

全天球画像のデータ収集と雲形と状態判定

Collecting whole sky images and classification of cloud genera and conditions

森川 優*¹ 中西 波瑠*¹ 稲村 直樹*⁵ 近藤 伸明*⁵ 小渕 浩希*⁴ 大澤 輝夫*²
 Yu MORIKAWA Haru NAKANISHI Naoki INAMURA Nobuaki KONDO Hiroki OBUCHI Teruo OHSAWA
 松原 崇*³ 上原 邦昭*³
 Takashi MATSUBARA Kuniaki UEHARA

*¹株式会社神戸デジタル・ラボ
 Kobe Digital Labo Inc

*²神戸大学大学院海事科学研究科
 Graduate School of Maritime Sciences, Kobe University

*³神戸大学大学院 システム情報学研究科
 Graduate School of System Informatics, Kobe University

*⁴スカパー JSAT 株式会社
 SKY Perfect JSAT Corporation

*⁵バニヤン・パートナーズ株式会社
 BANYAN PARTNERS Inc

Maritime meteorological observation is critical for a safe voyage, and general ships are required in Japan to report the observations to parties concerned. Since it is difficult to recognize the meteorological conditions for non-experts, the demand of automatic recognition arises. Many studies have tackled the classification of cloud genera and the regression of cloud cover. However, less attention has been paid for cloud conditions. Thus, we developed a machine learning system for classification of cloud conditions. We first developed a dedicated equipment for photographing whole sky images and collected data samples. Then, we tagged cloud genera and conditions in each cloud layer (high, middle, and low). Using the dataset, we built a deep convolutional neural network to classify the cloud genera and conditions via fine-tuning ResNet50. The network achieved accuracies higher than 0.9 for both cloud genera and conditions.

1. はじめに

現在、一般船舶では3時間毎に海上の気象観測を行い、「船舶気象観測表」を気象庁へ報告することが気象業務法で定められている。[5] 海上気象観測の報告はこれまで電報にて実施されていたが、気象測器の開発や通信インフラの発達によって自動測定およびデータ通信が可能となり、気象観測や報告の自動化が進められている。しかし、雲の観測だけは従来の気象測器での測定が困難であり、乗組員による目視で観測されているが正確性や人的負荷の問題が存在する。

このように雲の観測はまだ研究の途にあるが、雲の自動観測を行うにあたって解決しなければならない3つの課題が存在している。まず第1に、船上に設置して長期の観測に耐えるデバイスの開発が必要である。第2に、収集された画像を自動的に分析して雲形・状態や雲量を測定する手法の開発が必要である。また、近年の畳み込みニューラルネットワーク(CNN)の発達を見るに深層学習の利用が有望であるが、雲の観測ではまだ研究の域を出ていない。そもそも学習に用いることができる雲形のデータセットは存在しているが、日本の海上気象観測で求められている、雲の状態と雲量に関するラベル付けされたデータセットは存在していない。これが第3の課題である。本論文ではこれら3つの課題にアプローチする。

先行研究において、全天球画像における雲形に関しては高い識別率で識別できている[1, 2, 3]。しかし、状態に関しては先行事例がなく、また実際の現場では下層・中層・上層にお

る雲形と状態を報告しなければならないため、先行事例において作成された学習器[1, 2, 3]を今回のシステムに適用することはできない。このため、雲形と状態を同時に判定する学習器が必要である。

本研究では、オープンデータとして存在しない雲形のみではなく、状態までラベル付けされたデータセットの構築を行い、雲形と状態の学習器をCNNによって構築するアプローチを取る。

2. 全天球画像のデータ収集

データセットを開発するにあたって、撮影デバイスが存在していないことから、撮影デバイスのプロトタイプ開発を行っている。図1は撮影デバイス本体の写真で、図2は実際に撮影デバイスを設置した船舶、図3は撮影デバイスを防水ケースに収納し船上に設置した際の写真である。



図1: デバイス稼働実験時

連絡先: 稲村 直樹, バニヤン・パートナーズ株式会社,
〒650-0031 神戸市中央区東町116-2 オールドブライビル 7F, 078-335-5955, 078-335-5431,
inamura@banyanpartners.co.jp



図 2: 神戸大学大学院海事科学研究科所属・深江丸



図 3: 撮影デバイス設置時

当初は、デバイスとしてコダック アクションカメラ PIXPRO SP360 4K (以下アクションカメラ) 単体で撮影を行う予定であった。しかし、アクションカメラ単体ではバッテリーの容量が一日の撮影を行うには足りず、その上、撮影時の熱により 30 分程度で撮影不可能になってしまう問題点が存在する。

そのため、アクションカメラにデスクトップ PC 用の水冷キットを接合することによって冷却し、バッテリーに 12 万 mAh の物を用意し接続をしている。また、デバイスの撮影タイミングやデバイスの稼働監視、カメラの再起動装置等の制御を Raspberry Pi 3 Model B で行っている。結果として、図 1 のような大きなデバイスとなってしまっている。図 2 の深江丸に搭載することを考慮すると、潮風や風雨による防水加工されていない部分へのダメージが大きいことは明白である。このため、図 3 のように防水加工されたケースを改造し、デバイスを収納して撮影を行っている。定期的な保守等は必要ではあるが、安定した撮影を行うことに成功している。

全天球画像収集のため、撮影デバイスを用いて構築した自動撮影システムを、神戸大学大学院海事科学研究科所属の深江丸に搭載して 1 分毎の撮影を行っている。1 分毎の写真だけでは写真の部分の雲や、重なった部分の雲が判定できないため、10 分毎に 1 度判定することによりラベル付けを行っている。また、夜間は人間の目では判定できないため、ラベル付けを行っていない。

ラベルとしては、世界気象機関 (WMO) により定められた 10 基本形と国際式天気記号 [5] における状態のラベル付けを行っている。雲の層ごとに雲形と状態のラベル付けを行ったため、一つの画像に対してラベルは 6 種類存在している。表 1 は層ごとに出現する雲形と状態のラベルの範囲 (0 から 9 まで) を示している。

表 1: 雲形と状態

層	雲形	状態
下層	積雲	0~9
	層積雲	
	積乱雲	
	層雲	
中層	高積雲	0~9
	乱層雲	
	高層雲	
上層	巻雲	0~9
	巻積雲	
	巻層雲	

今回は深江丸において撮影できなかった、もしくは枚数が少なすぎる雲形と状態が存在したため、学習の際はこれらを除外し表 2 と表 3 の雲形と状態を対象にして学習を行っている。

表 2 と表 3 を見ると、雲形と状態の枚数が揃っていない状況にある。この理由としては、撮影した範囲が限られていることや、季節による雲の出現の違いが理由としてあげられる。撮影開始の時期が秋頃からだったため、夏場によく現れる積乱雲や乱層雲を撮影できなかったことや、上層の雲が出現しやすい海域での撮影が出来ていないことにより、枚数に偏りが出てしまっている。

表 2: 雲形

層	雲形	図	枚数
下層	積雲	図 4a	388 枚
	層積雲	図 4b	120 枚
	層雲	図 4c	67 枚
中層	高積雲	図 5a	257 枚
	高層雲	図 5b	336 枚
上層	巻雲	図 6a	118 枚
	巻積雲	図 6b	91 枚
	巻層雲	図 6c	73 枚

表 3: 状態

層	状態	図	枚数
下層	0		640 枚
	1	図 4d	332 枚
	4	図 4e	116 枚
	7	図 4f	123 枚
中層	0		257 枚
	1	図 5c	134 枚
	2	図 5d	219 枚
	3	図 5e	113 枚
	7	図 5f	89 枚
上層	0		933 枚
	1	図 6d	86 枚
	2	図 6e	24 枚
	7	図 6f	30 枚
	8	図 6g	24 枚
	9	図 6h	91 枚

雲形に「なし」がない理由としては、先に状態を識別してから雲形を識別する方法をとったため、雲形の「なし」を学習させていない。このような方法を取った理由としては、2 回判定することにより整合が取れない場合 (例: 下層に雲が存在するにも関わらず雲形はなしと識別) を回避するためである。結果として、計 6 つの学習器を作成し、識別システムの実装を試みている。

図 4~図 6 では実際に撮影し、ラベル付けされた雲の画像の例を示している。雲ごとに十分な差異があることが明らかであるが、図 5c と図 5d のように人間の目では区別がつかない雲も存在している。また 3 層分の雲を判別することから別の層の雲が重なり合う事例が発生するため、ラベル付けの作業はさらに困難である。このように、雲自体の判別が観測未経験者では困難である上、雲の層ごとの高さの判別はさらに困難であるため、専門に教育された人材が必要である。

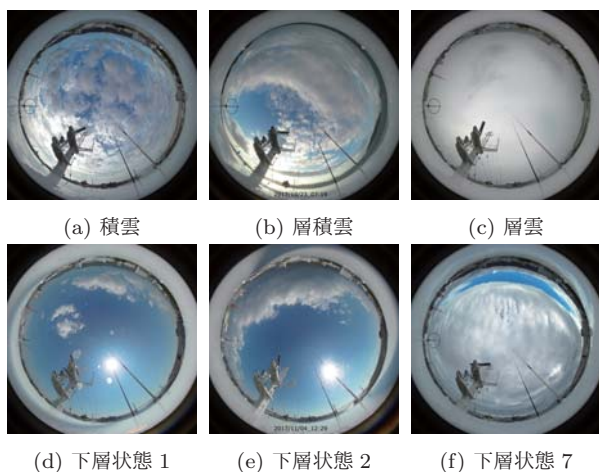


図 4: 下層雲形・状態

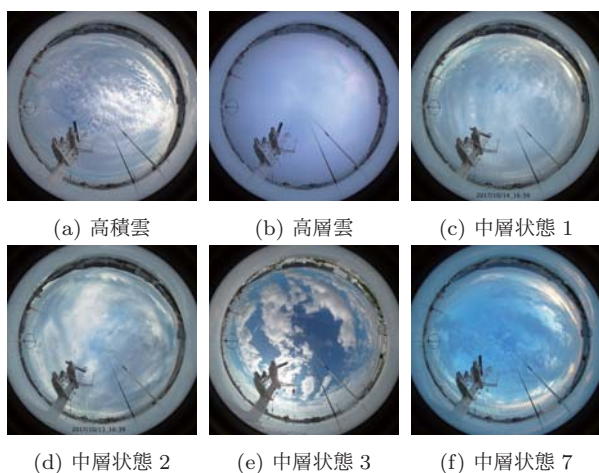


図 5: 中層雲形・状態

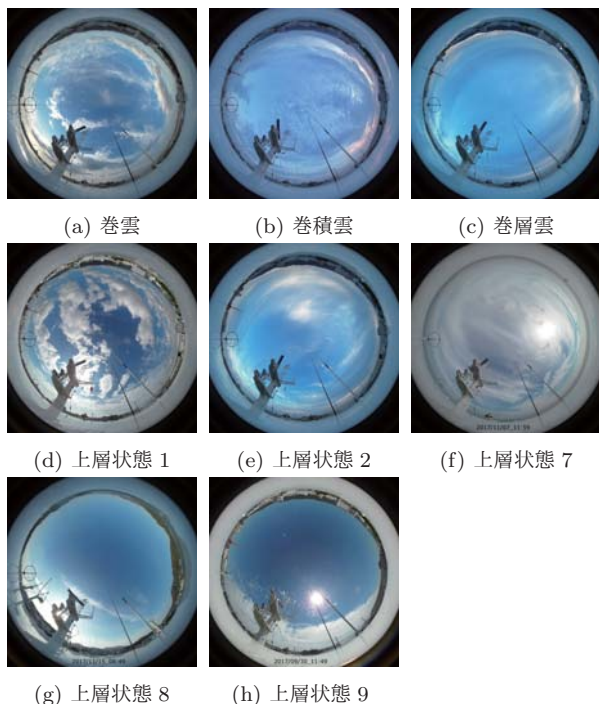


図 6: 上層雲形・状態

3. 学習環境と実験結果

Deep learning のライブラリは TensorFlow-gpu(1.4.1)[4] と Keras(2.0.6) を用いている。また、既存の学習モデルである ResNet50 を用いて finetuning を行っている。49 層目まで固定し、全結合層のみをデータに対応した形に変形し、学習を行っている。学習時は 5-fold の cross-validation により accuracy を計算している。学習率は 0.0002 から始めて、200 epoch 時に 0.0001 になるよう、指数関数にしたがって減少させながら学習を行っている。勾配法は確率的勾配降下法を用いて momentum を 0.9 に設定している。さらに、Keras の Data augmentation の機能を使って訓練データの変形を行っている。

- データ拡張は横・縦方向に幅に対して 0~0.1 の間で変形
- 反時計回りシアー変換を 0~0.2 π の間で変形
- 画像中心に対してズームを 0.6~1.4 の間で変形
- 画像を -180~180 の間でランダムに回転
- RGB の値を -0.2~0.2 の間でランダムに変動
- 水平・垂直方向にランダムで反転

以下の図 7 と図 8 は 1 epoch ごとのテストデータに対する精度である。

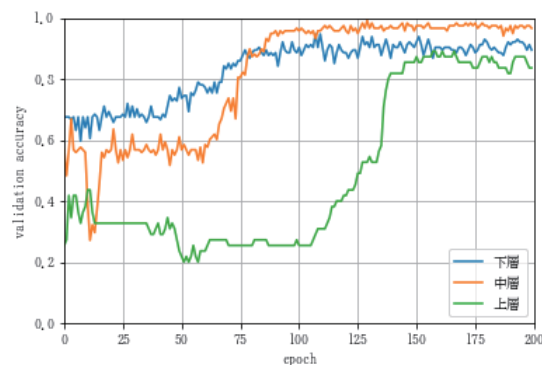


図 7: 雲形

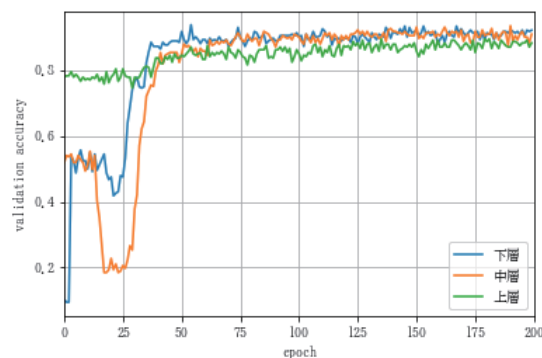


図 8: 状態

図 7 と図 8 では、雲形・雲の状態共に上層以外は 0.9 を超える高い accuracy を示している。まだすべてのラベルの全天球画像が集まっていない問題はありますが、先行研究において雲形は高い精度であることを考慮すると、全ラベルの場合は若干

分類しにくい程度で収まる可能性を示している。この事実は、今後、訓練データが完全な形で収集できれば、チューニング次第では、通常航海においても実用に足る学習器の生成が可能であることも示している。

実運用では未知のデータを扱うことから、図9～図11では過学習の度合いを示す指標である、validation loss が最も低い場合におけるテストデータに対する accuracy の confusion-matrix を示している。下層・中層の精度は上層よりも高い結果となった。表2や表3を見ると分かるが、各データの偏りが大きく、特に上層は枚数も少ない上に大きなデータの偏りが生じていることが大きな原因であると考えられる。上層の枚数が少ない理由としては、下層と中層の雲に遮られることで収集画像に写り込まないことである。このことから、今後も収集できる上層の枚数は大幅に増えない可能性があるため、収集量を増やす方法が必要となる。

		Predicted label			Predicted label			
		Cumulus	Stratocumulus	Stratus	0	1	4	7
True label	Cumulus	363	22	3	620	12	1	7
	Stratocumulus	26	94	0	14	301	16	1
	Stratus	6	0	61	10	20	81	5
					True label			
					0	1	4	7
					14	301	16	1
					10	20	81	5
					30	0	1	92

(a) 下層雲形

(b) 下層状態

図9: validation loss の最も低い学習器における下層の validation accuracy

		Predicted label		Predicted label				
		Altostratus	Altostratus	0	1	2	3	7
True label	Altostratus	241	16	584	6	4	9	1
	Altostratus	8	328	4	114	11	2	3
				True label				
				0	1	2	3	7
				4	114	11	2	3
				0	5	213	0	1
				34	0	0	75	4
				1	5	1	7	75

(a) 中層雲形

(b) 中層状態

図10: validation loss の最も低い学習器における中層の validation accuracy

		Predicted label			Predicted label									
		cirrostratus	cirrocumulus	cirrus	0	1	2	7	8	9				
True label	cirrostratus	58	4	11	904	15	0	7	1	6				
	cirrocumulus	10	73	8	11	63	5	0	4	3				
	cirrus	9	7	102	0	5	17	0	1	1				
					True label									
					0	1	2	7	8	9				
					11	63	5	0	4	3				
					0	5	17	0	1	1				
					7	9	0	1	18	2				
					8	7	3	3	1	7				
					9	19	9	3	0	3				
					9	19	9	3	0	3				

(a) 上層雲形

(b) 上層状態

図11: validation loss の最も低い学習器における上層の validation accuracy

4. 結論と今後の課題

本論文では、ラベルは不完全ではあったが、収集した雲の全天球画像データにラベル付けを行って構築したデータセットにて、雲形と状態の両方の識別において高い精度を出すことに成功した。これにより現状のデータ収集方法と学習方法に問題がないことを示すことができた。また、先行研究では取り組まれていなかった雲の状態においても識別が可能であることを示すことができた。しかし、学習精度の向上や一般船舶にて実運用

を行うために解決しなければならない課題が多数存在する。

まず第1に、データ収集のための撮影は深江丸と沖縄県伊是名島の2箇所で行っているが、網羅的なデータセット構築のために観測地点を増やす必要がある。また、データ収集量の増大に合わせて、効率的なラベル付け作業の実行方法も同時に検討が必要である。

第2に、全天球画像の特性上、1枚の画像だけでは雲の高さの判別が困難である。気象測器として雲底高度計があるが測定距離の限界と直上地点のみという制限がある。正確なラベル付けおよび学習精度向上のために、全天球画像の中でどの範囲にどの高さの雲が存在するかを判別する方法の検討が必要である。

第3に、乗組員が雲の観測をする際には前回の観測結果を参考に時系列に測定を行なっている。今回の学習では、1枚の画像だけで学習を行っているが、前後の画像も併せて学習することで、雲の流れによる雲の高度や雲形・状態の変化を検証する必要がある。

第4に、本研究では全天球画像をそのまま利用しているが、画像のセンター部分やエッジ部分などをマスクしたり、目視観測により近い矩形画像に変換して学習させた場合の全天球画像との精度比較を検証する必要がある。

5. 謝辞

田中謙次氏にコーディングの作業の実地や学習方法について議論を行って頂いた。兵庫県立大学大学院応用情報科学研究科申吉浩教授にGPUマシンの貸出や学習方法について指摘を頂いた。深江丸での観測の実地は、矢野吉治船長、青山克己一等航海士をはじめとする乗組員の皆様にご協力頂いた。なお、本研究は、国土交通省の「先進船舶・造船技術研究開発費補助事業（先進船舶技術研究開発）」の支援対象事業における「海上気象観測の自動観測・自動送信システムの開発^{*1}」の一環として行われたものである。

参考文献

- [1] Jinrui Gan, Weitao Lu, “Cloud Type Classification of Total-Sky Images Using Duplex Norm-Bounded Sparse Coding” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017.
- [2] A. Kazantzidis, P. Tzoumanikas, “Cloud detection and classification with the use of whole-skyground-based images” Atmospheric Research, vol.113 pp.80–88, 2012
- [3] A. Heinle, A. Macke, “Automatic cloud classification of whole sky images”, Atmos. Meas. Tech., vol.3, p557–567, 2010
- [4] Martn Abadi et al., “TensorFlow: A System for Large-Scale Machine Learning”, 2016
- [5] “第5章 雲の観測”. 船舶気象観測指針, 改定第7版(2013), 気象庁, <https://marine.kishou.go.jp/brochure/guide-jp.pdf>, (参照 2018-03-06)

*1 株式会社商船三井、古野電気株式会社、スカパー JSAT 株式会社を主体とする産学官連携の共同研究