

2次元翼モデルの可視化CFD結果に関する 深層学習と位相的データ解析

Deep learning and Topological Data Analysis
for Visualized CFD Results of 2-dimensional Wing Model

伊藤雅仁¹⁾, 奥脇弘次¹⁾, 望月祐志^{1,2)}, 小杉範仁³⁾
Masato Ito, Koji Okuwaki, Yuji Mochizuki, and Norihito Kosugi

- 1) 立教大学理学部 (〒171-8501 東京都豊島区西池袋3-34-1, E-mail (MI, KO & YM): sigure_3030@rikkyo.ac.jp, okuwaki@rikkyo.ac.jp & fullmoon@rikkyo.ac.jp)
2) 東京大学生産技術研究所 (〒153-8503 東京都目黒区駒場4-6-1)
3) (株)ヴァイナス (〒530-0003 大阪府大阪市北区2-1-31 京阪堂島ビル, E-mail (NK): kosugi@vinas.com)

In this study, we applied machine learning (ML) techniques to the visualized results (or velocity distribution maps) of computational fluid dynamics (CFD) on a two-dimensional NACA airfoil models. First, the transfer learning was employed to predict the transitions of vortex situations in an efficient learning time by avoiding an issue of overfitting. Second, a combination of the persistent homology (as a representative topological data analysis method) and the K-means clustering was used to make labels for the situations of velocity distributions without arbitrariness via human judgement, suggesting a promising potential of automatic labeling for unsupervised learning. In the presentation, these two works are summarized.

Key Words : CFD, Transfer Learning, Topological Data Analysis, Persistent Homology, Keras, TensorFlow-GPU.

1. 緒言

近年の機械学習手法の発達に伴い、大規模な計算データを効率的に解析する需要が高まっている。特に数値シミュレーションでは、得た大量のデータをいかにして効率よく解析・解釈するかが重要となる。数値流体力学 (Computational Fluid Dynamics: CFD) 分野でも、シミュレーション後の可視化したポストデータに対して機械学習的アプローチを施し解析する事例[1,2]がある。このようにCFDをはじめとした計算シミュレーションは、結果回収や解析の面で機械学習手法と非常に相性がよい。

本研究では、CFDのポストデータである流速分布図(以下、流速図)に対し、深層学習手法[3]や他のデータ解析手法を適用した事例について紹介する。前者は、2次元化した翼形モデルの流速図を入力に転移学習 (Transfer Learning: TL) と教師あり学習をベースにした剥離現象の遷移予測、後者は近年注目されている位相的データ解析

(Topological Data Analysis: TDA) を応用した流速図の自動分類についての内容である。

2. 全体の背景・目的

2-1. CFD×深層学習 これまでに、二次元化した対称型翼形の流速図(Fig. 1)に対し、CNNモデルによる教師あり学習で種々の剥離現象を高精度で判定する試みを実施し、学習モデルの推論によって剥離遷移の予測も可能とした[4,5]。その研究では、NACA0018型[6]をはじめとした計3種のモデルに対してCFD計算を施し、取得した流速図データセットで学習を行っていた。正解クラスは{0: 剥離前、1: 後縁剥離、2: 前縁剥離、3: 剥離後}の計4種類である。

しかし先行研究のデータセットには非対称性の高い翼形での流速図が存在せず、汎用的に翼形剥離現象を判定

しにくいという問題があった。そこで本研究から、非対称型翼形であるNACA6306型[6]に対し、同様のCFD計算を実施したところ、翼形の前後で同時に流体剥離が生じるケースが確認された{4: 前後同時剥離、Fig. 2}。よって新たなクラスを考慮したデータセット作成と再学習の必要が生じる。しかしながら、再学習には時間的なコストがかかることもあり、より効率的に類似したデータを学習できる学習手法が必要だと考える。

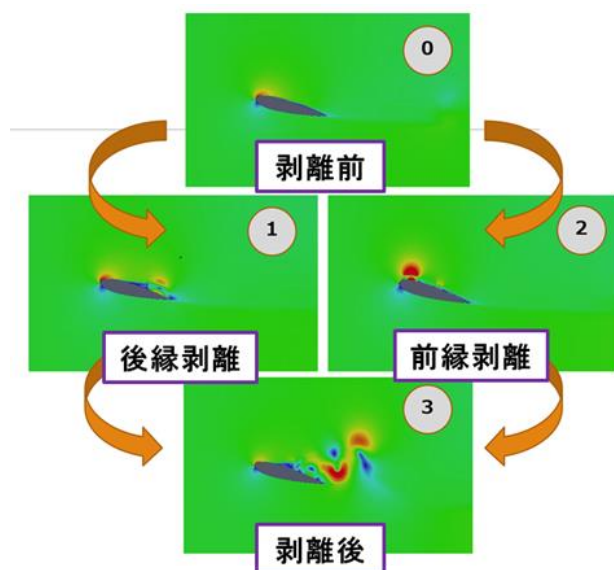


Fig. 1. NACA0018 型流速図とその剥離現象クラ

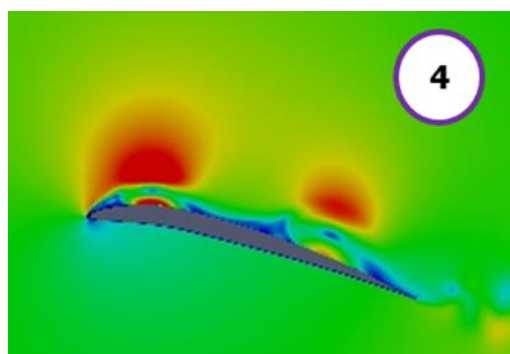


Fig. 2. NACA6306 型流速図。前後同時剥離。

2-2. 転移学習 翼形が対称型・非対称型であろうが、使用するデータは同じ形式の流速図である。このような類似したデータを効率的に学習する方法として、本研究では転移学習[7]を採用した。これはあるデータを事前に学習したモデルの特徴抽出器の一部を転用することで、

学習したデータと類似したデータを再学習する方法である。利点としては学習時間を短縮化でき、小規模なデータセットでも高精度で学習を行える。

本研究では、対称型NACA0018型流速図(4クラス)のデータセットをCNNモデルで学習し、その全結合層以外を転用し、新たに全結合層を取り付けたモデルで非対称型NACA6306型流速図(5クラス)のデータセットを転移学習させることを目的とする[8,9]。

2-3. CFD×位相的データ解析(TDA) 教師あり学習では、予め使用するデータを任意クラスに分類する必要がある。しかし分類するのは人間であり、そのデータセットには分類者の主観や偏見が介在し、恣意性のあるデータセットとなる恐れがある。無論そのようなデータセットで学習した場合、得られた結果にも恣意性が反映される。恣意性を排除したデータセット作成のためには、データから特徴を抽出し、それを考慮してデータを分類するフローが必要だと言える。

本手法ではデータからの特徴抽出に関して、近年注目されている位相的データ解析[10]を用いて、データ内の構造に由来する不変量の特徴として取得する。

2-4. TDAとパーシステントホモロジー 位相的データ解析とは、トポロジー理論をバックエンドに持つ構造情報に対する解析手法である。任意ホモロジー次元の構造的特徴量(位相不変量)である点(連結成分)、リング構造、空洞構造を本手法で解析可能である。TDAを使用した構造認識の一つにサイクロンの目を特定するという事例[9]があり、画像データに対しても位相不変量を取得・解析できることがTDAの利点である。

TDAは画像だけでなく、三次元点群データに対しても適用できる。パーシステントホモロジー(Persistent Homology)[11,12]と呼ばれる手法では、個々の点群を球体とみなし、その半径を増加させていくフィルトレーション計算によって位相不変量を特定できる。隙間が生じた場合の半径を生成半径(Birth time)、消滅した際の半径を消

減半径(Death time)という。しかしそれらが元データのどこに対応しているのかが不明では、特徴を見つけても解析できない。そこで逆解析(Inverse Analysis)することによって、その対応付けを行っている。

本研究では、逆解析を用いて流速図中の剥離が生じる箇所で1次元の位相不変量であるリング構造が生成するとして、剥離生成を画像データ上に特徴づける。さらに深層学習の特徴抽出力で剥離生成の特徴量を取得し、非階層クラスタリングによって、剥離生成に基づいた流速図の自動分類を実施する。

3. 転移学習を用いた剥離遷移予測

3-1. CFD計算とデータセット作成 NACA6306型翼

形のCFD計算では、非定常・非圧縮性のLES-Smagorinsky乱流モデルを想定した。流体は流速20 m/sの空気、翼形については、迎角{14, 15, 15.5, 16, 17}度の5種類、レイノルズ数が 2.0×10^6 となるようなモデル長に設定した。ソルバーアルゴリズムにはPIMPLE法を採用し、0.01秒刻みで2秒間の流速場と圧力の近似計算を実施した。計算後は、その結果から1迎角につき200枚の流速分布図を取得した(合計1,000枚)。CFD計算ソルバーにはOpenFOAM/HELIX-OS 2.0.4[14,15]を、流速分布図の取得にはParaView 4.0.1[16]を使用した。

取得した流速分布図1,000枚をTrainデータ640枚、Validationデータ160枚、Testデータ200枚に分割した。この分割はScikit-Learnライブラリのメソッドを使用してランダムに行っている。またNACA0018型流速分布図については、先行研究時に同条件で取得されたデータセットがあるため、そちらを使用した。

3-2. 事前学習と転移学習 Fig. 3に本研究で使用した学習モデルを示す。

まず畳み込み層を4つ持つ直列CNNモデルでNACA0018型流速分布図(56×56のRGB画像)を学習させた。その学習済みモデルの畳み込み層を別途保存し、NACA6306型流速分布図を転移学習する際に、その

モデルの重みを一部固定し、新しい全結合層を取り付けて再学習を行った。事前学習並びに転移学習では、損失関数に多クラス交差エントロピー誤差、勾配計算アルゴリズムにAdam[17](学習率 5.0×10^{-5})を採用し、学習回数100回、バッチサイズ20でTrainデータを学習した。

本学習では、Keras 2.2.4(TensorFlow-GPU 1.14.0 バックエンド)のフレームワークを採用し、GPUとしてNvidia GeForce 2080 Tiを2枚使用した。

また転移学習モデルと通常の教師あり分類モデルとで性能を比較するため、転移学習モデルと同じアーキテクチャーで、重みを固定しないモデルも用意し、NACA6306流速分布図を同様に学習させた。

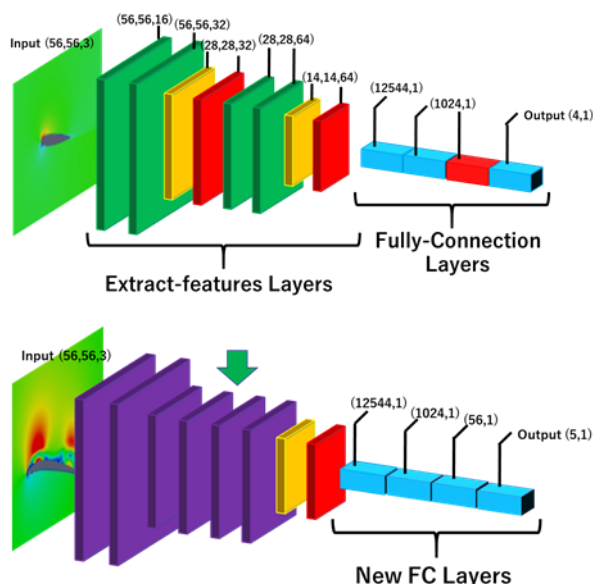


Fig. 3. 学習モデル。上段: NACA0018 型事前学習用 CNN モデル、下段: NACA6306 型転移学習用 CNN モデル。緑: Conv2D、黄: MaxPooling2D、赤: Dropout、青: Dense、紫: 重み固定層。

3-3. Testデータによる剥離遷移の予測 前述の通り、非対称翼形では前後同時剥離が生じたため、CFDシミュレーションから想定される剥離遷移は対称型の場合と異なる。Fig. 4にシミュレーションから目視で導いたNACA6306型の剥離遷移を示す。この遷移図をもとに、特定の2クラス間について、学習モデルの推論結果から剥離遷移の予測を試みた。実装時はルールベースで以下の条

件を満たす場合に剥離遷移を検出するよう実装した。

- Softmax関数の出力行列に0.3以上の要素が2つある
- それら要素に大小関係が成立する
- 値の大きい方から小さい方に遷移する

実際の流れとして、まずTestデータで転移学習後モデルによる推論を実施した。その出力に対し、上記のルールで剥離遷移を検出した。また各Testデータにも正解となるクラスがあるため、推論したクラスと正解クラスとの一致率も算出した。

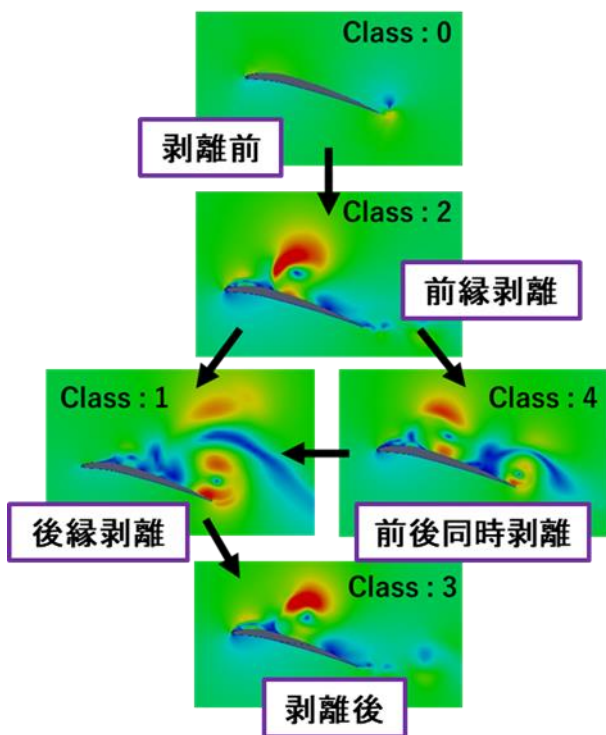


Fig. 4. CFD シミュレーションに基づく NACA6306 型の剥離遷移。

3-4. 結果 - 性能評価 - Fig. 5に学習回数に対する正答精度(Accuracy)、学習損失(Loss)の学習曲線を示す。左列が転移学習をしないモデルでの、右列が転移学習をしたモデルでの指標曲線を表す。またTable. 1には学習の最終ステップにおける各指標値、過学習と学習にかかった時間をFig. 5と同じ形式でまとめている。特に過学習は、20回目から100回目までのTrain・Validationデータでの指

標値同士の差を算出し、その平均と不偏標準偏差で定量的に評価する。

Accuracyについては、両モデル共に学習回数に伴って緩やかに上昇し、値は9割で収束した。Validationデータ数はTrainデータよりも少ないため、Validationデータでの指標値は小さく見積もられる。しかしそれを考慮しても、転移学習での正答性能は通常の教師あり分類モデルと同様の精度を持つことが確認できた。Lossについても同様で、転移学習の有無に関わらず、学習損失の緩やかな減衰が確認できた。

ここで過学習と学習実行時間を確認する。曲線から判断すると、Accuracy・Loss共に曲線間の離れ具合は転移学習をした場合に小さくなることが判明した。数値的にも転移学習をしなかった場合と比較してAccuracyで約2%、Lossで約7%の抑制が見られた。また学習時間は、転移学習をしないモデルでの学習が641.87秒であり、転移学習することによって約18%高速化した526.05秒をマークした。

今回は翼形の形状は異なるものの、同じRGB形式の流速図であったこともあり、非常に転移学習しやすいドメインデータであったことが転移学習の性能を高める要因だと考える。

3-5. 結果 - 剥離遷移予測 - Fig. 6に転移学習モデルで推論を行った場合のNACA6306型流速分布図の剥離遷移フレームを示す。本項では、迎角16度の第175フレームでの剥離遷移結果について記載する。

結果的にこのフレームでは、クラス4から1への遷移、つまり前後同時剥離から後縁剥離への遷移が検出された。

Softmax関数出力値を参照すると、クラス1の確率が35.94%、クラス4の確率が64.96%であった。本実装におけるルールに従うと、共に確率が30%以上で大小関係が成り立っているため、剥離の検出は妥当だと判明した。さらにその直前と直後のフレームを確認したところ、遷移前の正解クラスは4(前後同時剥離)、遷移後の正解クラスは1(後縁剥離)であり、遷移予測が妥当であると判断できた。

またTestデータの正答性能は転移学習をしない場合で

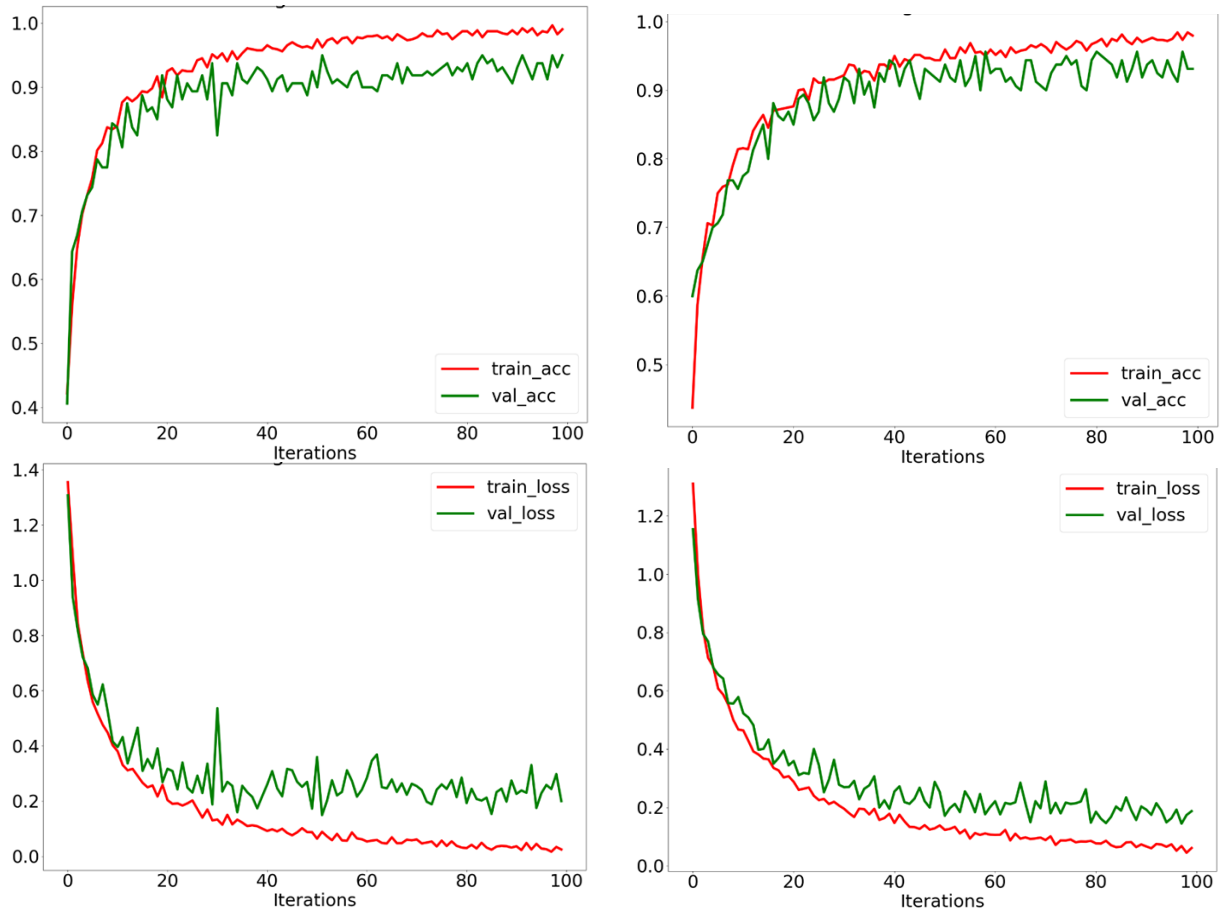


Fig. 5. 学習回数に対する Accuracy・Loss 曲線 (赤: Train データ、緑: Validation データ)。左列: 転移学習をしない CNN モデル、右列: 転移学習モデル。

Table. 1 学習終了時の各指標値と過学習・学習時間。左列: 転移学習をしない CNN モデル、右列: 転移学習モデル。

Model	CNN(Non-TL)	CNN(TL)
Train Accuracy	0.991	0.980
Validation Accuracy	0.950	0.931
Accuracy Overfitting (%)	5.34 ± 1.90	3.15 ± 1.92
Train Loss	0.025	0.061
Validation Loss	0.200	0.187
Loss Overfitting (%)	17.37 ± 6.60	10.14 ± 3.72
Run-time (sec)	641.87	526.05

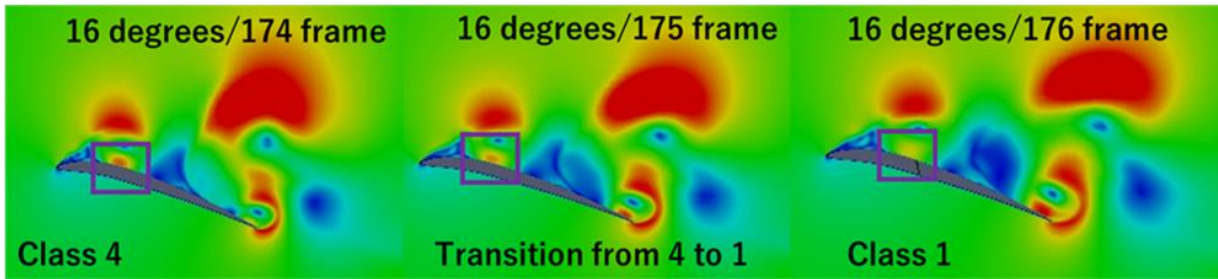


Fig. 6. Test データ (NACA6306/迎角 16 度) の剥離遷移予測結果。左: 遷移前のフレーム、中: 遷移予測フレーム、右: 遷移後のフレーム。枠で囲まれた部分は剥離遷移のポイントと考えられる。

90.00%、転移学習をした場合で89.00%であった。数値的には通常の教師あり学習における判定と同等の精度であり、転移学習での判定率の減少は遷移検出数の多さに依存すると推測する。

4. 位相的データ解析による流速図自動分類

4-1. パーシステントホモロジーによる流速図の解析・

逆解析 NACA6306型流速図(630×390、RGB、計1,000枚)をグレースケールに変換し、グレースケール化した流速図を入力として、パーシステントホモロジーによるフィルトレーション計算を行った。ここで取得したバイナリファイルから、入力データ内の生成半径と消滅半径を記録したテキストファイル(計1,000データ)を出力した。次にグレースケール流速図と対応するテキストファイルを読み込み、消滅半径と生成半径との差(リング構造の寿命)が20より大きくなる場合、グレースケール流速図に赤点で生成半径をマッピングすることで、逆解析を行った(Fig. 7)。

パーシステントホモロジーによる計算では、Homcloud[18]というフリーソフトウェアを使用し、フィルトレーション計算部分をDIPHA[19]バックエンドで演算している。

4-2. 特徴抽出と自動分類 Fig. 7のような逆解析後の流速図(224×224、RGB)を入力として、VGG16[20]モデル(ImageNet[21]学習済み、全結合層なし)による特徴抽出

[22]を行った。本モデルの最終出力層は4次元テンソルを与えるためにこれを平坦化した。これら一連の操作を計1,000データに対して実行し、得た計1,000個のベクトルを連結することで行列に成形した。

この行列に対し、非階層クラスタリング手法であるK-means[23]を用い、初期クラスター数5でクラスタリングを実施して予測したクラスター番号{0,1,2,3,4}と入力データ名を格納したCSVファイルを出力した。最後にCSVファイルを参照して、入力データ名と同じオリジナルの流速図を各クラスター番号のディレクトリに配置することで、1,000データの流速図を自動分類した。

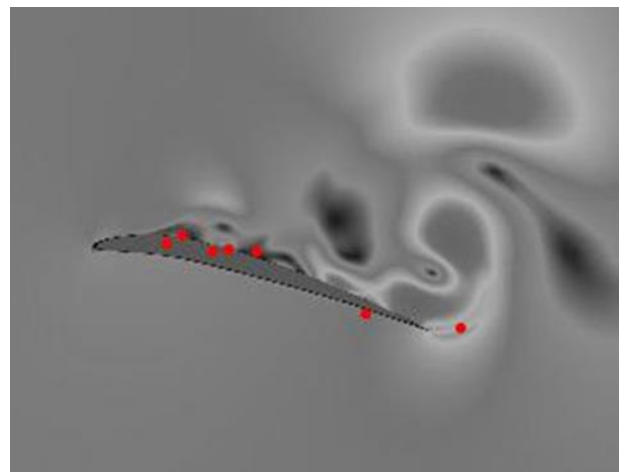


Fig. 7. パーシステントホモロジーによる逆解析後の流速図。赤点: 消滅半径と生成半径の差が20を超える場合のリング構造生成点。

4-3. 結果 - 逆解析ありの自動分類 - Fig. 8にパーシステントホモロジーで逆解析を実施した流速図を用いた自動分類の結果を示す。本研究ではK-means法の初期クラスター数を5に設定したため、計5種のクラスに流速図

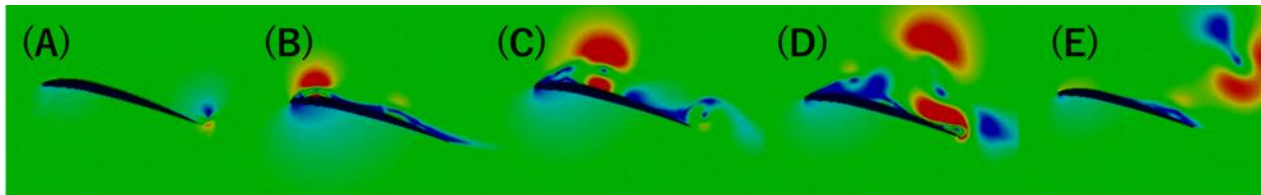


Fig. 8. パーシステントホモロジー逆解析を使用した流速図の自動分類結果。A: 剥離前、B: 前縁剥離、C: 前後同時剥離、D: 後縁剥離、E: 剥離後のフレーム。

を分類している。これは目視によるクラス分けと同数のクラスになる。

CFDシミュレーションの時系列を考慮すると、剥離前から前縁剥離を得て、前後同時剥離になり、後縁剥離に遷移して最終的に剥離後になるという目視でクラス分けを行った場合と近い分類がされた。本研究では非対称型であるNACA6306流速図を用いており、追加で確認された前後同時剥離はシミュレーション結果をもとに、前縁剥離から後縁剥離への中間に位置する剥離現象であることが確認できている。したがってパーシステントホモロジーを用いたデータ自動分類によって、膨大なデータに対して目視で教師を付ける手間を省略できると考える。

5. 結言

転移学習を用いた翼形剥離遷移予測で用いた学習モデルは、通常の転移学習をしない教師あり分類と比較して、過学習を抑制し、時間的コストを約18%抑制できるものであった。また転移学習によって対称型翼形で調整した重みを転用し、事前学習時に学習していない剥離現象を判定、その遷移を検出することができた。判定性も高く、コストも削減できることから転移学習は深層学習の効率化に大きく寄与する手法であると言える。

教師あり学習の恣意性は、位相的データ解析手法を適用し、深層学習の特徴抽出とクラスタリング手法を併用することで抑制でき、人間の手を介さない特徴量に依存した自動的な分類が可能となった。剥離生成点を位相不変量であるリング構造の生成点とみなして逆解析で可視化、それを抽出することで、純粋に特徴量を獲得できる。

6. 今後の展望

6.1 転移学習 機械学習を用いた場合、説明解釈性の難しさは課題として直面する。特に画像認識用の深層学習モデルでは、画像のどこを特徴として捉えているかが判断しがたい。そこでGrad-CAM[24]を用いた特徴量の可視化を行い、本モデルでの推論の妥当性をより定性的に評価していく。

また深層学習には様々なハイパーパラメーターがあり、より精度の高い予測や推論を行うための適切なパラメーター設定は必須である。Keras専用のパラメーター探索ライブラリHyperas[25]では、バッチサイズや学習率だけでなく、学習モデル内のノード数等も調整できるため、それを用いたパラメーター探索フロー実装も検討している。

6.2. 位相的データ解析 本研究では分類するクラス数を決める手法にK-means法を採用したが、これは初期クラスター数に依存するクラスタリング手法である。当研究室では、そのクラスター数を自動決定して自動分類を行えるよう、X-means法やG-means法といった非階層クラスタリング手法[26,27]の適用を進めている。これによって分類するクラス数決定に関しても、人の手が介さない自動分類を行えると考えている。

6.3. 共通する展望 本研究では、翼形形状を変えて取得したデータで転移学習を実施したが、翼形以外の形状にCFDシミュレーションを実施して取得した流速図に対し、転移学習のドメイン適応ができるかを評価する必要があると言える。今回同様、二次元化した乗用車[28]流速図も取得できているため、翼形から乗用車への転移学習を試作している。しかし乗用車の剥離現象は目視で分

類することが難しいため、本自動分類手法で分類したデータを使用した転移学習を試作している。一連の手法をひとつながりのフローとすることで、大規模なポストデータ解析に大きく貢献できると期待している。

7. 謝辞

本研究は、文科省ポスト「京」プロジェクト重点課題6の支援を得て進められている。併せて、本研究に対してご教示と励ましを頂いている東京大学生産技術研究所の加藤千幸先生、東京大学名誉教授の小林敏夫先生、東京大学工学部の吉村忍先生、神戸大学大学院システム情報学研究科の坪倉誠先生、九州大学情報基盤センターの小野謙二先生、非対称翼形への試みをご教示頂いたJAXA宇宙科学研究所の高木亮治先生、位相的データ解析手法についてご教示頂いた理化学研究所の大林一平先生に深謝する。また、OpenFOAM等のCFD環境を提供いただいた(株)ヴァイナスの藤川泰彦社長にも感謝したい。

参考文献

[1] Taira, K. et al., “Modal Analysis of Fluid Flows: An Overview,” *AIAA JOURNAL*, Vol.55, No.12 (Dec 2017), pp.4013-4041.
[2] 平邦彦, “固有直交分解による流体解析: 1. 基礎,” *ながれ*, Vol.30 (2011), pp.115-123.
[3] LeCun, Y. et al., “Deep Learning,” *Nature*, Vol.521 (28 May 2015), pp.436-444.
[4] 伊藤, 齊藤, 土居, 他, “二次元翼周り流れのCFD解析結果のTensorFlowによる自動解釈の試み”, G-03-02, 計算工学講演会, 名古屋, (2018/6/7).
[5] 小杉, 近藤, 秋永, 他, “AIを活用した流体解析シミュレーション技術の開発”, *機械設計*, 2018年7月号, 62 (2018) pp. 42-46.
[6] <http://www.cenav.org/wing_conv/>.
[7] Chuanqi, T. et al., “A Survey on Deep Transfer Learning,” *arXiv* (2018), <arXiv:1808.01974v1>.
[8] 伊藤ら, “転移学習と生成ネットワークの試行事例,” *J. Comp. Chem. Jpn.* Vol.18, No.3 (2019), pp. 132-135.
[9] 伊藤, 遠藤, 望月, 泰岡, “深層学習とシミュレーションの連携事例,” *計算工学*, Vol.24 No.4 (2019), pp.3983-3987.
[10] Wasserman, L. “Topological Data Analysis,” *Annual Review of Statistics and Its Application*, Vol. 5, No. 1 (2018),

pp. 501-532.

[11] Inatsu, M. et al., “A Cyclone Identification Algorithm with Persistent Homology and Merge-Tree,” *SOLA*, Vol.13 (2017), pp. 214-218.
[12] Zomorodian, A. et al., “Computing Persistent Homology,” *Discrete & Computational Geometry*, Vol.33, (Feb 2005), pp. 249-274.
[13] Obayashi, I. et al., “Persistence diagrams with linear machine learning models,” *J Appl. and Comput. Topology*, Vol.1 (Jun 2018), pp. 421-449.
[14] <<https://openfoam.org/>>.
[15] <<https://engys.com/>>.
[16] <<https://www.paraview.org/>>.
[17] Diederik P. K. et al, “Adam: A Method for Stochastic Optimization,” *arXiv* (2014), <arXiv:1412.6980v9>.
[18] <https://www.wpiaimr.tohoku.ac.jp/hiraoka_lab/homcloud/>.
[19] <<https://github.com/DIPHA/dipha>>.
[20] Simonyan, K., Zisserman, A., “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv*, Apr 2015, <arXiv:1409.1556v6>, pp.1-14.
[21] Deng, J. et al., “ImageNet: A Large-Scale Hierarchical Image Database,” *IEEE Computer Vision and Pattern Recognition (CVPR)*, (2009), pp.1-8.
[22] <<https://keras.io/ja/applications/>>.
[23] MacQueen, J., “Some methods for classification and analysis of multivariate observations,” *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, Vol.1 (Univ. of Calif. Press, 1967), pp.281-297.
[24] Ramprasaath R.S. et al., “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,” *Int J Comput Vis* (2019), pp 1-24.
[25] <<https://github.com/maxpumperla/hyperas>>.
[26] Pelleg, D. et al., “X-means: Extending K-means with Efficient Estimation of the Number of Clusters,” *ICML-2000*.
[27] Greg H. et al, “Learning the k in k-means,” *NIPS'03* (2003), pp. 281-288.
[28] <<https://www.myminifactory.com/object/3d-print-honda-kiwami-893>>.