i=1}^n (a_i - b_i)^2$$

$$Loss = Loss_{cos} + \alpha \cdot Loss_{MSE} \quad (1)$$

α の値を網羅的に変えて損失関数の値を検討した結果から、 $\alpha = 0.85$ を最適値として採用した。

2-4. 検証モデルの仕様

予測精度を段階的に評価するため、時系列データの学習法(季節別/通年)と画像のノイズ対策(トリミング半径:標準/短)を組み合わせた 3 つのモデル(表 1)を構築した。

表 1. モデル 1～3 の仕様

	特徴量 次元数	過去 データ 参照	参照 地点数	時系列データ の扱い (学習時)	トリミング 半径	太陽の 位置 情報
モデル 1	128	○	2	季節別	標準	×
モデル 2	128	○	2	季節別	短	×
モデル 3	1,024	○	2	通年	標準	○

【モデル 1】(時系列: 季節別、半径: 標準)

データを四季に分割して学習させ、全天画像は標準半径で円形トリミングした。主成分分析 (PCA) により画像特徴量を 128 次元に縮約して学習を行った。

【モデル 2】(時系列: 季節別、半径: 短)

全天画像の周囲に写り込む建物の影響を排除するため、トリミング半径を短くして天頂付近の空の領域のみを切り出した(図 1)。

【モデル 3】(時系列: 通年、半径: 標準)

一年中安定した推論を行うため、全季節のデータを一括して扱う通年学習を採用した。標準半径でトリミングし、PCA を 1,024 次元へ拡張した。また、太陽の位置(高度・方位)を入力特徴量に追加した。

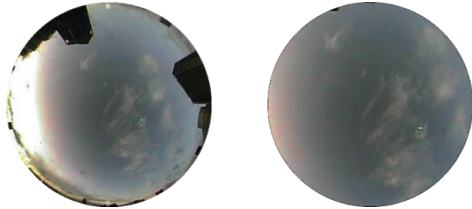


図 1. 標準半径(左)と短半径(右)でトリミングした例

なお、すべてのモデルにおいて、共通の基盤である参照地点数およびその他の仕様は次の通りである。

① 参照地点数

気象データの空間解像度を最適化するため、撮影地点に近い A、B の 2 地点を採用した(図 2)。

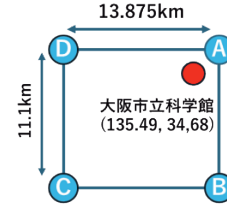


図 2. 参照地点のイメージ図

② 過去データの参照

時間的な連続性をモデルに学習させるため、撮影時刻だけでなく、3 時間前および 6 時間前の気象データを入力特徴量に加えた。その際、エラーデータによる歯抜けや日付を跨ぐ等により参照する過去時刻のデータが存在しない場合は、当時刻の気象データをもとに補間を行った。

③ 時系列ベースのデータ分割

モデルの汎化性能を評価するため、ランダムな抽出ではなく時間軸に沿ったホールドアウト分割を採用した。通年学習(モデル 3)の場合は、データ全体(1,761 件)のうち、各月の 1～24 日を学習データ、各月の 25 日以降(357 件)をテストデータとした。季節別モデル(モデル 1、2)の場合は各季節の後半約 2 割をテストデータとし、モデル 3 と同じ約 8:2 にデータを分割した。

2-5. 評価指標

各モデルの精度を 2 つの指標によって評価した。

(1) Recall@K

モデルは、入力された気象データから推論した予測ベクトル $\mathbf{a}$ に対して、画像データベース内の各正解ベクトル $\mathbf{b}$ との類似度(ベクトル間の距離としてユークリッド距離を正規化するとコサイン類似度と等価)を算出する。Recall@K は、類似度の高い順に候補画像を K 枚提示した際、その K 枚の中に正解画像が含まれる確率を示す指標である。学習データを入力データとした場合は、正解画像が学習済の画像データベースに存在する。テストデータを入力データとした場合は、入力に対応する画像が学習済データベースに存在しないので、テストデータの各画像の正解ベクトル $\mathbf{b}'$ と学習データの各画像の正解ベクトル $\mathbf{b}$ の類似度を算出しておき、正解ベクトル $\mathbf{b}'$ と最も類似度の高い正解ベクトル $\mathbf{b}$ に対応する画像を正解画像とする。本研究では K の値を段階的に拡張し、モデルの実用性を評価した。

(2) 平均検索順位

予測ベクトル $\mathbf{a}$ に基づく近傍検索の結果、対象時刻の正解画像が検索結果全体の第何位にランクインしたかをすべての学習データ、テストデータに対して平均した数値である。この値が1に近いほど、気象データから空の景観を高い精度で特定できていることを意味する。

3. 実験結果および考察

3-1. 構築したシステムの概要

各時刻における気象データと、全天画像および正解ベクトル $\mathbf{b}$ とをペアリングした画像データベースを作成し、これを用いてモデル1～3をそれぞれ学習させた。学習後のモデルに入力データとして気象データを与え、得られた予測ベクトル $\mathbf{a}$ との類似度の高い正解ベクトル $\mathbf{b}$ を近傍検索し、正解ベクトル $\mathbf{b}$ に対応する全天画像を出力させた。これにより、気象データから全天画像を予測するシステム(モデル1～3)を構築した。

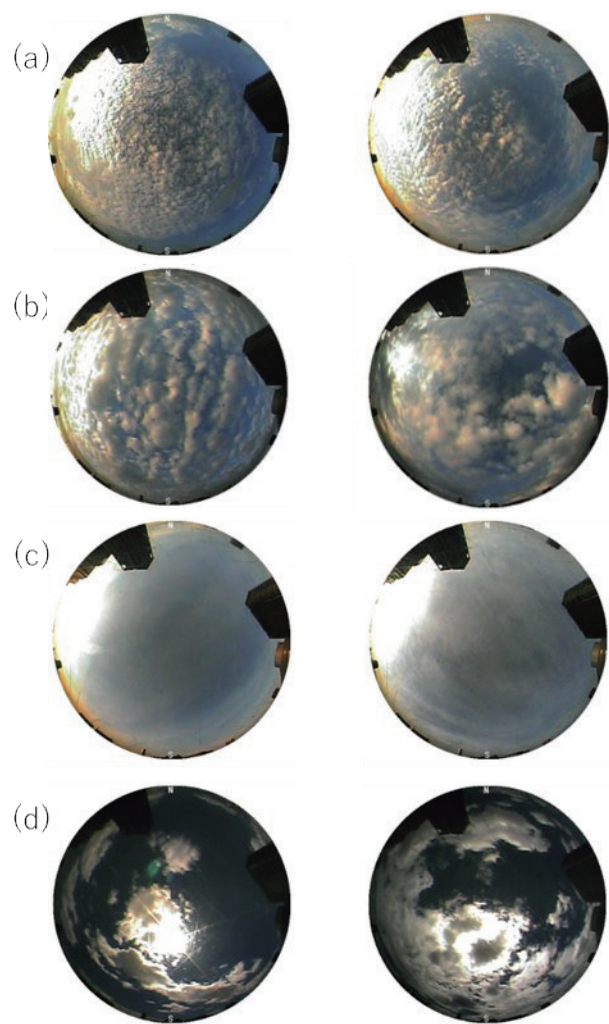


図 3. 妥当な推論結果が得られた例
(左: 正解画像 右: 推論画像)

このシステムを使って適切に推論できた例を図3に示す。左は気象データとペアリングされた正解画像であり、モデルの学習には用いていない。右は、気象データを入力して得られた推論画像である。正解画像は(a)鱗雲(巻積雲)、(b)羊雲(高積雲)、(c)巻層雲、(d)積雲、と大別され、これらに対して妥当な推論画像が得られたことがわかる。

このシステムは、緯度・経度と GRIB2 気象予報データを入力することで、大阪市立科学館だけでなく、任意の地点に応用することができる。その例として、相模原市淵野辺における推論結果を図4に示す。正解画像(気象予報データに対応する日時に青山学院大学相模原キャンパスにて撮影したもの)に対しておおむね妥当な推論結果が得られた。また、他の地点に対しても、公開されているライブカメラ画像と推論結果を比較して、妥当な推論結果が得られることを確認した。



図 4. 相模原市の例(左: 正解画像 右: 推論画像)

このように、GRIB2 形式の気象予報データを入力すると全天画像を出力する予測推論システムを構築し、おおむね妥当な推論結果が得られることを確認できた[3]。しかし、正解画像と特徴の異なる画像が出力される場合もあった。以下の節では、学習と推論の精度評価と、精度向上に向けた検討の結果を報告する。

3-2. 主要モデルにおける学習曲線の評価

モデル1～3に対して学習プロセスにおける訓練損失(Train Loss)および検証損失(Val Loss)の推移を調べた。四季別に学習したモデル1、2(いずれも夏の例)と通年学習したモデル3について、学習曲線を示す(図5)。

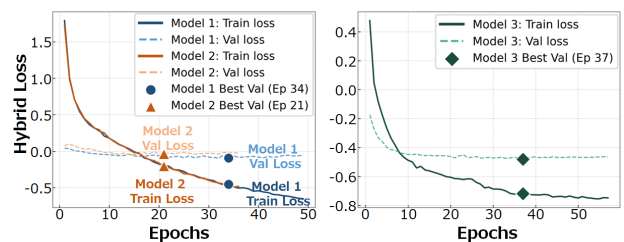


図 5. モデル1,2 夏(左)とモデル3(右)の学習曲線

モデル 1～3 のいずれも学習の進行に伴って Train Loss は単調減少するものの、Val Loss は早い段階で減少が停滞し、以降は横ばいとなる典型的な過学習の挙動を示した。モデル 3 は、過学習の傾向を示すものの、Train Loss (-0.7165) と Val Loss (-0.4787) がモデル 1、2 と比べて小さな値を示した。

3-3. 主要モデルの推論性能比較と考察

モデル 1～3 の推論精度を表 2 に示す。各モデルの仕様が推論性能に与えた影響について考察する。

表 2. モデル 1～3 の推論結果

カテゴリ	評価項目	モデル 1	モデル 2	モデル 3
損失関数の最適値	Train Loss	-0.4527	-0.4365	-0.7165
	Val Loss	-0.1017	-0.0057	-0.4787
学習データに対する精度	Recall@1 (%)	45.62	50.32	41.88
	Recall@10 (%)	71.35	73.55	77.99
	Recall@100 (%)	91.30	92.14	97.72
	平均検索順位	18.15	16.84	10.57
未知データに対する精度	Recall@1 (%)	21.35	18.87	2.24
	Recall@10 (%)	52.81	54.93	15.69
	Recall@100 (%)	86.80	85.07	49.58
	平均検索順位	28.40	29.49	67.69

(1) トリミングに伴う光学的な疑似適合

全天画像の円形トリミング半径を標準(モデル 1)から短く変更(モデル 2)した結果、学習データに対する Recall@1 は 45.62 % から 50.32 % へと向上し、平均検索順位も 18.15 位から 16.84 位へと改善した。しかし、未知のテストデータに対しては Recall@1 が 18.87 %、平均検索順位が 29.49 位と低下した。

この原因として、画像を中心付近に切り詰めた結果、「昼間の曇り空」と「夜間の快晴の空」など、本来は異なる画像から、類似した特徴量ベクトルが抽出された可能性が考えられる。実際、昼間の曇りと夜間の快晴の画像において、類似度が高くなった例が確認された。モデル 2 の推論が失敗した例を図 6 に示す。正解画像が夜 6 時の快晴であるのに対し、推論画像は昼間の曇りである。モデル 2 では、このような見かけ上の学習精度上昇(疑似適合)が起きていたと推察される。

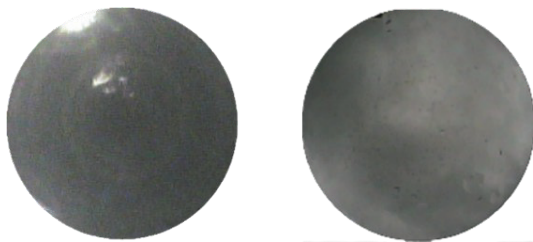


図 6. モデル 2 の推論失敗例
(左:正解画像 右:推論画像)

(2) 太陽位置の追加に伴う過学習とその考察

Train Loss と Val Loss が最小になるようにモデル 1 と 2 を改良したモデル 3 において、学習データに対する平均検索順位や Recall@100 が向上した。しかし、未知のテストデータに対する平均検索順位や Recall@1 は低下した。この原因としては、まず 1,024 次元という高次元空間における過学習が考えられる。

また、太陽が含まれる画像では、雲の形状や質感が似ていなくても類似度が高くなる現象が確認された(図 7)。モデル 3 において推論が失敗した例(図 8)からも、太陽の影響が示唆される。

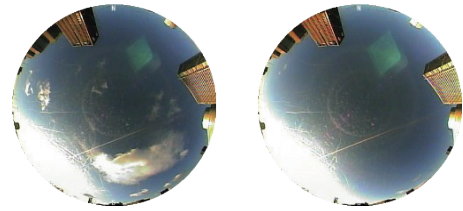


図 7. 基準画像(左)とコサイン類似度 Top1 画像(右)

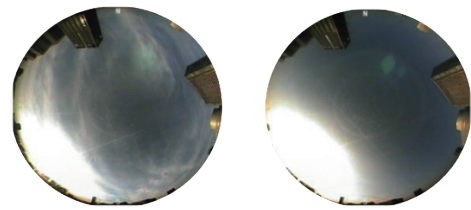


図 8. モデル 3 の推論失敗例
(左:正解画像 右:推論画像)

3-4. 精度の頭打ちに対するパラメータ感度分析

モデル 3 における過学習の緩和を目指し、特徴量ベクトルの次元数と気象データの参照地点数(図 2)を段階的に変化させたパラメータ感度分析を実施した。

(1) 圧縮次元数の検討

PCAを用いて特徴量ベクトルの次元数を 128 から圧縮なし(1,280)へ段階的に変化させ、学習プロセスへの影響を評価した(表 3)。次元を拡張するほどモデルの表現量が増大するため、学習時の Train Loss は低下した。しかし、未知データに対する Val Loss は約 -0.4 となり、過学習の緩和は限定的であった。

(2) 参照気象地点数の検討

気象予報値の空間解像度が学習に与える影響を評価するため、入力する GRIB2 データを大阪市立科学館に近接する 1 地点(A)、2 地点(A, B)、4 地点(A ~ D)と変化させた(図 2)。

損失関数の値を比較した結果(表 3)、地点数の変更が Train Loss と Val Loss の乖離に与える影響は小さいことが判明した。格子点予報情報の入力としては、標準 2 地点の段階で説明変数としての情報量が実質的に飽和していることがわかった。

以上の結果が示す通り、次元数や参照気象地点数などのパラメータ調整だけでは、未知データに対する過学習傾向および損失関数の減少停滞を解決することはできなかった。これは、問題の原因がパラメータ設定ではなく、太陽の直射ノイズや、予報値と画像の不整合など、データセット側にあることを示唆している。

表 3. 次元数・参照地点数に対する損失関数の推移

参照 地点数	指標	画像特徴量の圧縮次元数				
		128	256	512	1024	1280
1	Train Loss	-0.6025	-0.6317	-0.6817	-0.6424	-0.6456
	Val Loss	-0.4206	-0.4544	-0.4693	-0.4810	-0.4862
2	Train Loss	-0.6520	-0.6516	-0.6642	-0.6575	-0.6697
	Val Loss	-0.4208	-0.4550	-0.4698	-0.4794	-0.4804
3	Train Loss	-0.6430	-0.6431	-0.6780	-0.6875	-0.6757
	Val Loss	-0.4190	-0.4544	-0.4703	-0.4751	-0.4786

3-5. 精度改善に向けたデータクレンジングの強化

パラメータ調整によって過学習の傾向が改善しなかったため、データの純度そのものを高める発展的なデータクレンジングを試みた。

(1) 太陽画像全除去の背景: プーリング層の工夫

雲の微細な質感を捉えるために、EfficientNetV2-Lの最終層付近における特徴量抽出アルゴリズムを改良した。EfficientNetV2-Lでは、特徴量マップの各領域内の値を1つの値に圧縮する際に平均値を用いる平均プーリングが採用される。本研究では、画素値のばらつきを抽出する標準偏差プーリングを平均プーリングに結合し、計 2,560 次元の複合特徴量ベクトルを生成・学習させるシステムを考案した。

また、全天画像において最大の輝度ノイズとなる「太陽の直射光」への対策として、太陽の領域を「輝度 0」として黒く塗りつぶすマスク化処理を検討した。しかし、マスク化された境界においては、青空や雲という高輝度画素と輝度 0 の画素が隣り合うため、画素値の標準偏差が極端に大きくなる。その結果、雲の形状ではなく、季節や時刻によって位置が変わるマスクの形状を、最大の特徴としてモデルが誤学習するというトレードオフに直面した。

以上の理由から、単なる部分マスク化は無効であると判断し、太陽が大きく映り込んでいる画像自体をデータベースから除去することとした。

(2) 太陽を含む画像の除去

太陽を含む画像を除くため、輝度が 245 以上の画素を含むすべての画像を削除した(図 9)。これにより、純粋な空領域が担保された 503 枚からなる「Clean503 データセット」へと絞り込みを行った。

太陽が映った画像を削除すると、データの時間的連続性は失われる。そこで、過去データの参照(2-4

節)を撤廃し、Clean503 データセットの各時刻の気象データのみを入力として、再度学習を行った。

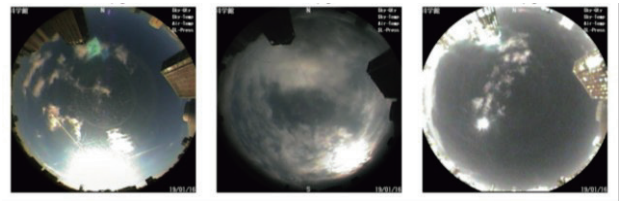


図 9. 輝度 245 以上の部位を含む画像の例

また、輝度が極端に大きな画素は、太陽が映っている部分だけでなく、日光が建物に当たっている部分にも生じる。短い半径でトリミングしたモデル 2 で、昼間の曇りと夜の晴れの類似度が高くなったことを鑑みて、エッジを検出して建物のみを適切にマスクした画像を作成した(図 10)。建物の位置はどの画像でも変わらないため、このマスク化が、隣り合う高輝度な空や雲よりも優先されて誤学習される可能性は小さいと考えられる。そこで、Clean503 データセットに建物のマスク化とプーリング層の工夫を併用して学習の精度向上を図った。



図 10. 建物のマスク化

(3) 画像と整合しないデータの除去

データセットを精査すると、気象データと全天画像の特徴が整合しないデータが確認された。そこで、気象データに含まれる雲量と全天画像から読み取れる雲量が整合しないデータの除去を検討した。手法としては、Ridge 回帰モデルを構築して、特徴量ベクトルから雲量を簡易的に予測した。算出された雲量と GRIB2 の予報値の乖離があるデータを AI に指摘させ(図 11)、目視で確認した上で太陽を除去したデータからそれらを排除した。このように雲量不整合データを除去した 474 件のデータセットを用いて、再度学習を行った。

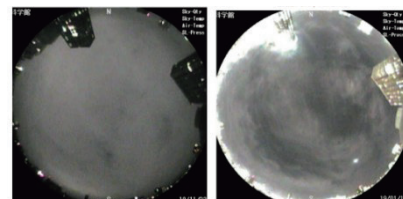


図 11. 全天画像と気象データに雲量の乖離のある例

左: GRIB2 雲量 10 % 簡易予測雲量 90 %

右: GRIB2 雲量 0 % 簡易予測雲量 51 %

(4) データクレンジングの結果と考察

データクレンジング後の学習曲線（図 12）と推論結果（表 4）を示す。

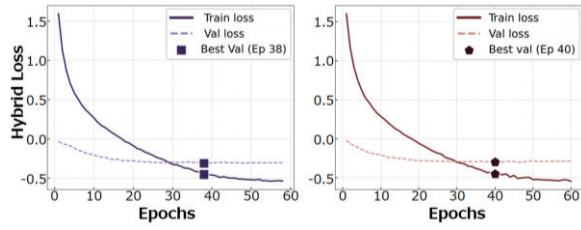


図 12. データクレンジング後の学習曲線

（左：太陽画像除去 右：雲量不整合データ除去）

表 4. データクレンジング後の推論結果

カテゴリ	評価項目	太陽画像除去	雲量不整合データ除去
損失関数の最適値	Train Loss	-0.4443	-0.4495
	Val Loss	-0.3109	-0.3019
学習データに対する精度	Recall@1 (%)	39.70	43.55
	Recall@10 (%)	76.88	80.65
	Recall@100 (%)	98.49	98.92
	平均検索順位	10.22	8.17
未知データに対する精度	Recall@1 (%)	1.90	0.00
	Recall@10 (%)	17.14	18.63
	Recall@100 (%)	60.95	58.82
	平均検索順位	59.95	60.61

太陽を含む画像を除去すると、Train Loss (-0.4443) と Val Loss (-0.3109) の差に改善が見られた。また、雲量が整合しないデータを除去すると、学習データに対する平均検索順位はわずかに向上した。しかし、未知データに対する Recall@K や平均検索順位は、いずれの場合も、モデル 3 の結果と同程度となった。

このようにデータクレンジングによる精度の向上が限定的であった背景には、データセットにおける本質的な限界が絡んでいると考える。

人間が数字を見て判断・クレンジングできるのは、雲量のように画像から視認・算出できる指標のみである。しかし、気象データに含まれる風速や上空の湿度など、全天画像から視認できない「不可視の気象変数」が、撮影地点における局所的な雲の形状や質感と整合しているとは限らない。この不可視な変数の潜在的ミスマッチが、データクレンジングによる精度向上を阻む最大の障壁であると考えられる。

その他の要因としては、CNN による画像の特徴抽出の限界が挙げられる。本研究で試みた平均プーリングや標準偏差プーリングは、画像全体の輝度トーンやコントラストの激しさを圧縮・数値化するには有効であると考えられる。しかし、これらが雲の微細な形状や質感を完全に捉えているとはいえない。また、全天

画像においては、天頂付近と地平付近とで雲のスケールが異なるため、これが CNN による特徴抽出を困難にした可能性もある。さらに、より本質的な問題として、気象予報データに含まれる風速、雲量、気温などの変数だけでは雲の形状を予測しきれない可能性がある。

以上を踏まえ、精度向上に向けた方策を提案する。

第一は、特徴量ベクトルの質を向上させる撮影法の検討である。たとえば太陽の直射光を回避しやすい横向き固定カメラによる局所景観撮影は検討に値する。

第二は、画像と整合した気象データを用いた上で、局所的な気象変数のみでなく、気象トレンドや大気の動的な状態を示す環境変数を入力データに加えることである。これらのアプローチが、本システムの景観予測能力の向上につながると考える。

4. おわりに

本研究では、GRIB2 形式の気象予報データから局所的な雲の景観を全天画像の形で出力する検索型の推論システムを構築した。このシステムは任意の地点に応用できる。精度に改善の余地があるものの、おおむね妥当な推論結果が得られることを確認した。

過学習の緩和と推論精度の向上を目指し、圧縮次元数や参照地点数に対する感度分析、特徴量プーリング層の改良、太陽画像や雲量不整合データを排除するデータクレンジングを行った。推論精度の向上には至らなかったものの、その背景に、風速や上空の湿度などの目視不可能な気象変数と局所景観の潜在的ミスマッチや CNN による特徴抽出の限界が関わっている可能性が示唆された。

今後の精度向上に向けた指針として、太陽光の影響を軽減する撮影法の検討や、天候トレンドを補う環境変数の導入が挙げられる。データ純度の向上と説明変数の多角化を融合させるアプローチが、気象予報値から雲の景観を推論する本システムの改善につながると考える。

参考文献

- [1] 京都大学生存圏研究所, 生存圏データベース <http://database.rish.kyoto-u.ac.jp>
- [2] 江越 航, 大阪市立科学館研究報告, 33, p.1-4 (2023).
- [3] 水野 大運, 卒業論文, 青山学院大学(2026).