

実機振動診断カルテに基づいたマルチモーダル・ ディープラーニングによる中・小型回転機械の診断効率化

An approach to building an efficient diagnosis system for middle range rotational machinery using multimodal deep learning and diagnosis database.

株式会社プライア 長野 達朗 Tatsuro NAGANO Member
株式会社沢田テクニカルサービス 沢田 作雄 Sakuo SAWADA Member

Abstract

For the purpose of detecting anomaly of rotational machinery, the vibration diagnostic method is often used. However, it requires experts to understand insights of cause or sign of abnormal progression. We have developed SA-FRONTIER-PRO, a diagnosis system that assists operators to acquire vibration signals, analyze measured data, and examine machinery through an inquiry approach. To step forward towards further efficiency, we employ multimodal deep learning techniques to mimic decision-making process of experts especially focusing on how to perceive and extract features from measurement data.

Keywords: Vibration Analysis, Anomaly Detection, Multimodal Deep Learning, Machine Learning

1. はじめに

回転機器の状態を把握する上で振動診断法はよく用いられる手法の一つである。しかし、異常原因や進行程度の推定は難易度が高く、熟練者や専門家の知識を要する。また設備診断の実施には、従来型の振動計では機能的に限度があり、取り扱いにも技術が求められる。総じて振動診断は熟練者の経験と勘によるところが依然多い。一方、熟練者の不足と後継者の育成は常に課題であり、それに対処すべく、業界各社は様々な自動診断装置を製品化している。しかし現状は、各々に得意不得意があり、異常要因の特定には限りがある。

この状況を踏まえ我々は、熟練者の思考・判断プロセスを分析し、計測から解析、診断機能までを実装したSA-FRONTIER-PRO（以下、SA）を開発した[1]。SAには、熟練者による過去の診断結果の実例をデータベースとして登録してあるため、ユーザは問診形式にて分析を進めることが出来る。この問診法により、SAは診断の効率化を可能にすると同時に、ユーザの診断スキルも向上出来るよう工夫されている。

連絡先: 長野達朗 〒181-0013 東京都三鷹市下連雀
3-38-4 三鷹産業プラザ B1F-B 株式会社プライア
E-mail: info@prior-japan.com
沢田作雄 〒721-0941 広島県福山市引野町北 4-1-1
株式会社沢田テクニカルサービス
E-mail: sawada@sawada-ts.co.jp

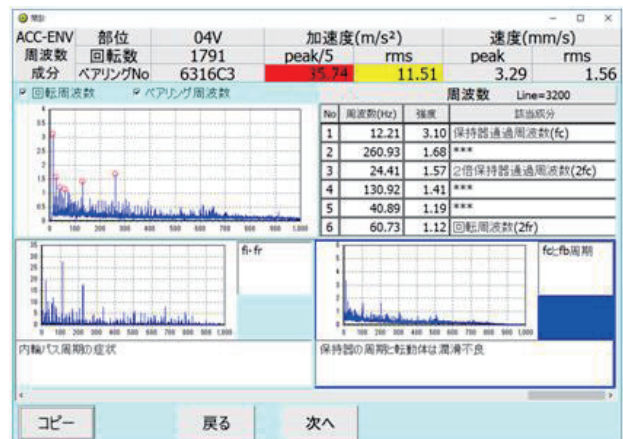


Fig.1 Appearance of SA-FRONTIER-PRO

2. SA-FRONTIER-PRO による診断

熟練者は、設備を診断するにあたり、全体と部分的な視診を行い外観的な異常や特徴を捉え、聴音棒等を用いて異常の察知や潤滑状態の良否など聴診を行う。更に、軸受温度や指先に伝わる揺れを触診で感じ取り、問診にて異常発生の時期、保全履歴、負荷状況などを把握する。つまり五感を働かせて総合的に判断をしている。

このような熟練者の技術を全て診断装置に置き換えることは難しいが、SAは簡易的な計測で効率良く主要な診断が出来るよう設計されている。例えば、設備の状態は、振動信号として、複数の加速度センサ間で同期を取りな

がリアルタイムにPCに記録される。複数の解析プロセスを経て、回転軸受部の異常やその兆候を可視化し、更に傾向的な振動推移からユーザーに適切な対策を促す。特にSAは、問診形式による診断を特徴としており、回転機器の属性や軸受等主要な構成部品の情報を入力することで、特徴量を自動抽出し判定する機能を備え持つ。しかし一方で、振動の時系列波形や周波数スペクトルのパターン認識などは、定量的に特定・自動化することが難しく、人手を介してデータベースと比較をして選択する必要がある。

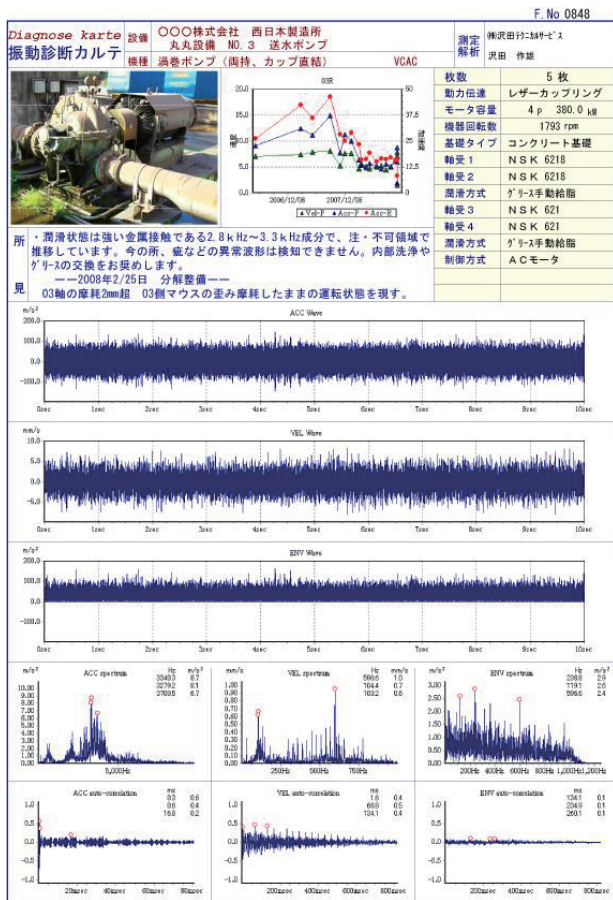


Fig.2 Example of SA-FRONTIER-PRO diagnosis report

3. パターン認識

パターン認識において、ディープラーニング技術は近年著しい成果を挙げている。2012年、大規模画像認識コンテストのImageNet Large Scale Visual Recognition Challenge (ILSVRC)にて、トロント大学のSuperVisionチームが認識率で他チームに大差を付けて優勝した。その際に使用したアルゴリズムは、畳み込みニューラルネットワーク(以下、CNN)と呼ばれる人間の脳の視覚野に

ヒントを得たディープラーニングの技術である[2]。以降、CNNを始めとしたディープラーニングは、機械翻訳[3]から画像の自動キャプチャ[4]、自律走行車[5]に至るまで、機械認識における主要な技術の一つとして様々な分野で応用されている。

アプリケーションの増加と共に、アルゴリズムも日々改良が繰り返されており、振動分野においても例外ではない。例えば、回転設備の振動をスペクトログラムとして表現しCNNを適用することで、正常と異常のパターン分類が可能であることが報告されている[6]。設備を常時監視し、異常の兆候が表れた時点で適切な処置を行うことが出来るようになれば、TBM(時間基準保全)からCBM(状態基準保全)への保全方法の移行もしくは融合も現実的なものになると考えられる[7]。

4. マルチモーダル・ディープラーニング

人は、1つの事象を感覚器毎に特定種類の刺激として受け取り、末梢神経系を経て中枢神経系に情報を伝達する。脳はその情報を集約、分析、統合して事象を認識する。振動診断の熟練者もまた、同様のプロセスを経て総合的に判断している。これを踏まえ、ここではSAの診断過程で手作業を要している波形データのパターン認識において、マルチモーダルなニューラルネットワークの適用を考える。

4.1 特徴抽出

熟練者は、収録した加速度データから微細な疵や衝撃により発生する振動を検知し、また速度データから摩耗や劣化情報を読み取る。これに倣い、加速度波形(以下、ACC)及び、その信号処理により得られる速度波形(以下、VEL)、そして加速度の包絡線波形(以下、ENV)もパターン認識の対象に含めて考える。

ACC、VEL、ENVの各波形から、それぞれが持つ詳細な特徴を捉えるべく、基となる時系列波形(以下、WAVE)と併せて、高速フーリエ変換によるスペクトル波形(以下、FFT)及び波形振幅の確率分布(以下、AMP)を、前処理を施して扱う。FFTは周波数情報、AMPは時系列波形の長期的な振る舞いを捉えることを目的とする。

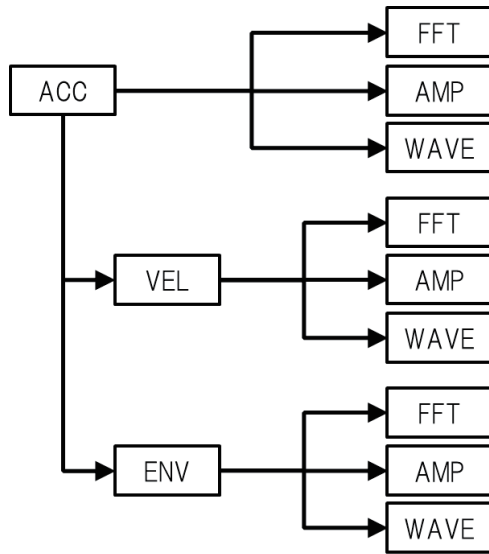


Fig.3 Features and corresponding time-domain signals

4.2 スライド窓によるデータ拡張

各特微量は、収録した時系列波形に対して、一定サイズの窓をスライドさせながらデータを切り出し、前処理を施す。窓サイズの設定は、FFTにおいて周波数分解能を決め、また AMP に対しては時系列波形の情報量に関係する[8]。窓のスライド量は、1つの時系列波形データから得られるデータ量に繋がり、後に述べるニューラルネットワークの学習データ数に関わる。重なりを多く取ればデータ数を増やすことが出来るが、一方で個々のデータを平均化してしまうため、ニューラルネットワークの学習状況を確認しながら設定する必要がある。

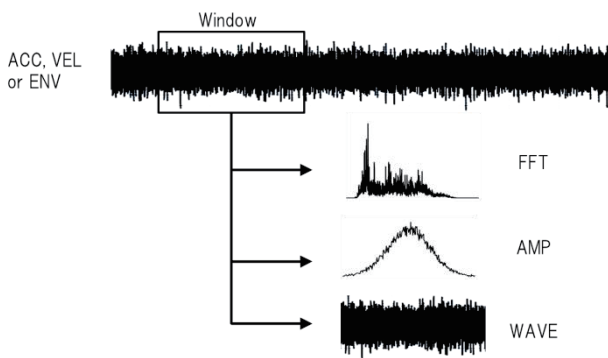


Fig.4 Sliding-window technique for the feature extraction

4.3 ニューラルネットワーク構造

一般に、実設備の診断データは、その運用面から、正常状態のものが圧倒的に多く、異常データが少ないことが普通である。そのため、実機データを使用してディープラーニングを行う場合、教師なし学習の手法を用いて

異常検知や外れ値検知を行うことが多い[9, 10]。幸い我々は、過去の実機診断カルテをデータベースとして所有しており、異常を含め十分な量のデータが手元にある。診断結果からデータを「現象」毎に振り分け、教師ラベルを付加することで詳細分類が可能な教師あり学習を実施することが出来る。

ディープニューラルネットワークを用いて、FFT、AMP、WAVEの形状を認識するにあたり、最も単純な方法は、それぞれに独立したネットワークを定義し学習を行うことである。この場合、入力ベクトルの次元やデータセット数が異なっても学習を実施する上では問題にならない。しかし推論時、それぞれのネットワークから出力される分類クラスは、必ずしも一致するとは限らず、総合結果は、各出力の多数決や平均値等から判断する必要がある。クラス毎のデータに明確な特徴が見られる場合はこれで問題無いが、実際には特徴抽出方法により捉えやすさは異なる。これは学習モデルの精度が上がらない要因となり得るため別のアプローチを検討する。

並列ネットワーク[11]や、ランダムマルチモデル[12]他幾つか方法は考えられるが、ここでは複数入力を統合して同時に処理可能なマルチモーダル・ディープニューラルネットワークを採用する[13]。マルチモーダルの手法は、ロバスト性が高く、1つのデータで特徴を捉えられなくとも、他のデータにて補完出来ることが強みである[14]。例えば、時系列波形に長期的で緩やかな変動がある場合、FFTやWAVEでその特徴を掴むことは難しいが、AMPがそれを捉えることで、全体として学習モデルにその特徴を反映させることが出来る。

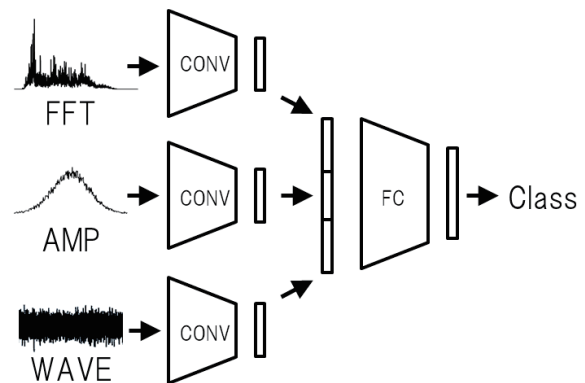


Fig.5 Multimodal deep neural network structure

具体的には、図5に示すよう、ニューラルネットワーク前段で各入力に適した構造を用意し、個別に特徴抽出を行う。各特徴ベクトルを結合した上で、後段の分類用

ネットワークにてクラスを予測する構造とした。前段は1次元の畳み込みニューラルネットワーク (CONV 部) を定義し、後段では全結合層 (FC 部) を重ねた構造とした。各ネットワークの詳細を表1、2、3及び4に示す。

Table 1 FFT feature extraction layers of the network

Layer	Input Size	Filter Number	Filter Size	Stride
CONV1	2,048	32	32	2
POOL1	1,024	32	4	4
CONV2	256	64	16	2
POOL2	128	64	4	4
CONV3	32	128	8	2
POOL3	16	128	2	2

Table 2 AMP feature extraction layers of the network

Layer	Input Size	Filter Number	Filter Size	Stride
CONV1	256	32	32	2
POOL1	128	32	2	2
CONV2	64	64	16	2
POOL2	32	64	2	2
CONV3	16	128	8	1
POOL3	16	128	2	2

Table 3 WAVE feature extraction layers of the network

Layer	Input Size	Filter Number	Filter Size	Stride
CONV1	25,600	16	64	2
POOL1	12,800	16	5	5
CONV2	2,560	32	32	2
POOL2	1,280	32	4	4
CONV3	320	64	16	2
POOL3	160	64	4	4
CONV4	40	128	8	2
POOL4	20	128	2	2

Table 4 Classification layers of the network

Layer	Input Size	Output Size
FC1	3,328	1600
FC2	1,600	800
Output	800	16

4.4 ニューラルネットワークの学習

全層の活性化関数はReLUとし、プーリング層にはMax Poolingを用いた。また、各層にBatch Normalization、全結合層にはL2正則化とDropoutを使用した。各層の重みはHe初期化を行い、AdamOptimizerを最適化関数とし、学習は誤差逆伝搬法にて行った。

学習データは、規模が同程度の汎用片持ち送風機の診断データを4機種、計56セット用いた。各セットには、ACC、VEL、ENVの波形データが、それぞれ約10秒間以上収録されており、スライド窓により、FFT、AMP、WAVEデータを各々4,000セット以上抽出した。クラス分類は、ACC、VEL、ENVそれぞれ個別の現象に紐づけられるよう、各々クラスを定義した。

学習は、ACC、VEL、ENV共にデータセットの70%を学習用、10%を評価用、残り20%をテスト用とした。ENVの学習モデルをテストデータで評価した際の混同行列を図6、学習モデルの精度を表5に示す。ENVの結果を示したのは、分類クラス数が最も多いためである。ACC及びVELも傾向は同様である。クラスは、良好域 (2)、f1アンバランス (3)、f12倍 (4)、その他 (7)、f0周期 (8)、柔構造の偏芯の影響 (10)、軸受部摩耗の兆候、揺れ (11)、軸受部摩耗の症状 (12) である。混同行列より、「軸受部摩耗の兆候、揺れ」のラベルが付加されたデータを「良好域」と誤認識したケースがいくつか見られる。学習モデルの精度は数値こそ高いものの、次節に述べる理由で汎化性に課題があるため、参考値として記す。

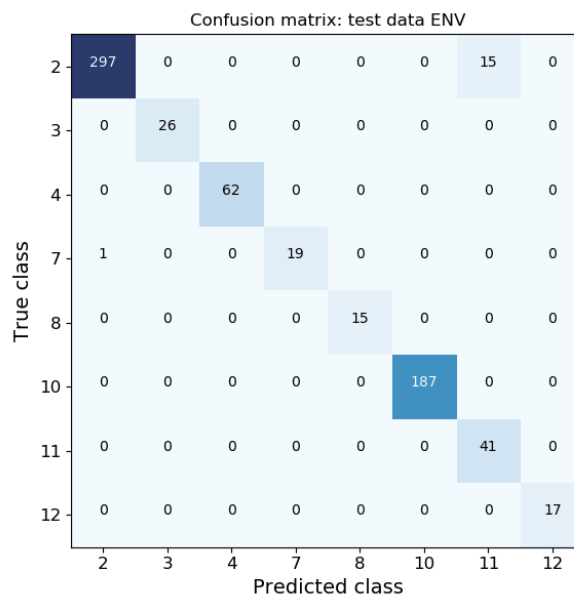


Fig.6 Confusion Matrix of ENV evaluated with test data

Table 5 Accuracy of ENV evaluated with test data

Class	Precision	Recall	F-measure	Accuracy	Entry
2	0.997	0.952	0.974	0.976	312
3	1.000	1.000	1.000	1.000	26
4	1.000	1.000	1.000	1.000	62
7	1.000	0.950	0.974	0.999	20
8	1.000	1.000	1.000	1.000	15
10	1.000	1.000	1.000	1.000	187
11	0.732	1.000	0.845	0.978	41
12	1.000	1.000	1.000	1.000	17

4.5 学習モデルの精度

今回使用したデータベースは、教師ラベル付加が完了していないため、学習に使用出来るデータ量に限りがある。また基となる振動測定データは、機器の定常稼働状態を記録したものであり、収録中の変化はほぼ見られない。そのため、スライド窓で拡張したデータセットは、数こそ増すが、テストデータを含め学習データ全体が特徴空間内でクラス毎に局在化していると考えられる。また、窓サイズ (25600 点) とスライド幅 (4096 点) の設定により重複部分が比較的大きいため、データが平均化され、これも局在化に影響していると考えられる。一般に、局在化したデータ分布において、境界の決定は比較的容易である。モデル評価に使用したテストデータも、局所化された領域から大きく外れないため、高い精度が得られたと考えられる。学習モデルの精度は、直接汎化性とは関係性がなく教師データ拡充の課題が残る。

4.6 マルチモーダル手法のロバスト性

今回は、FFT、AMP、WAVE の入力データを1つのマルチモーダルなネットワーク内で統合し同時に扱った。ここではそのロバスト性を考える。ロバスト性の検証は、本来、様々なデータを用意して実際にモデルを適用した際の振る舞いを確認することが理想であるが、現時点では教師データ量に限りがあるため、既存のデータを用いて以下のように行う。

まず、FFT、AMP、WAVE のデータセットに対して各々独立したユニモーダルなネットワークを別途用意し学習モデルを作成する。次に、FFT、AMP、WAVE から何れか1つを選択し、マルチモーダル及びユニモーダルの両方に同じテストデータを入力する。テストデータは、一定の割合でランダムに選択した指標の値を0に置き換え

る。置換率を0%から100%まで変化させ、その際の両モデルの精度 (多クラス分類であるため、Macro F1 Score を使用) の推移を比較する。

ACC データにおいて、FFT の置換を行った場合の精度推移を図7に示す。ユニモーダルの学習モデルは、置換率が20%を超えたあたりから精度低下が見られるが、マルチモーダルでは置換の影響を殆ど受けず精度を維持している。置換率100% (Ratio=1.0) は、FFT の入力全てが0であることを意味するが、他モード (AMP 及び WAVE) の補完により、精度低下を回避していると考えられる。

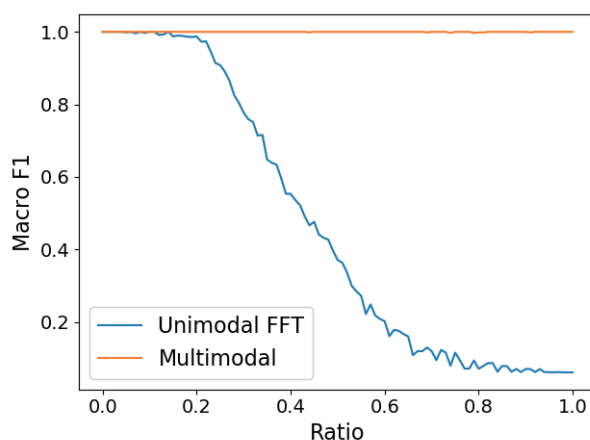


Fig.7 Macro F1 score under disturbance on FFT (ACC)

次に、AMP のみ置換を行った場合を図8に示す。ユニモーダルは僅かな置換に対して精度低下が顕著に表れるが、マルチモーダルでは置換率が90%程度まで精度を維持している。この場合も、他モードによる補完が働いていると考えられる。

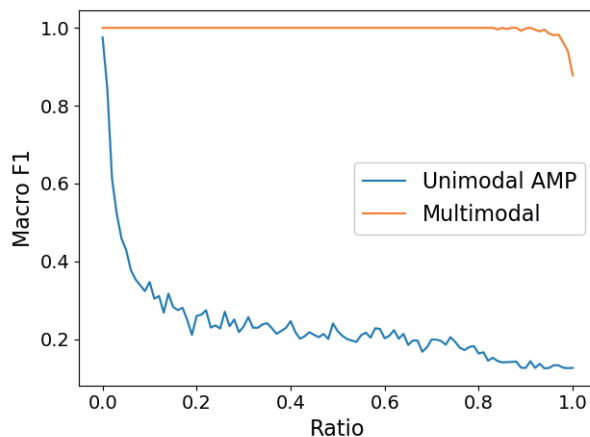


Fig.8 Macro F1 score under disturbance on AMP (ACC)

続いて WAVE のみ置換を行った場合を図 9 に示す。ユニモーダルは置換率が 20%あたりから、またマルチモーダルは 30%あたりから精度低下が見られる。WAVE における置換は、マルチモーダルにおいても精度低下を招き、FFT や AMP とは傾向が異なる。ただし、精度はマルチモーダルの方がユニモーダルと比べて全域で高い。以上より、ACC のクラス分類では、WAVE が主な役割を果たしていると言える。

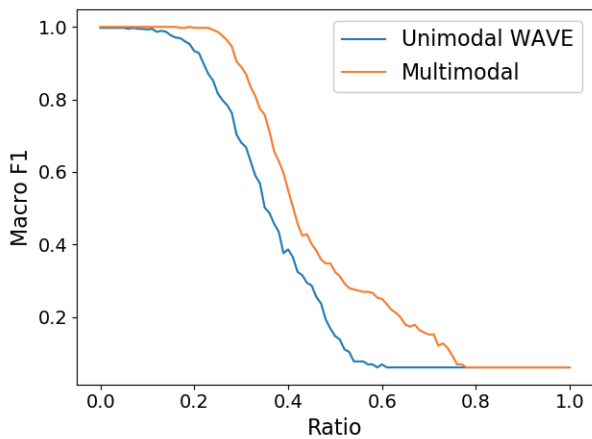


Fig.9 Macro F1 score under disturbance on WAVE (ACC)

次に、VEL データにおける比較を図 10、11、12 に示す。何れもマルチモーダルの方が外乱に強いことが分かる。特に VEL では、マルチモーダルにおいて WAVE だけでなく、AMP にも置換の影響が見られる。また、その分 WAVE の精度低下が抑えられている可能性もある。これはクラス分類を行う上で、時系列波形の形状 (WAVE) と波形振幅の長期的変動 (AMP) がその役割を担っていることを意味する。

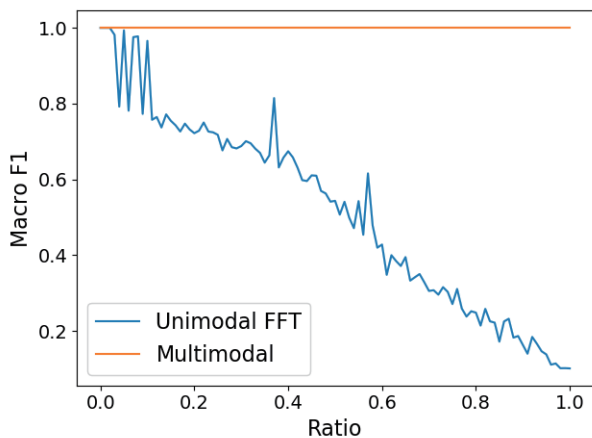


Fig.10 Macro F1 score under disturbance on FFT (VEL)

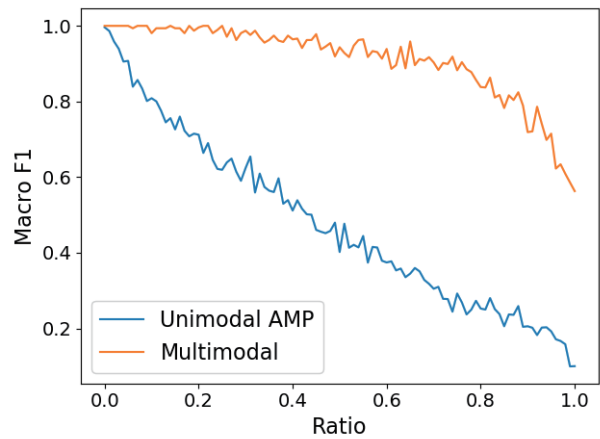


Fig.11 Macro F1 score under disturbance on AMP (VEL)

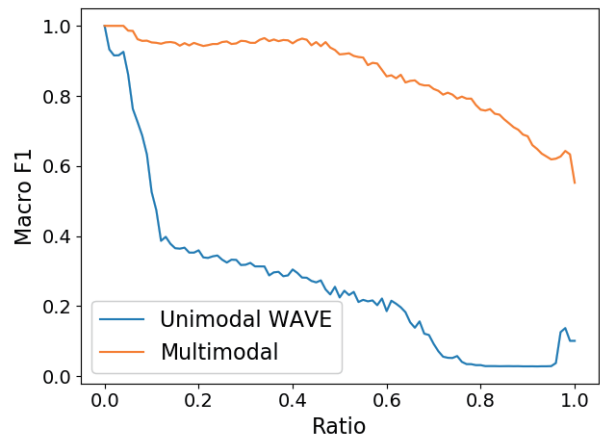


Fig.12 Macro F1 score under disturbance on WAVE (VEL)

最後に、ENV データにおける比較を図 13、14、15 に示す。ここでは、FFT と WAVE における精度低下が見られるため、時系列波形の形状と周波数の特徴が、クラス分類に影響していると考えられる。

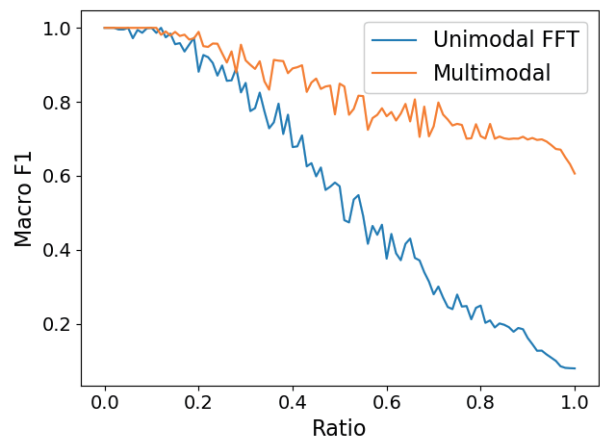


Fig.13 Macro F1 score under disturbance on FFT (ENV)

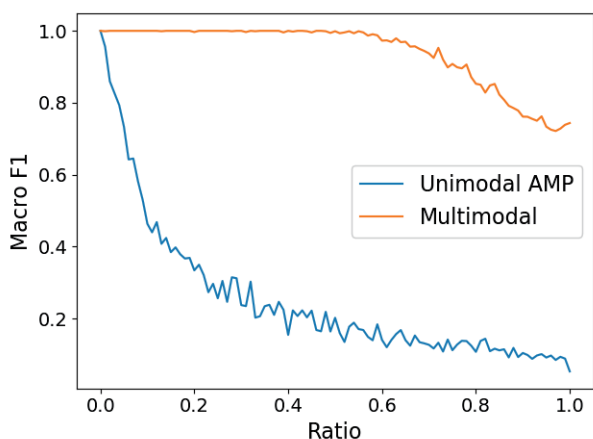


Fig.14 Macro F1 score under disturbance on AMP (ENV)

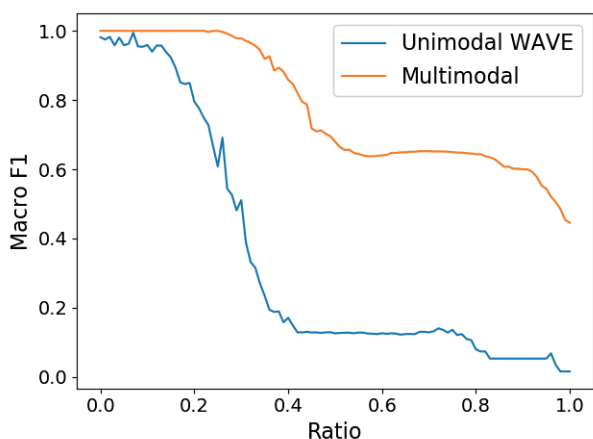


Fig.15 Macro F1 score under disturbance on WAVE (ENV)

以上より、扱ったデータにおいて、マルチモーダルアプローチは、ユニモーダルと比較をしてロバスト性が高いことが分かる。更に、振動データから現象を特定する上で、加速度 (ACC) から時系列波形の形状 (WAVE)、速度 (VEL) から波形形状 (WAVE) と波形振幅分布 (AMP)、包絡線 (ENV) から波形形状 (WAVE) と周波数スペクトル (FFT) が分類に影響していることが分かる。

5. 実機シミュレータにおける検証

5.1 実機シミュレータ

今回作成した学習モデルの有効性を確認すべく、データを取得した装置と同規模の実機シミュレータ (出力 1.5kW、極数 4P の三相誘導電動機を搭載) を用いて実験を行った (図 16)。シミュレータは、再現性を確保しながら、軸受部の摩耗やアンバランス、外輪クリープ等の不具合を任意に発生させられる構造を持つ。

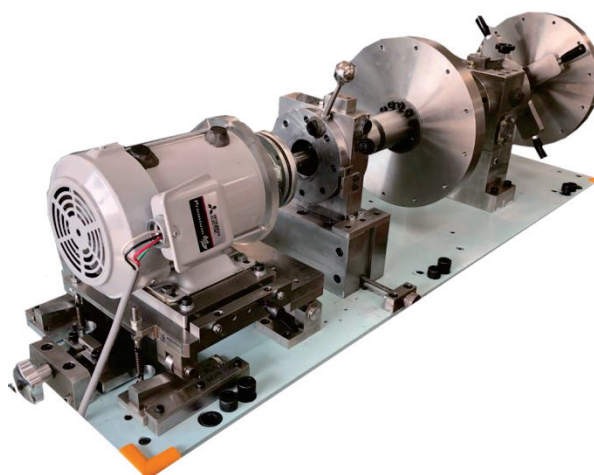


Fig.16 Appearance of the rotation machinery simulator

5.2 転移学習

汎用片持ち送風機の測定データを用いて作成した学習モデルを、シミュレータの状態判定に適用するにあたり、ドメイン適用を考慮する必要がある。近年、ドメイン適用に関して様々な報告があげられているが[15, 16, 17]、ここでは一般的な転移学習の手法を用いた。具体的には、マルチモーダル・ディープニューラルネットワークの前段である畳み込み層の重みを固定した上で、後段の全結合層を、シミュレータより測定した新たなデータを用いて学習させた。

シミュレータにおいても、ACC、VEL、ENV を取得し、各々に対して学習モデルを作成した。何れも同様の傾向であるが、ここでは前述のとおり ENV の計算結果を示す。分類クラスは、新たなドメインで再定義しており、良好域 (2)、fr アンバランス (3)、fr2 倍 (4)、fr3 倍 (6)、軸受部摩耗の兆候、揺れ (11)、軸受部摩耗の兆候、アンバランス (13)、ベルト駆動の揺れやすい構造 (15) である。

転移学習における学習モデルをテストデータにて評価した際の混同行列を図 17 に、そして精度一覧を表 6 に示す。本来、良好域にあるデータが、その他のクラスとして誤認識されたケースがいくつか見られるが、おおよそは正しく分類されている。しかし、シミュレータの動作安定性や不具合再現性を考慮すると、一定の予測精度を求める際は、必ずしも汎化性の高い学習モデルは必要ない。これは以前と同様に特徴空間でのデータ分布がクラス毎に局在化していると考えられるためである。これより、今回のモデルは精度こそ高いが、このまま一般の設備に当てはまるとは言えず今後の課題である。

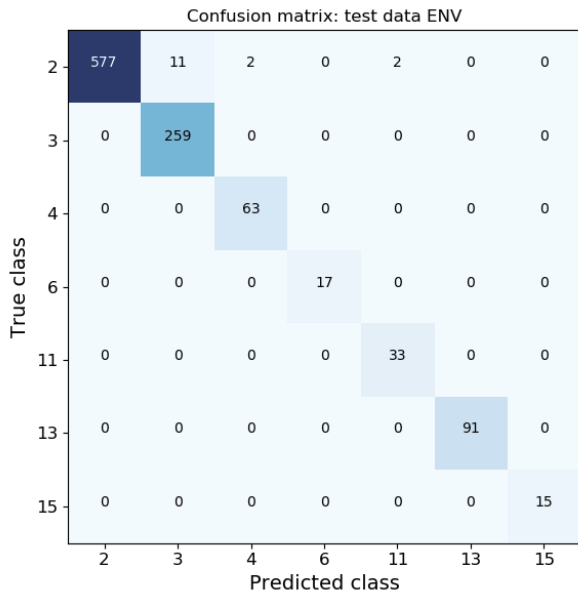


Fig.17 Confusion Matrix of ENV simulator data

Table 6 Accuracy of ENV simulator data

Class	Precision	Recall	F-measure	Accuracy	Entry
2	1.000	0.975	0.987	0.986	592
3	0.959	1.000	0.979	0.990	259
4	0.969	1.000	0.984	0.998	63
6	1.000	1.000	1.000	1.000	17
11	0.943	1.000	0.971	0.998	33
13	1.000	1.000	1.000	1.000	91
15	1.000	1.000	1.000	1.000	15

6. まとめ

熟練者頼みの振動診断の現状改善に少しでも貢献すべく、SA-FRONTIER-PROの開発からその課題と対策に付き考えた。特に人手を必要とする波形形状のパターン認識を自動化する上で、マルチモーダルディープニューラルネットワークの手法の有用性を検討した。

現時点では、振動診断カルテのデータベースから一部のデータに対して教師ラベルの付加が完了した状況であり、現実的な汎化性のある学習モデルの作成には至っていない。しかし、マルチモーダルのロバスト性の検証結果からは、そのアプローチに一定の効果があると考えている。また、転移学習と併せて実機シミュレータを用いた実験では、十分な精度で現象を推論することが確認出来た。

今後は、教師データを拡充し、本手法にてどの程度汎化性のあるモデルが作成可能であるか検討していく。また、今回はACC、VEL、ENVを個別に扱ったが、次の段階では総合的に判定できるようマルチモーダルのアプローチを広げて実装していく。

謝辞

本研究は、株式会社ジェイテック及び株式会社JFEメカフロント福山の協力のもと行われた。ここに感謝の意を表します。

参考文献

- [1] SA-FRONTIER-PRO, <http://sawada-ts.co.jp>
- [2] Hinton, G. E., Osindero, S. and Teh, Y. "A fast learning algorithm for deep belief nets", *Neural Computation*, Vol.18, 2006, pp.1527-1554.
- [3] Y. Wu, et al., "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation", arXiv preprint, arXiv:1609.08144v2
- [4] O. Vinyals, A. Toshev, S. Bengio, D. Erhan, "Show and Tell: A Neural Image Caption Generator", *ICML*, 2015
- [5] M. Bojarski, et al., "End to End Learning for Self-Driving Cars", arXiv preprint, 2016, arXiv:1604.07316v1
- [6] A. Shaheryar, X. Yin, W. Y. Ramay, "Robust Feature Extraction on Vibration Data under Deep-Learning Framework: An Application for Fault Identification in Rotary Machines", *International Journal of Computer Applications*, 2017, Vol.167, No.4, pp.37-45
- [7] 凌元錦, 真木紘, 遊佐訓孝, 宮健三, "TBMとCBMを併用した保全方式におけるTBM実施間隔最適化モデルの開発", *日本原子力学会和文論文誌*, 2007, Vol.6, No.3, pp.312-319
- [8] 豊田利夫, "設備診断技術(その2) 振動音響解析による回転機械の診断技術", *安全工学*, 1997, Vol.36, No.5, pp.321-328
- [9] D. Abati, A. Porrello, S. Calderara, R. Cucchiara, "Latent Space Autoregression for Novelty Detection", arXiv preprint, 2019, arXiv:1807.01653v2
- [10] R. Chalapathy, S. Chawla, "Deep Learning for Anomaly Detection: A Survey", arXiv preprint, 2019, arXiv:1901.03407v2
- [11] S. Yang, G. Peng, "D-PCN: Parallel Convolutional Networks for Image Recognition via a Discriminator",

arXiv preprint, 2018, arXiv:1711.04237v3

- [12] K. Kowsar, M. Heidarysafa, D. E. Brown, K. J. Meimandi, L. E. Barnes, "RMDL: Random Multimodel Deep Learning for Classification", Proceedings of the 2nd International Conference on Information System and Data Mining, 2018, pp.19-28
- [13] J. Choi, J. Lee, "EmbraceNet: A robust deep learning architecture for multimodal classification", Information Fusion, Vol.51, 2019, pp.259-270
- [14] K. Noda, H. Arie, Y. Suga, T. Ogata, "Multimodal integration learning of robot behavior using deep neural networks", Robotics and Autonomous Systems, Vol.62, 2014, pp.721-736
- [15] A. R. Zamir, A. Sax, W. Shen, L. Guibas, J. Malik, S. Savarese, "Taskonomy: Disentangling Task Transfer Learning", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018
- [16] K. Saenko, B. Kulis, M. Fritz T. Darrell, "Adapting Visual Category Models to New Domains", Computer Vision – ECCV, 2010, pp.213-226
- [17] K. Saito, K. Watanabe, Y. Ushiku, T. Harada, "Maximum Classifier Discrepancy for Unsupervised Domain Adaptation", CVPR.2018.00392, 2018, pp.3723-3732