

令和 6 年 5 月 1 日現在

機関番号：10101

研究種目：若手研究

研究期間：2020～2023

課題番号：20K19856

研究課題名（和文）技術者の認知・判断・行動プロセスに基づくハイパーマルチモーダル画像分類技術の構築

研究課題名（英文）Construction of hyper-multimodal image classification technology based on the cognitive, decision-making, and behavioral processes of engineers

研究代表者

前田 圭介（Maeda, Keisuke）

北海道大学・情報科学研究所・特任助教

研究者番号：20798243

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では、AI技術の実社会応用へ向けて、技術者から取得する様々な生体情報・専門画像・それに対する判断結果を用い、認知・判断・行動プロセスを組み込んだハイパーマルチモーダル画像分類技術を構築した。具体的に、複数の生体情報から技術者に共通の知識及び経験を抽出することで、技術者に近い判断が可能な信頼性が担保された専門画像分類技術を構築した。これにより、技術者がどこに注目し・どのように判断したかを学習可能となることに加え、実際に構築したモデルからその判断根拠を提示可能となり、高い解釈性・信頼性を有するAIを実現した。

研究成果の学術的意義や社会的意義

従来のAI分野では、人間の視覚野における局所受容野の働きを模擬した学習手法が提案されているものの、これらは構造の模倣に留まっていた。近年、脳波や視線情報などの人間から得られる生体情報を用いて、画像分類の精度向上を目的とする研究が進められているが、これらは生体情報の利用による判断結果の取得のみに留まっている。そこで、本研究では、技術者の認知・判断・行動プロセスに注目することで、生体特徴と判断結果の関係性の抽出とAIの判断結果の抽出を融合した画像分類技術を構築した。本技術は、様々な専門分野に横展開可能であり、技術継承問題の解決策としての利活用も期待されることから、高い学術的意義を有する。

研究成果の概要（英文）：This study constructed a hyper-multimodal image classification technology that incorporates cognitive, decision-making, and behavioral processes, using various biometric data, images, and classification results obtained from engineers, for real-world applications of AI technology. Specifically, by extracting knowledge and experience common to engineers from multiple biometric data, a reliable image classification technology that enables judgments similar to those of engineers can be constructed. This enables the AI to learn what engineers paid attention to and how they made judgments, and to present the basis for their judgments from the actually constructed model, thus realizing an AI with high interpretability and reliability.

研究分野：画像認識

キーワード：画像分類 生体情報 機械学習 解釈性 暗黙知 深層学習 信号処理 マルチモーダル

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

現在、AI 技術を用いたデータ解析の需要は土木工学・医学・惑星科学などの専門分野においてますます高まっており、業務効率化のために、専門性の高い画像（以降、専門画像）を対象とした分類技術の実現が求められている [1]。高度な業務への AI 導入が期待される専門分野では、高い精度のみならず、結果に対する確かな信頼性を有するモデルの構築が不可欠である。しかしながら、従来の機械学習では、機械と技術者の判断の関係性を明らかにする理論の構築には至っておらず、機械の判断に対する信頼性が担保されないという問題が存在する。この問題を解決し、AI 技術の実用化を促進するためには、以下の 4 点を組み込んだ機械学習理論の構築が必要となる。

【要点 1】 経験や知識の基となる情報を機械に入力可能な特徴へ変換するモデル

【要点 2】 経験や知識を表す生体特徴を抽出することで技術者に近い判断が可能なモデル

【要点 3】 画像分類で判断する際の判断根拠を技術者が理解可能な形で提示するモデル

【要点 4】 得られた結果が誤っていた場合に、効率よく再学習可能なモデル

上記 4 点の技術を構築することで、技術者がどこに注目し・何を考えたかを学習可能となることに加え、実際に構築したモデルからその判断根拠を提示することで高い解釈性・信頼性を有する AI 技術を実現できると考え、研究を遂行してきた。

[1] 平成 28 年版情報通信白書 第 1 部 第 4 章 第 2 節 人工知能 (AI) の現状と未来

### 2. 研究の目的

本研究の目的は、AI 技術の実社会応用へ向けて、技術者から取得する様々な生体情報・専門画像・それに対する判断結果を用い、認知・判断・行動プロセスを組み込んだハイパーマルチモーダル画像分類技術を構築することで、人間の経験や知識を機械学習に導入した、高い解釈性・信頼性を提供する AI 技術を実現することである。そこで、以下の 4 つの【解決策】により、「1. 研究開始当初の背景」に示した【要点】を満たす機械学習理論を導出する。

【解決策 1】 複数の技術者から得られる種々の生体情報からノイズ・個人差を除去

【解決策 2】 複数の技術者に共通する特徴を見出し、技術者の判断との間の因果関係を説明

【解決策 3】 判断・予想結果に対する判断根拠の可視化機構を導入

【解決策 4】 技術者からのフィードバック結果を用いたモデルパラメータの転移学習

本研究で構築する AI 理論は、今後懸念されている熟練技術者不足に伴う若手技術者への技術継承問題の解決策としての利活用が期待されるだけでなく、分野を問わず適用可能、つまり、様々な専門分野に対して横展開可能な極めて汎用性の高い技術基盤となることから、創造性の高い研究である。

### 3. 研究の方法

本研究では、AI 技術の実社会応用へ向けて、技術者の生体情報・専門画像・判断結果の間の関連性分析により、経験や知識に関連する特徴を抽出し、技術者の判断との因果関係を学習可能なモデルを構築する。さらに、モデルの判断根拠の可視化技術を構築する。最後に、構築した AI を持続的に学習可能とするために、モデルの再学習手法を構築する。本研究は、社会インフラ維持管理の分野を実験フィールドとし、その維持管理業務における技術者を対象とする。本研究は、「1. 研究開始当初の背景」および「2. 研究の目的」に記載した【要点】【解決策】に対応しており、以下の 4 つのフェーズから構成されている。

【フェーズ 1】 複数の技術者から多種多様な生体情報の取得およびその特徴量化

技術者から、視線・脳活動・動作などの生体情報と遂行業務に対する判断を取得した。さらに、技術者に共通する特徴を明らかにするために、複数の技術者から得られる生体データと遂行業務に対する判断との間の関連性について分析を実施した。これにより、一部の生体データが AI 構築に特に有効であることが示唆され、これらのデータの特徴量化を行った。

【フェーズ 2】 生体特徴・専門画像・技術者の判断から推論可能なモデルの構築

【フェーズ 1】 で取得した複数の技術者の生体特徴とその技術者が熟練であるか否かを表す情報を用い、技術者が共通に有する特徴を算出可能なモデルを構築することで、技術者の判断と熟練度合いの関係性および種々の結果に関する推論を可能とした。

【フェーズ 3】 構築したモデルの判断根拠の可視化

モデルが予測するクラスの自信の度合いが大きい部分を推定可能な勾配計算理論を確立し、判断根拠を可視化する深層学習モデルを構築した。さらに、文字形式で判断理由を提示可能とするためのモデルを構築した。

#### 【フェーズ4】技術者からのフィードバックを再学習可能な転移学習モデルの構築

一度学習したモデルについて、新たに取得したデータを追加で学習可能な転移学習の枠組みを構築することで、結果に対するフィードバックに基づく再学習を可能とした。

以上、これまでに構築してきた理論によって、複数の生体情報から技術者に共通の知識及び経験を抽出することが可能になった。これにより、高い解釈性・信頼性を有するAI技術が実現した。

#### 4. 研究成果

本研究では、技術者の認知・判断・行動プロセスを生体情報解析により抽出し、それに基づくハイパーマルチモーダル画像分類技術を構築した。具体的に、複数の技術者から取得した多種多様な生体情報について、ノイズ・個人差を低減しながら、経験や知識に関する情報を抽出するための相関分析理論を導出した。また、構築した分類モデルにおいて、「何を見て」、「どのように考えたか」を明らかとするために、モデルが分類するクラス的确信度が高くなるように領域に重み付けし学習することで、正確に注目領域を予測可能とした。最後に、判断根拠を文字形式として出力することで、技術者への提示を可能とし、フィードバックを新たなデータとして学習可能なモデルを実現した。本研究の遂行によって、人間の知識の導入理論および実社会応用に耐え得る信頼性が担保された機械学習理論が構築され、高精度な画像分類技術が実現されている。さらに、本研究の基礎技術は、分野を問わず適用可能であり、他分野へ横展開可能な極めて汎用性の高い技術基盤となっている。

この研究を実施する過程で、インフラ維持管理現場におけるデータ取得実験において、技術者間で点検行動に大きな差があることが明らかとなった。さらに、実験協力者の技術者に対するインタビューによって、この差は“個人差”ではなく、技術者がノウハウを獲得するまでに培ってきた経験に大きく依存することが判明した。したがって、AIへ入力するための画像の撮影方法自体に技術者毎に差が生じることから、撮影された損傷画像をそのまま活用することを前提とする従来のAIモデルでは、精度向上に限界が存在することが新たな課題として表出した。そこで、申請者は、当該研究で進めてきたAIの注目領域を推定可能なアプローチを拡張し、点検時に得られる映像に適用することで、AI構築に適した画像を撮影できるのではないかと考えた。このように、当該研究で構築したアプローチを基盤とすることで、AIの実社会応用で生じる画像撮影方法の多様性の問題を解決する新たな研究課題の遂行に関する目途が立ったことから、基盤研究(C)への前年度応募を提案し採択された。現在は、基盤研究(C)「画像認識の高度化に向けた画像の撮影方法を最適化する異環境異種データ適応型AIの構築」にて本研究成果の発展的内容について取り組んでいる。

5. 主な発表論文等

〔雑誌論文〕 計25件（うち査読付論文 25件 / うち国際共著 0件 / うちオープンアクセス 21件）

1. 著者名 小川 直輝、前田 圭介、小川 貴弘、長谷山 美紀	4. 巻 3
2. 論文標題 異種特徴間の相関およびAttention Mapの確信度を考慮可能な変状画像の劣化レベル分類	5. 発行年 2022年
3. 雑誌名 AI・データサイエンス論文集	6. 最初と最後の頁 704 ~ 713
掲載論文のDOI (デジタルオブジェクト識別子) 10.11532/jsceiii.3.J2_704	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 櫻井 慶悟、前田 圭介、藤後 廉、小川 貴弘、長谷山 美紀	4. 巻 3
2. 論文標題 地下鉄トンネル点検時の一人称視点映像を用いたVision Transformerに基づく変状検出	5. 発行年 2022年
3. 雑誌名 AI・データサイエンス論文集	6. 最初と最後の頁 470 ~ 478
掲載論文のDOI (デジタルオブジェクト識別子) 10.11532/jsceiii.3.J2_470	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Maeda Keisuke, Takada Saya, Haruyama Tomoki, Togo Ren, Ogawa Takahiro, Haseyama Miki	4. 巻 22
2. 論文標題 Distress Detection in Subway Tunnel Images via Data Augmentation Based on Selective Image Cropping and Patching	5. 発行年 2022年
3. 雑誌名 Sensors	6. 最初と最後の頁 8932 ~ 8932
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s22228932	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Higashi Takaaki, Ogawa Naoki, Maeda Keisuke, Ogawa Takahiro, Haseyama Miki	4. 巻 23
2. 論文標題 Estimation of Degradation Degree in Road Infrastructure Based on Multi-Modal ABN Using Contrastive Learning	5. 発行年 2023年
3. 雑誌名 Sensors	6. 最初と最後の頁 1657 ~ 1657
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s23031657	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ogawa Naoki, Maeda Keisuke, Ogawa Takahiro, Haseyama Miki	4. 巻 22
2. 論文標題 Deterioration Level Estimation Based on Convolutional Neural Network Using Confidence-Aware Attention Mechanism for Infrastructure Inspection	5. 発行年 2022年
3. 雑誌名 Sensors	6. 最初と最後の頁 382 ~ 382
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s22010382	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ogawa Naoki, Maeda Keisuke, Ogawa Takahiro, Haseyama Miki	4. 巻 9
2. 論文標題 Distress Image Retrieval for Infrastructure Maintenance via Self-Trained Deep Metric Learning Using Experts' Knowledge	5. 発行年 2021年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 65234 ~ 65245
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2021.3074019	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Maeda Keisuke, Takahashi Sho, Ogawa Takahiro, Haseyama Miki	4. 巻 -
2. 論文標題 Deterioration level estimation via neural network maximizing category-based ordinally supervised multi-view canonical correlation	5. 発行年 2020年
3. 雑誌名 Multimedia Tools and Applications	6. 最初と最後の頁 1 ~ 22
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s11042-020-10040-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Moroto Yuya, Maeda Keisuke, Ogawa Takahiro, Haseyama Miki	4. 巻 8
2. 論文標題 Human-Centric Emotion Estimation Based on Correlation Maximization Considering Changes With Time in Visual Attention and Brain Activity	5. 発行年 2020年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 203358 ~ 203368
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2020.3036908	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 東 孝明、小川 直輝、前田 圭介、小川 貴弘、長谷山 美紀	4. 巻 4
2. 論文標題 データ横断型対照学習を用いた道路構造物における変状画像の劣化レベル分類	5. 発行年 2023年
3. 雑誌名 AI・データサイエンス論文集	6. 最初と最後の頁 44～57
掲載論文のDOI (デジタルオブジェクト識別子) 10.11532/jsceiii.4.2_44	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

1. 著者名 Feng Yuhu, Maeda Keisuke, Ogawa Takahiro, Haseyama Miki	4. 巻 13
2. 論文標題 Gaze-Dependent Image Re-Ranking Technique for Enhancing Content-Based Image Retrieval	5. 発行年 2023年
3. 雑誌名 Applied Sciences	6. 最初と最後の頁 5948～5948
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/app13105948	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

〔学会発表〕 計66件(うち招待講演 0件/うち国際学会 34件)

1. 発表者名 Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Gaussian distributed graph constrained multi-modal Gaussian process latent variable model for ordinal labeled data
3. 学会等名 2022 IEEE International Conference on Image Processing (ICIP 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 Kaito Hirasawa, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 A trial of fine-grained classification of expert-novice level using bio-signals while inspecting subway tunnels
3. 学会等名 IEEE Global Conference on Consumer Electronics (国際学会)
4. 発表年 2021年

1. 発表者名 Yuya Moroto, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Few-shot personalized saliency prediction using person similarity based on collaborative multi-output Gaussian process regression
3. 学会等名 IEEE International Conference on Image Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Naoki Ogawa, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Correlation-aware attention branch network using multi-modal data for deterioration level estimation of infrastructures
3. 学会等名 IEEE International Conference on Image Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Kaito Hirasawa, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Time-lag aware multi-modal variational autoencoder using baseball videos and tweets for prediction of important scenes
3. 学会等名 IEEE International Conference on Image Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Kyohei Kamikawa, Maeda Keisuke, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Interest level estimation via multi-modal Gaussian process latent variable factorization
3. 学会等名 IEEE International Conference on Image Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Takaaki Higashi, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Estimation of visual features of viewed image from individual and shared brain information based on fMRI data using probabilistic generative model
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Yusuke Akamatsu, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Classification of expert-novice level using eye tracking and motion data via conditional multimodal variational autoencoder
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (国際学会)
4. 発表年 2021年

1. 発表者名 Kyohei Kamikawa, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
2. 発表標題 Feature integration via semi-supervised ordinally multi-modal Gaussian process latent variable model
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (国際学会)
4. 発表年 2021年

1. 発表者名 平澤 魁人, 前田 圭介, 藤後 廉, 小川 貴弘, 長谷山 美紀
2. 発表標題 地下鉄トンネル点検時の技術者から取得される生体信号と技術者の点検行動の関連性分析
3. 学会等名 映像情報メディア学会技術報告
4. 発表年 2022年



1. 発表者名 小川 直輝, 前田 圭介, 小川 貴弘, 長谷山 美紀
2. 発表標題 Attention map に対する確信度を考慮可能な深層学習を用いた変状分類の高精度化に関する検討
3. 学会等名 映像情報メディア学会技術報告
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<https://www-lmd.ist.hokudai.ac.jp/member/keisuke-maeda/>  
北海道大学大学院情報科学研究院メディアダイナミクス研究室 研究室メンバー

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関