HOKKAIDO UNIVERSITY

Title	構造物の維持管理支援のための異種特徴を用いた機械学習に基づく変状分類に関する研究
Author(s)	前田, 圭介
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第13523号
Issue Date	2019-03-25
DOI	https://doi.org/10.14943/doctoral.k13523
Doc URL	https://hdl.handle.net/2115/74261
Type	doctoral thesis
File Information	Keisuke_Maeda.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

博 士 論 文

構造物の維持管理支援のための異種特徴を用いた機械学習に基づく変状分類に関する研究



北海道大学 大学院情報科学研究科
メディアネットワーク専攻

前田 圭介

2019 年 3 月

目 次

第 1 章　序論　1

1.1　本研究の背景 1

1.2　本研究の目的 2

1.3　論文の構成 4

第 2 章　機械学習を用いた構造物の維持管理支援に関する従来研究　6

2.1　はじめに 6

2.2　パッチ分割を用いた従来研究 7

2.3　Data augmentation を導入した従来研究 9

2.4　学習済み CNN を用いた転移学習に関する従来研究 10

2.5　本論文で解決すべき課題と解決法 12

2.6　まとめ 14

第 3 章　異種特徴間の関係を表現可能なベイジアンネットワークに基づく変状分類　15

3.1　はじめに 15

3.2　Text-based BN の構築と信頼度の算出 17

3.3　Multi-feature BN の構築 17

3.4　適応的なベイジアンネットワークモデルの選択 23

3.5　実験 25

3.5.1　実験準備 25

3.5.2　実験結果および考察 29

3.6　まとめ 37

第 4 章　異種特徴の正準相関を考慮した変状分類　38
4.1　はじめに 38
4.2　画像特徴の算出およびテキスト特徴の算出 40
4.3　相関関係を考慮することによる変状に関連のある特徴の算出 . . . 41
4.4　得られた特徴を用いた変状分類 43
4.5　実験 46
4.5.1　実験準備 46
4.5.2　実験結果および考察 53
4.6　まとめ 57

第 5 章　異種特徴間の関連性を考慮可能な深層学習に基づく変状分類　60
5.1　はじめに 60
5.2　LRF に基づく特徴抽出 62
5.3　正準相関を最大化する隠れ層の構築 64
5.4　DELM の構築 67
5.5　実験 70
5.5.1　実験準備 70
5.5.2　実験結果および考察 77
5.6　まとめ 82

第 6 章　結論　86
6.1　本研究の総括 86
6.2　本研究の今後の課題 87

謝辞　88

備考　90

参考文献　91

著者の研究業績　100

図 目 次

1.1 本研究における課題とその解決策および章構成の関係. 4

3.1 実際の維持管理業務における点検の流れ. 16
3.2 提案手法全体の概要図. 16
3.3 顕著性マップに基づく変状領域の抽出の例. 18
3.4 本手法の有効点である適応的なベイジアンネットワークモデル選択の詳細な処理. 25
3.5 変状画像の例. (a) および (c) は近景, (b) および (d) は遠景 27
3.6 変状画像の例. 31
3.7 比較手法 1 から 4 および提案手法の正解率の比較. 34
3.8 本手法により正しく分類された変状画像: (a) – (h) は元画像であり, (a') – (h') は領域が抽出された後の画像である. 35
3.9 比較手法 A, B および提案手法の正解率の比較. 37

4.1 提案手法の概要図. 本手法は 3 つのステップから成る. 1 つ目は, 特徴算出. 2 つ目は画像特徴とテキスト特徴間の正準相関を最大化する変換を行う隠れ層の構築. 最後に, WELM に基づく最終層の重みの導出を行う. 39
4.2 変状画像の例. ただし, 表 4.2 に示す ID に対応している. 42
4.3 正準特徴の各次元に対応する正準相関係数. 54
4.4 提案手法によって正しく正解された変状画像の例. 58
4.5 異なるアングルから撮影された腐食画像の例. 59

5.1 本手法の概要図を示す. まず, LRF を用いて画像特徴を算出する. 次に, 画像特徴およびテキスト特徴間の正準相関を最大化する層を構築する. 最後に, 順伝搬型の深層学習である Deep Extreme Learning Machine (以降, DELM) を構築する. 61
5.2 テキスト特徴の算出方法の一例. 図における画像 ID は各画像を表し, それぞれにテキストデータが対応している. 64
5.3 提案手法によって正しく推定された画像の例. ただし, 本実験では, 画像中央から 50×50 $(= d_w \times d_h)$ ピクセルを用いている. 78
5.4 元特徴とクラスラベルとの間のピアソンの相関係数を示す. また, “Var” は係数の分散を表し, “AveA” は係数の絶対値の和の平均を表す. (a), (b) および (c) は比較手法 3, 4 および 5 にそれぞれ対応している. 84
5.5 正準特徴とクラスラベルとの間のピアソンの相関係数を示す. また, “Var” は係数の分散を表し, “AveA” は係数の絶対値の和の平均を表す. (a), (b) および (c) は比較手法 1, 2 および提案手法にそれぞれ対応している. 85

表 目 次

3.1 本手法で用いる特徴とその次元数. 18
3.2 点検データの例. 20
3.3 実験で用いる点検データ数. 26
3.4 1 つ目の実験で用いる比較手法. 28
3.5 2 つ目の実験で用いる比較手法. 29
3.6 各手法の Recall, Precision および F 値 ($\Phi = 5$). 32
3.7 各手法の Recall, Precision および F 値 ($\Phi = 7$). 33
3.8 各手法の Recall, Precision および F 値 ($\Phi = 5$). 36
3.9 各手法の Recall, Precision および F 値 ($\Phi = 7$). 36
3.10 本手法における計算時間. 36

4.1 本手法で用いる画像特徴の詳細とその次元. 40
4.2 点検データ中の点検項目の例. 表中の ID は, 図 4.2 の変状画像の例と対応している. 41
4.3 本実験で用いる点検データの数. 47
4.4 パラメータの探索範囲. 47
4.5 提案手法で用いられるパラメータの詳細. 47
4.6 正解のクラスと推定クラスとの間の関係. 48
4.7 全ての手法における Recall, Precision および F 値. 52
4.8 クラス均衡とクラス不均衡なデータにおける CMWELM の精度比較. 53
4.9 提案手法の混同行列. 55
4.10 WELM の混同行列. 55

4.11 CMWELM と VGG16 の計算時間 (sec). 56

5.1 テキストデータの一例およびその次元数を示す. 実際には, 送電線には名前が記録されており, 塩害区分は塩害の影響を受ける地域を表す. 63
5.2 本実験で用いる画像数. 71
5.3 提案手法で用いるパラメータの詳細. 72
5.4 本実験で用いる手法の詳細. 73
5.5 本実験で用いる手法のネットワーク構成を示す. “conv”, “norm”, “pool” は, 畳み込み層, 局所応答正規化層, プーリング層を表し, “pca”, “cca” は, PCA, CCA を表す. “fc-elm”, “fc” は ELM に基づく全結合層, 一般的な全結合層を表す. 74
5.6 全ての手法における Recall, Precision および F 値. 76
5.7 それぞれの手法における計算コストを示す. 78
5.8 提案手法, Caffe-Net および VGG16 の学習時間 (sec). 79
5.9 提案手法, Caffe-Net および VGG16 のテスト時間 (sec). 79

第1章　序論

1.1　本研究の背景

現存する橋梁やトンネルなどの構造物の多くは, 高度経済成長期に集中的に建設され, 現在, 急速な老朽化の時期を迎えている. 平成 29 年度国土交通白書では, 建設から 50 年以上が経過する構造物の割合は今後 15 年間で急増すると述べられている [1]. 例えば, 橋梁におけるその割合は, 2018 年に約 25%であったものが, 2033 年には 63%に達するとされている. また, 少子高齢化の進行により生産年齢人口が 1995 年をピークに減少に転じており，構造物の維持管理の分野においても，高齢者の大量離職による中長期的な技術者不足が生じることが懸念されている [1]．そこで，構造物の急速な老朽化に伴う維持管理・更新費用の増大および将来的な技術者不足に対処するため，国土交通省では，予防保全等の計画的なメンテナンスによる費用の削減や業務の効率化を目的として,「インフラメンテナンス革命」を推進し，異分野との連携による効率的な構造物の維持管理の実現を目指している [2]．このような背景から，構造物の点検の無人化や診断の自動化に向けた取り組みがなされており，その実現のため，機械学習を用いたデータ解析技術の創出が求められている．従来より, 構造物に発生する異常 (以降, 変状) が撮影された画像 (以降，変状画像) を用いて, 機械学習により, 構造物に発生する頻度の高いひびわれやポットホール等の特定の変状に注目した検出が盛んにおこなわれてきた [3, 4, 5]. 一方, 構造物には上述の変状以外にも様々な種類の変状が発生し, その劣化の度合いも多様であることから, 近年では, 変状の種類や劣化の度合いを自動分類 (以降, 変状分類) する手法が提案されている [6, 7]. 以上より, 変状分類の実現により技術者の負担を軽減することが可能となるため, 構造物の維持管理支援において有効な手段の一つである.

画像の自動分類は従来より一般物体認識の分野で活発に研究が行われており, これらは変状分類を行う際にも基礎的な技術となる. これらの画像分類においては, 深層学習の導入により様々なタスクの精度向上が報告されている. 特に, Convolutional Neural Network (以降, CNN) [8] は, 入力画像と対応する出力ラベルの組を用い, 多数の隠れ層を構築することで分類に有効な弁別性の高い特徴への自動変換を実現する. 一方, 効果的な特徴変換を行うためには, 隠れ層を構成する膨大なパラメータを最適に決定する必要があることから, 大量の学習画像を必要とする. しかしながら, 構造物の維持管理の分野では, 実際の維持管理業務の際に技術者が手動で変状画像を撮影するため, 大量の画像を用意することは困難である. このような問題を解決するために, 従来より, 変状領域のパッチ分割による学習画像の充足 [9, 10, 11] や変状画像に対する回転・コントラスト調整などによる学習画像の充足 [12, 13, 14], 一般物体認識用の画像で学習されたパラメータを用いた転移学習 [15, 16, 17] など, 少量画像に対応した様々なアプローチが研究されてきた. しかしながら, 変状画像には, 背景などの不要な領域が多く存在することや同一の変状においても多様な視覚的特徴を有することから, 学習画像の充足や転移学習の効果が低く [18], 画像のみを用いた変状分類には精度向上に限界が存在した. ところで, 実際の維持管理業務では, 技術者は変状画像の撮影に加えて, 構造物の部位や部材の名称等をテキストデータとして記録し, 変状画像とテキストデータを総合的に評価することで最終的な判断を下している. したがって, 変状分類にテキストデータを導入することはその高精度化に有効であると考えられる. そこで, これらの異なる種類のデータから得られる特徴 (以降, 異種特徴) の関連性に基づき, 分類に有効な特徴へ変換可能な射影を導出し, 深層学習の隠れ層に導入することで, 上記課題が解決可能となることが期待される.

1.2　本研究の目的

構造物の維持管理の分野では, 撮影される画像に限りがあるため, 従来より様々なタスクにおいて高精度な画像分類を達成している CNN を用いた分類は困難である. この問題を解決し, 高精度な変状分類を実現するため, 本研究では, 実際の

維持管理業務で撮影される変状画像と記録されるテキストデータから算出する異種特徴を用いて, 分類に有効な弁別性の高い特徴へ変換可能な枠組みを導入した機械学習による変状分類手法を実現することを目的とする. 本手法を構築するためには, 異種特徴および変状との間に関連性が存在することを明らかにし, その関連性に基づき効果的な特徴変換を行う射影を導出し, 深層学習へ導入する必要がある. したがって, 本論文は, 1）異種特徴間の関連性の評価, 2）分類に有効な特徴へ変換可能な射影の導出, 3）異種特徴間の関連性を考慮可能な深層学習手法の構築により構成される.

1) 異種特徴間の関連性の評価

実際の維持管理業務においては, 技術者は変状画像とテキストデータを用いることで最終的な判断を下すが, 機械学習を用いた変状分類においては, これらのデータから得られる異種特徴が有効であるかは明らかとなっていない. そこで, 異種特徴および変状間の関連性を定量的に評価可能な分類器を構築することで, 変状分類において, 異種特徴および変状間に関連性が存在することを明らかにする.

2) 分類に有効な特徴へ変換可能な射影の導出

1) により, 異種特徴および変状間に関連性が存在することが明らかとなるため, これらの関連性に基づき, 分類に有効な特徴への変換を可能にする射影を導出する. 具体的に, 異種特徴間の相関関係を考慮した特徴変換により得られる特徴は, 分類精度向上に寄与するという報告 [19] に基づき, これらの特徴間の相関関係を考慮可能な手法を導入する. これにより, 少量の学習画像を用いる場合でも, 弁別性の高い特徴への変換が可能となる.

3) 異種特徴間の関連性を考慮可能な深層学習手法の構築

2) で得られた射影を隠れ層に導入した順伝搬型の深層学習を新たに構築する. 順伝搬型の深層学習の隠れ層の重みとして異種特徴の相関関係を考慮して導出した射影を用いることで, 少ない層で分類に有効な特徴への変換を可能とすることから, 最適に決定すべきパラメータ数の削減が可能となる.

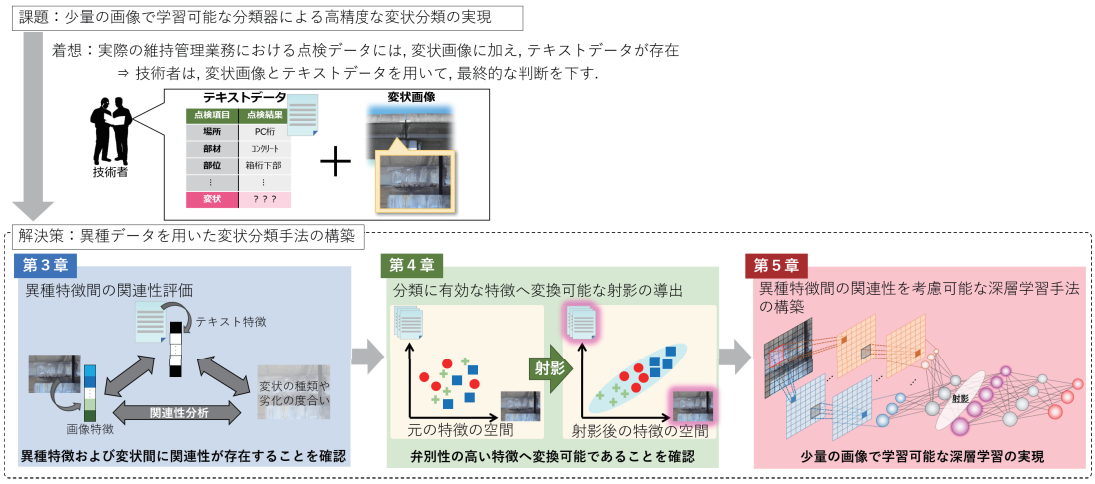


図 1.1: 本研究における課題とその解決策および章構成の関係．

以上により，異種特徴間の関連性に基づいて，高い弁別性を有する特徴への変換の枠組みを導入した深層学習を構築することで，少量の学習画像を用いた高精度な変状分類を実現する．

上記の本研究における課題とその解決策および章構成の関係を図 1.1 に示す．

1.3　論文の構成

本論文は，6 章から成り，以下に各章の概要を述べる．第 1 章では，研究背景および目的について述べた．

第 2 章では，構造物の維持管理支援を目的とした，機械学習に基づく従来研究について紹介し，本論文で解決すべき課題を明らかにする．

第 3 章では，異種特徴と変状を変数とみなすことで，それらの因果関係を確率値として表現可能なベイジアンネットワーク [20] に基づく変状分類手法を構築する．具体的に，実際の維持管理業務において，変状画像とテキストデータが技術者の判断に有効であることに着想を得て，維持管理業務を模擬した枠組みにより変状分類を行う．これにより，どの特徴がどの変状に関連があるかなど，異種特徴間の関連性の存在を明らかにすることが可能となる．

第 4 章では，異種特徴に正準相関分析 [21] を適用することで変状分類に有効な特徴を算出し, それらの特徴を用いた変状分類手法を構築する. 具体的に, 互いに高相関な特徴は分類精度向上に効果的であることから, 画像特徴とテキスト特徴を互いの相関が高くなるような空間へ変換する射影を導出する. これにより, 少量の学習画像を用いた場合でも, 弁別性の高い特徴へ変換可能となる.

第 5 章では，得られた射影を隠れ層に導入した順伝搬型の深層学習として新たに Correlation-maximizing Deep Extreme Learning Machine-local Receptive Field [22] を構築する. 順伝搬型の深層学習の隠れ層の重みとして異種特徴の相関関係を考慮して算出した射影を用いることで, 少ない層で分類に有効な特徴への変換を可能とすることから, 最適に決定すべきパラメータ数の削減が可能となる. したがって, 少量の学習画像を用いた高精度な変状分類が実現する.

第 6 章では，本論文の結論を示し, 今後の課題について述べる.

第2章　機械学習を用いた構造物の維持管理支援に関する従来研究

2.1　はじめに

本章では, 構造物の維持管理支援を目的とした, 機械学習を用いた画像解析に関する従来研究について述べる. 機械学習を用いた構造物の維持管理支援に関する研究は, 点検時に撮影される画像を用いて 1990 年代より行われてきた. 特に, 計算機を用いた構造物に発生する変状の検出, 変状の種類の分類および変状の劣化レベルの分類などに関する研究が盛んに行われてきた. 2000 年に入ると, 画像から対象物体を表すことが可能な特徴量を人手で設計し, Support Vector Machine (以降, SVM) [23] 等による機械学習手法が画像分類手法として一般的であった. 近年では, 2012 年に行われた一般画像認識のコンペティション (ILSVRC 2012) において, ニューラルネットワークの一種である CNN [8] が大幅な認識精度向上を達成したことをきっかけに, 深層学習へ注目が一斉に集まった. 2012 年以降も, 活発な研究が行われ，文献 [24, 25, 26] に代表される様々な CNN モデルが提案されており, 非常に高い精度で物体を認識することが可能となった. それ以来, 構造物の維持管理支援のための画像解析に関する研究においても CNN の導入が検討されてきた.

CNN モデルを構築するためには, 膨大な数のパラメータを最適に設定する必要があるため, 大量の学習画像を必要とする. しかしながら, 構造物の維持管理において撮影される画像の数は CNN に必要な画像数よりも少ないことから, さまざまな工夫により, CNN モデルが構築されてきた. 具体的には, パッチ分割や data augumentation による学習画像数の増加, 学習済み CNN モデルを用いた転移学習に基づく手法の 3 つの手法に大別される. そのため, 本章では, 上記 3 つの研究に

ついて説明する.

本章の構成は，以下の通りである．2.2 節では, パッチ分割を用いた従来研究について説明する．2.3 節では, data augmentation を導入した従来研究について説明する．2.4 節では, 学習済み CNN モデルを用いた転移学習に関する従来研究について説明する．2.5 節では, 従来研究より考えられる課題と本論文の解決法について検討する.

2.2　パッチ分割を用いた従来研究

2.1 節で述べた CNN の学習では大量の画像を必要とする問題に対処するために, 画像をパッチ分割することで CNN の学習で用いるサンプル数を増加させる手法が種々提案されている [9, 10, 11]. これらの手法では, 一般物体認識において高精度な画像認識を達成したと報告されている CNN モデルを用いて学習するだけではなく, 著者らの設定した問題に適したモデルを新たに構築し, 学習する手法なども提案されている. これらのモデルの入力画像サイズは 200 ピクセルから 300 ピクセル四方であるため, より多くのパッチを取得するためには, 元画像が高解像度である必要がある. 以下に示す手法は, 高解像度な画像から膨大なパッチを算出することで, CNN の効果的な学習を可能とした.

文献 [27]: CNN を用いたひびわれ検出

文献 [27] では, 構造物において最も発生頻度の高い変状であるひびわれを自動検出することを目的としている. 従来の画像処理技術では, 自らひびわれを表す特徴を設計する必要があったため, 画像を撮影する環境に大きく依存し, 実際のアプリケーションへの導入は困難であった. そこで, 本手法では, $4,928 \times 3,264$ ピクセルの高解像度画像から 256×256 のパッチを $40,000$ 枚用意し, 文献 [8] に基づき著者らが新たに構築した CNN を学習することで 98% の精度でひびわれ検出を可能にしたと報告されている.

文献 [9]: CNN を用いたひびわれ検出

文献 [9] では, 文献 [27] と同様に, パッチ分割された画像を用いて CNN を構築し, 手動による特徴設計が必要な従来手法の精度を上回ることを目的としている. 本手法では, $3,264 \times 2,448$ ピクセルの画像から 99×99 ピクセルのパッチ $640,000$ 枚を作成し, CNN を学習する. 実験では, 従来手法よりひびわれ検出で用いられてきた SVM や Boosting [28] の精度が約 75%であるのに対し, 著者らの CNN を用いた手法では約 90%の精度を達成したと報告されている.

文献 [10]: CNN とナイーブベイズを用いた動画像からのひびわれ検出

文献 [10] では, 原子力発電所を点検しているロボットが撮影した動画像から, 金属に発生するひびわれを自動検出することを目的としている. 本手法では, 30fps で撮影された 720×540 ピクセルの画像からひびわれを含むパッチ $147,344$ 枚, ひびわれを含まないパッチ $149,460$ 枚を取得し, 新たに CNN モデルを構築することで動画像の各フレームの各パッチから検出結果を取得する. 本手法では, 静止画ではなく動画像を用いることから, 各フレームからパッチの算出が可能となるため, 比較的低解像度画像ではあるものの, データ数の増加が可能となる. 最後に, 得られた結果をナイーブベイズに基づき統合することで, 動画像の外乱等の影響を抑制することが可能になり, 最終的に約 98%の検出精度を達成したと報告されている.

文献 [11]: CNN とマルチタスク学習に基づく鉄道路線点検のための物体検出

文献 [11] では, 鉄道路線の点検時に得られる画像から結び目と留め具などの複数の物体を自動検出することを目的としている. 具体的には, 撮影された $4,096 \times 320$ ピクセルの画像中から $2,000,000$ 枚のパッチを算出し, 素材ごとに 10 種類のクラス分類を行う. さらに, その素材の一つである留め具においては, 劣化レベルの分類も行っている. 本手法では, CNN の中間層から素材を分類するネットワークと留め具の劣化レベルを分類するネットワークに 2 つに分岐することで, 同時に複数のタスクの分類問題に対処してお

り, これが著者らの新規点である. 実験により, いずれの分類精度においても 95%以上の精度を達成したと報告されている.

2.3　Data augmentation を導入した従来研究

CNN の学習に必要な画像数を増やすことで, 学習画像のバリエーションを増やし, 学習を安定させて過学習を防ぐことを目的として data augmentation に関する研究が行われている [29]. 具体的に, data augmentation とは, 画像にランダムに回転, 水平に反転, コントラストの調整などの処理を加えることで, 元画像とは異なるデータを作成し, 学習データ数を補填する手法である. 構造物の維持管理の分野においても, 学習画像が少ない場合においては, data augmentation を導入することで, 深層学習の適用を可能にしている [12, 13, 14].

文献 [14]: CNN を用いた構造物に発生する変状の検出

文献 [14] では, 現実の状況を考慮し, 複雑な背景等が撮像されている画像から, 複数種類の変状を自動検出することを目的としている. 具体的に, 本手法では, 学習画像の一部をランダムに選択し, 回転, オクルージョン, モーションブラーリングの 3 種類の変換を施す. 回転は対象画像の上下左右を反転する処理であり, オクルージョンは対象画像の一部の領域を黒塗りすることで隠す処理である. これらの処理は data augmentation では一般的な処理である. 実際の環境では, 画像を撮影するカメラが動くことによって, 画像がぼけてしまう可能性もあるため, 本手法では, モーションブラーリングによって, 予めぼけた画像を作成し, 画像のぼけに頑健な CNN の学習を可能にする. これにより, 元画像数の 2 倍の画像を用いて CNN を学習することで, 物体の自動検出においてスタンダードな手法である Faster R-CNN [30] と比較し, 大幅な精度向上を報告している.

文献 [13]: Faster R-CNN を用いたリアルタイム処理による変状領域検出

文献 [13] では, 領域抽出手法として優れた性能を誇る Faster R-CNN を用い

て, 複数種類の変状の種類およびその領域をリアルタイムで自動検出することを目的としている. 本手法では, 元の学習画像に対し, 反転の処理を施すことで, data augmentation を行い, その画像をさらにパッチ分割することで学習データの充足を行っている. 実験により, 5 つの変状すべてに対し, 約 90%以上の精度で検出可能であることが確認された. さらに, 動画像に対してもテストを行うことで, リアルタイム処理が可能であることも報告されている.

2.4　学習済み CNN を用いた転移学習に関する従来研究

上述の手法の多くは, CNN モデルを著者らが新たに提案し, 大量のパッチ画像を用いて学習した. 対象のドメインに応じた画像を用いて CNN を学習することで, 弁別性の高い特徴の算出が可能となり, 人手により設計した特徴を用いた従来研究と比較して, 種々のタスクの精度向上を実現した. ここで, 近年では, 一般物体認識おいて高い認識精度を誇る CNN モデルが種々提案されている [31, 32, 24]. 一般物体認識用の大規模データセットである ImageNet を用いて学習された種々の CNN のモデルパラメータが公開されており, このパラメータを初期値として用い, 自らのデータセットで再学習を行う Fine-tuning が盛んにおこなわれている [16]. Fine-tuning では, 予め学習されたモデルを用い, 最終層に近い層のパラメータのみを, 対象のデータセットで再学習することから, 大量の学習画像がない場合において, 有効な分類手法の一つである. また, 予め学習された CNN モデルを特徴抽出器とみなした転移学習も存在する. 具体的に, ImageNet で学習された CNN モデルに対象の画像を入力することで特徴を抽出し, その後, 従来の SVM 等の分類器を学習する方法である. これらの手法は, 対象のドメインが学習で用いられた一般画像とは異なる場合でも, 対象画像を表現する能力 (以降, 表現能力) が比較的担保されていることに注目して構築されている. また, 特徴抽出以降の処理で, データに適した処理を加える, つまり, 対象画像から得られた特徴のみを用いて分類器を学習することで, 種々の精度の向上を実現する. CNN を特徴抽出器として用いる手法は, 特徴抽出のために CNN モデルを学習することなく, 比較的表現

能力の高い特徴を算出可能であることから, 構造物の維持管理の分野においても広く用いられている [15, 16, 17].

文献 [16]: 構造物を撮影した画像を用いた転移学習による変状分類

文献 [16] では, 構造物の画像を用いた種々のタスクの高精度化のために, Structured ImageNet と呼ばれるデータセットが提案され, それらを用いた転移学習により, 構造物を構成する成分の分類, 剥落の度合いの分類, 変状の劣化レベルの分類および変状の種類の分類の 4 つの画像認識タスクの高精度化を目的としている. 本手法では, ImageNet で学習された VGG16 [24] に対し, Structured ImageNet で Fine-tuning を行う. Fine-tuning による結果と, Fine-tuning 後の VGG16 を特徴抽出器として用いた場合の両者において, 分類精度が高いことから, 転移学習の有効性が認められたと報告されている.

文献 [15]: 無人航空機で撮影された画像を用いた転移学習によるひびわれ検出

文献 [15] では, 無人航空機によって撮影された構造物の画像を用いて, ひびわれを含む画像か否かを分類することを目的としている. 本手法では, 大規模データセットである ImageNet で予め学習された VGG16 に無人航空機で撮影された画像を入力し, 得られた特徴を用いて SVM, 3 層のニューラルネットワークおよびロジスティック回帰を学習することでひびわれ画像の検出を実現している. 分類器としてロジスティック回帰またはニューラルネットワークを用いた際の分類精度が最も高く, 前処理や data augmentation なしで, 90%の精度を達成したことが報告されている.

文献 [17]: 転移学習に基づき舗装道路の画像からひびわれを検出

文献 [17] では, 加熱混合アスファルトの表面画像とポルトランドセメントコンクリートの表面画像からひびわれを自動検出することが目的である. 本手法では, これらの画像を ImageNet で学習済みの VGG16 に入力することで, その中間層から画像特徴を算出する. 得られた特徴を用いて, SVM

や Random Forest [33] を学習することで精度比較する. 実験により, 3 層のニューラルネットワークを学習したときの精度が最も高いことが報告されている.

2.5　本論文で解決すべき課題と解決法

本節では, 構造物の効率的な維持管理支援の実現に向けて, 高精度な変状分類手法を構築するために, 本論文で解決すべき課題について明確にする. 上述した従来研究では, 変状画像が一般物体認識用の画像と比較して数が少ないという問題を解決するため, 様々なアプローチが提案されている. 2.2 節で述べた手法では, 画像数を増加させるために, 対象の変状画像をパッチに分割し, それらを入力画像として用いることで CNN の学習を行った. 画像から得られる大量のパッチを用いることで, 変状画像に含まれるひびわれを高精度に検出可能とした. ひびわれは, 基本的に構造物中のコンクリートに発生することから, これらの手法で用いられた変状画像はいずれも均一な背景を含む画像であった. しかしながら, 構造物に発生する変状にはひびわれ以外にも様々な種類が存在し, その発生部位や部材は変状ごとに異なることから, 一般的に撮影される変状画像には多様な背景が撮像される. したがって, パッチ分割した際, 多様なバリエーションを持つ画像が選択されるため, 変状分類において有効な特徴を算出することが困難となる. 次に, 2.3 節で述べた従来手法では, 元の画像に対して, 画像の回転や反転, オクルージョンなどの処理を施すことで, 画像を生成し, 学習画像の不足を補う手法であった. 文献 [14] では, 上記の生成方法に加えて, カメラのブレを考慮したモーションブラーリングを導入し, 実際に撮影される環境を想定した画像生成を行うことで, CNN の精度向上を実現している. しかしながら, これらの data augmentation の手法においても, 多様な背景の影響を軽減する手法は提案されておらず, 変状分類において本アプローチによる精度向上には限界が存在する. 以上の 2.2 節および 2.3 節では, 学習画像の不足という, 構造物の維持管理の分野における問題点を考慮した手法が提案されている. しかしながら, 変状画像は実際の点検時に技術者が手動で撮影することから,　多様な背景等の不要な領域が多く含まれることか

ら, 変状を表現する特徴の算出が困難となる. しかしながら, 従来研究では, これらの影響を考慮した手法の構築には至っていない.

続いて, 2.4 節で述べた従来手法では, 少量の画像でも学習可能な転移学習に基づく従来手法を提案している. 変状画像を表現可能な特徴を算出するためには, 深層学習中の隠れ層のパラメータを最適に決定する必要がある. しかしながら, 効果的な特徴変換を行うためには, 多層の隠れ層を必要とすることから, 最適に決定すべきパラメータ数が増大する. これにより, 学習画像が少量である変状画像では, このアプローチによる問題の解決は困難となる. そこで, これらの従来研究では, 大規模な一般物体認識用の画像で学習された CNN のパラメータを初期値とし, 変状画像を用いて深層学習のパラメータを微調整する手法と, 変状画像を入力した深層学習の隠れ層の出力を用いて少量の画像で学習可能な分類器を構築する手法の二つのアプローチがとられている. いずれのアプローチにおいても, 人手による特徴設計と比較して, 変状を表現可能な特徴の算出が実現しているものの, 変状画像は同一の変状においても多様な視覚的特徴を有するため, その分類精度向上に限界が存在する. 以上の問題点より, 少量の学習画像においては, 画像のみを用いて変状を表現可能な特徴の算出は困難であると考えられる.

そこで, 本論文では, 実際の維持管理業務において, 技術者は変状画像の撮影に加えて, 構造物の部位や部材をテキストデータとして記録し, 変状画像とテキストデータを総合的に評価することで最終的な判断を下す点に着目する. これらの異なる種類のデータを用いることで, 2.2 節, 2.3 節および 2.4 節で述べた問題点を解決することが可能であると期待できる. 具体的に, 変状画像には, 変状の発生している構造物の部位が背景として撮影されるものの, 部位や部材の情報を用いることで, 背景の影響を考慮した変状分類手法の構築が可能になると考えられる. また, CNN では, 大量の入力画像と出力の組を用いて, 表現能力の高い特徴を自動で抽出するために, 多数の隠れ層による特徴変換が必要であった. 一方で, 実際の技術者の判断に用いられるテキストデータと変状画像を用いることで, 表現能力の高い特徴への変換を行う射影を直接求めることが可能になると考えられる. これにより, 最適に決定すべきパラメータ数の削減も実現するため, 少量の学習画像を

用いた高精度な変状分類が可能となることが期待できる. したがって, 本論文では, 高精度な変状分類を実現するために, 少量の学習画像から高い弁別性を有する特徴へ変換可能とする深層学習手法を構築する. まず, 3 章で, 実際の維持管理業務のプロセスを模擬した変状分類手法を構築することで, 変状画像とテキストデータおよび変状との間に関連性があることを明らかにする. 次に, 4 章で, これらのデータから得られる異種特徴間の関連性に基づき, 分類に有効な特徴への変換を可能にする射影を導出する. さらに, 5 章では, 4 章で導出した射影を用いて, 順伝搬型の深層学習を構築することで, 高精度な変状分類を実現する手法を提案する. 実験により, 少量の学習画像を用いて, 弁別性の高い特徴への変換が可能となることを確認する.

2.6　まとめ

本章では, 構造物の維持管理支援を目的とした, 機械学習を用いた画像解析に関する従来研究について説明した. また, 従来手法における問題点を述べ, 高精度な変状分類を実現するために, 本論文で解決すべき課題とその解決法について明確にした.

第3章　異種特徴間の関係を表現可能なベイジアンネットワークに基づく変状分類

3.1　はじめに

本章では, 異種特徴と変状を変数とみなすことで, それらの因果関係を確率値として表現可能なベイジアンネットワーク [34] に基づく変状分類手法を構築する. 本手法では, 実際の維持管理業務において, 変状画像とテキストデータが技術者の判断に有効であることに着想を得て, 維持管理業務を模擬した枠組みにより変状分類を行う. 具体的に, テキスト特徴と変状を変数としたベイジアンネットワーク (以降, Text-based BN) と画像特徴・テキスト特徴・変状全てを変数としたベイジアンネットワーク (以降, Multi-feature BN) を構築する. 実際の点検業務では, 図 3.1 に示す通り, 技術者はテキストデータからの情報を優先的にみる傾向があることから, テキストデータのみで分類可能かどうかを判定する指標 (以降, 信頼度) を導入する. これにより, Text-based BN と Multi-feature BN の適応的なモデル選択が可能となる. 分類対象の画像ごとに最適なモデルが選択されるため, 各特徴がどの変状を分類する際に有効であるかを考慮することが可能になる. 具体的に, 本手法の概要図を図 3.2 に示す.

以降, 3.2 節で, Text-based BN の構築と適応的なモデル選択を行うための信頼度の算出を行い, 3.3 節で, Multi-feature BN の構築を行う. つづいて, 3.4 節で, 適応的なベイジアンネットワークモデルの選択方法について説明し, 3.5 節で, 実験を行い, 最後に 3.6 節で, まとめとする.

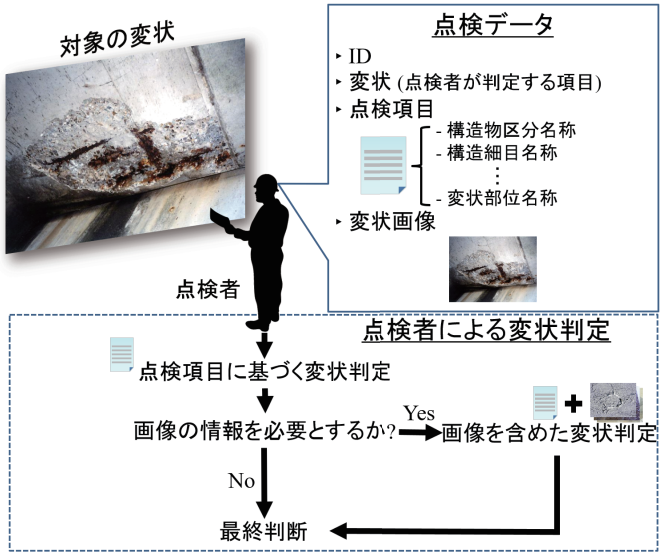


図 3.1: 実際の維持管理業務における点検の流れ.

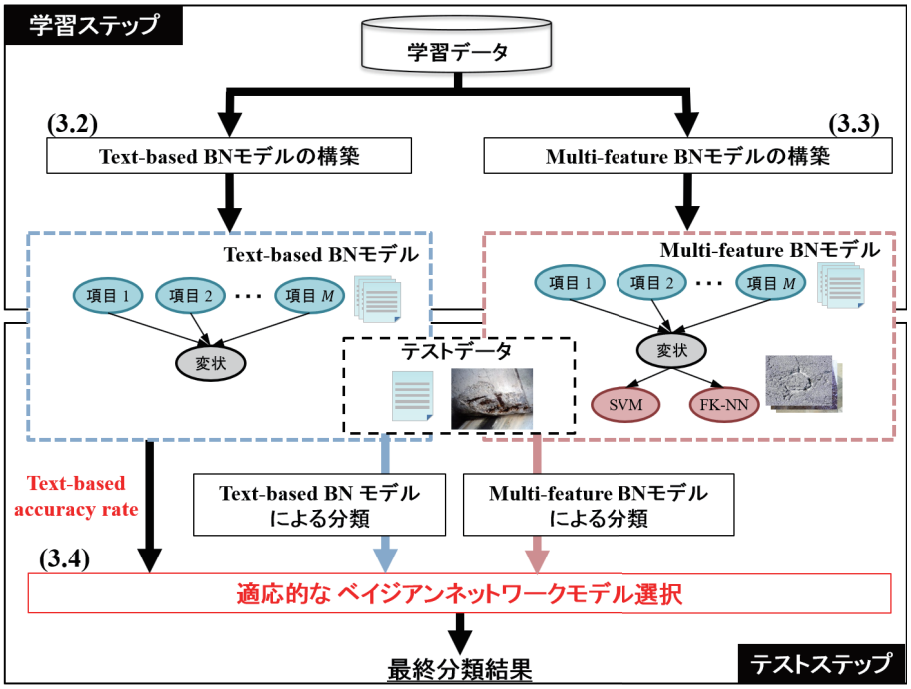


図 3.2: 提案手法全体の概要図.

3.2　Text-based BNの構築と信頼度の算出

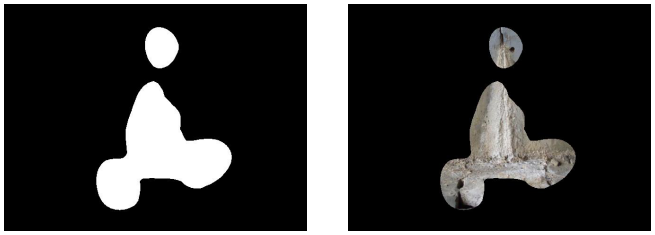
本節では, Text-based BN の構築方法と, Text-based BN に基づく信頼度の算出方法について説明する. まず, 学習データセット $\{X_i,\ T_i^1,\ T_i^2,\ ...,\ T_i^M\}_{i=1\sim N}$ (M は点検項目の数であり $(M\ =\ 6)$, N は学習データ数) が与えられた時, 条件付き確率表 (Conditional Probability Table; CPT) に基づき, x の条件付き確率 $P(x \mid t^1,\ t^2,\ ...,\ t^M)$ を取得する. ただし, X_i は変状の種類を表し, $T_i^1,\ T_i^2,\ ...,\ T_i^M$ はそれぞれ点検項目を表す. 具体的に, 投票処理を行うことで CPT を構築する. 一般的な BN では, 各変数が取りうる組み合わせの数が膨大となるため, 投票処理によって CPT を構築することは困難である. しかしながら, 道路構造物のデータの場合, 起こりうる組み合わせが少ないため, 投票処理による CPT の構築が可能となる. 次に, テストデータ $\{T^1,\ T^2,\ ...,\ T^M\}$ が与えられた際, 未知の変状 X の条件付き確率 $P(x = X \mid t^1 = T^1,\ t^2 = T^2,\ ...,\ t^M = T^M)$ が, 先に得られた CPT によって算出される. 続いて, 本手法の新規点である適応的モデル選択を行うために, Text-based BN に基づく信頼度を記述した表 (以降, Text-based accuracy rate table; TAT) を算出する. 具体的に, 学習データを用いて, Text-based BN に leave-one-out 法を適用することで, 算出する. これにより, テキストデータの組み合わせ $\{t^1,\ t^2,\ ...,\ t^M\}$ ごとに, 信頼度 $A(t^1,\ t^2,\ ...,\ t^M)$ が得られる. 以上により, テストデータが入力された際, そのテキストデータの組み合わせ $\{T^1,\ T^2,\ ...,\ T^M\}$ から, TAT を参照することで, $A(t^1 = T^1,\ t^2 = T^2,\ ...,\ t^M = T^M)$ が算出される.

3.3　Multi-feature BNの構築

本節では, Multi-feature BN の構築方法について説明する. 本手法では, まず, テキストデータからテキスト特徴を算出する. 次に, 図 3.3 に示すように, 変状画像から顕著性マップ [35] に基づき変状の領域を抽出し, 得られた領域から表 3.1 に示す画像特徴を算出する. その後, これらの特徴を一列に並べて得られる特徴を入力として, Support Vector Machine (SVM) [23] と Fuzzy K-Nearest Neighbor (FK-NN) [36] を構築する. その後得られた結果とテキストデータを入力として



(a) 元画像　(b) 重み付き顕著性マップ



(c) 注目された領域　(d) 抽出された変状の領域

図 3.3: 顕著性マップに基づく変状領域の抽出の例.

表 3.1: 本手法で用いる特徴とその次元数.

特徴	次元数
Gabor Wavelet に基づく特徴 [37]	48
HS ヒストグラムに基づく特徴 [38]	128
SIFT-BOF 特徴 [39]	2,000
テキスト特徴	141

Multi-feature BN を構築する.

学習データ i $(= 1, 2, ..., N; N$ は学習データ数 $)$ が与えられた際, まず, 特徴ベクトルを算出する. 点検データ (表 3.2) 中の “構造種別名称”, “構造細目名称” などの各点検項目 $s(= 1, 2, ..., M)$ が取りうる点検結果の種類を次元数 D_s とし, $\Sigma_{s=1}^{M} D_s$ 次元のバイナリベクトルを得る. 次に, 各変状の画像から, 表 3.1 に示す 2,176 次元の特徴を算出する. 得られた 2 つのベクトルを一列に並べることで, $D(= \Sigma_{s=1}^{M} D_s + 2,176)$ 次元の特徴ベクトル $\boldsymbol{v}_i \in \mathbb{R}^D$ を算出する.

算出された特徴ベクトルに対して, 各変状に対応したラベル X を付与し, minimum-

Redundancy-Maximum-Relevance (mRMR) アルゴリズム [40] による特徴選択を行い, 画像特徴の中で“変状の名称”との関連性の高い要素を並べた特徴ベクトル $\boldsymbol{\gamma}_i \in \mathbb{R}^d$ を新たに取得する. 具体的に, mRMR アルゴリズムでは, まず, d 個の特徴の集合 S_d としたとき, 特徴間の冗長性 $R(S_d)$ を下式により求める.

$$R(S_d) = \min_{S_d \subset S} \frac{1}{|S_d|^2} \sum_{\boldsymbol{v}(j), \boldsymbol{v}(l) \in S_d} I(\boldsymbol{v}(j); \boldsymbol{v}(l)) \tag{3.1}$$

ただし, $|S_d|$ は特徴集合 S_d の要素数である. S は全ての特徴の集合を表す. また, $I(\boldsymbol{v}(j); \boldsymbol{v}(l))$ は $\boldsymbol{v}(j)$ と $\boldsymbol{v}(l)$ の相互情報量を表す. 次に, 特徴とクラスラベルとの関連性 $L(S_d, X)$ を下式により求める.

$$L(S_d, X) = \max_{S_d \subset S} \frac{1}{|S_d|} \sum_{\boldsymbol{v}(j) \in S_d} I(\boldsymbol{v}(j); X) \tag{3.2}$$

ただし, $I(\boldsymbol{v}(j); X)$ は特徴 $\boldsymbol{v}(j)$ とクラスラベル X の関連性を表す.

以上により得られた冗長性と関連性を用いて下式を解くことで最終的な特徴ベクトル $\boldsymbol{\gamma}_i \in \mathbb{R}^d$ を取得する.

$$\max_{S_d \subset S} \Big\{ L(S_d, X) - R(S_d) \Big\} \tag{3.3}$$

最後に, 得られた特徴ベクトルを用いて SVM [23] および FK-NN [41] の構築を行う. N 枚の学習データから N 本の特徴ベクトル $\boldsymbol{\gamma}_i \in \mathbb{R}^d$ $(i = 1, 2, ..., N)$ とそれらに対応するクラス $y_i \in \{-1, 1\}$ が与えられたとき, $\psi(\boldsymbol{\gamma}_i) \in \mathbb{R}^{d'}$ $(d' \gg d)$ を得るために, 特徴ベクトル $\boldsymbol{\gamma}_i$ は高次元な特徴空間へ射影される. 重み係数 $\boldsymbol{w} \in \mathbb{R}^{d'}$ と

表 3.2: 点検データの例.

		点検項目			
ID	変状	構造種別名称	構造細目名称	···	変状部位名称
1	ひびわれ	地覆	高欄, 地覆	···	左・外・下
2	本体の損傷	鋼製フィンガージョイント	伸縮装置	···	その他
3	遊離石灰	RC 床板	コンクリート上部工	···	主版部
4	剥落	PC 桁	コンクリート上部工	···	Crossbeam
5	腐食	鋼桁	鋼構造物	···	主桁フランジ
6	漏水	RC 床板	コンクリート上部工	···	張出し部 (左)
7	鉄筋露出	PC 桁	コンクリート上部工	···	主桁フランジ
8	浮き	RC 床板	コンクリート上部工	···	主版部
⋮	⋮	⋮	⋮	⋮	⋮

バイアス $e \in \mathbb{R}^1$ を用い, 最適な超平面を定義することで, 二つの異なるクラスを分離することが可能となる. 分離するための式は, 以下のとおりである.

$$\hat{y} = \boldsymbol{w}^\top \psi(\boldsymbol{\gamma}) + e \tag{3.4}$$

ただし, $\boldsymbol{\gamma}$ はクラスが与えられていないテストサンプルの特徴ベクトルを表しており, そのクラスは $\hat{y}$ によって決定される. もし, $\hat{y}$ が正例であれば, $\boldsymbol{\gamma}$ はラベル 1 の正例クラスに属することになる. それ以外の場合, ラベルが-1 となり, 負例のクラスに属することになる. SVM の目的は, マージン最大化と推定誤差の最小化の二つの指標を考慮した目的関数を解くことで, 最適な超平面を求めることにある. 式 (3.4) の最終結果は, 下式により得られる.

$$\hat{y} = sign\Big(\sum_{i=1}^{N} c_i y_i \kappa(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}) + e\Big) \tag{3.5}$$

ただし, $\kappa(\cdot,\cdot)$ カーネル関数であり, d' 次元の特徴空間における内積により求められる. さらに, c_i は以下の凸二次計画問題を解くことで得られるラグランジュ乗

数である.

$$\text{Maximize} \sum_{i=1}^{N} c_i - \frac{1}{2} \sum_{i,\epsilon=1}^{N} c_i c_\epsilon y_i y_\epsilon \kappa(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_\epsilon) \tag{3.6}$$

$$\text{s.t.} \sum_{i=1}^{N} c_i y_i = 0, \;\; 0 \leq c_i \leq C \;\; (i = 1, 2, ..., N) \tag{3.7}$$

ただし, C は学習データにおいて, 誤分類のペナルティを調整するパラメータである. 式の導出等は, 文献 [23] に詳しい. 上記は, 二クラス分類問題に対応しており, 本研究では, 複数クラスに対応するために, one-vs-all approach [42] を採用している.

さらに, 分類結果 $f_{\text{fknn}}(\boldsymbol{\gamma})$ を算出するために, 我々は, FK-NN アルゴリズムを使う [36]. FK-NN は, ラベルにおけるファジィ理論が K-NN 分類器 [41] に組み込まれた手法である. 具体的に, ある対象のサンプルに対し, K 近傍のサンプルが各クラスに属する割合を算出し, これらを考慮して, メンバシップ値を割り当てる. ここで我々は, 学習データにおける特徴ベクトル $\boldsymbol{\gamma}_i$ とクラス X_i を用いる. FK-NN アルゴリズムを用いた分類では, 二つのステップが存在する. 一つ目のステップでは, 学習データセットにおける全サンプルの K 近傍を決定する. そして, クラス X_i に対する, 特徴ベクトル $\boldsymbol{\gamma}_i$ のファジーメンバーシップ値は以下の式によって推定される.

$$u_x(\boldsymbol{\gamma}_i) = \begin{cases} 0.51 + (n_i^x/K) \times 0.49 & \text{if } x = X_i \\ (n_i^x/K) \times 0.49 & \text{otherwise} \end{cases} \tag{3.8}$$

ただし, n_i^x は, $\boldsymbol{\gamma}_i$ の K 近傍内におけるクラス x に属するサンプルの数を表している. さらに, 0.51 と 0.49 は重みパラメータであり, 文献 [36] を踏襲して設定している. $u_x(\boldsymbol{\gamma}_i)$ は下式を満たす.

$$\sum_{x=1}^{\Phi} u_x(\boldsymbol{\gamma}_i) = 1 \tag{3.9}$$

ただし, $u_x(\boldsymbol{\gamma}_i) \in [0,1]$ であり, Φ は分類対象のクラス数を表す.

二つ目のステップでは, テストサンプル $\boldsymbol{\gamma}$ に対して, ファジーメンバーシップ値 $\hat{u}_x(\boldsymbol{\gamma})$ を下式により割り当てる.

$$\hat{u}_x(\boldsymbol{\gamma}) = \frac{\sum_{k=1}^{K} u_x(\boldsymbol{\gamma}^k)(1/\|\boldsymbol{\gamma} - \boldsymbol{\gamma}^k\|^2)}{\sum_{k'=1}^{K} (1/\|\boldsymbol{\gamma} - \boldsymbol{\gamma}^{k'}\|^2)} \tag{3.10}$$

ただし, $\boldsymbol{\gamma}^k$ $(k = 1,\ 2,\ ...,\ K)$ はテストデータ $\boldsymbol{\gamma}$ の k 近傍に存在する学習データの特徴ベクトルである. さらに, $\|\boldsymbol{\gamma} - \boldsymbol{\gamma}^k\|^2$ は, 特徴ベクトル $\boldsymbol{\gamma}$ と k 番目の近傍の特徴ベクトル $\boldsymbol{\gamma}^k$ との間の距離を表しており, ユークリッド距離に基づいて算出される. 結果的に, テストデータにおけるクラスラベル $f_{\mathrm{fknn}}(\boldsymbol{\gamma})$ が下式に従い, 算出される.

$$f_{\mathrm{fknn}}(\boldsymbol{\gamma}) = \arg \max_{x}\ \hat{u}_x(\boldsymbol{\gamma}) \tag{3.11}$$

以上より, 提案手法は二種類の分類結果 $f_{\mathrm{svm}}(\boldsymbol{\gamma})$ と $f_{\mathrm{fknn}}(\boldsymbol{\gamma})$ を算出する.

最後に, 各分類器の結果を BN により統合する. 統合の際は, 画像特徴だけでなく, テキストデータも用いる.

学習データ $\{X_i,\ T_i^1,\ T_i^2,\ ...,\ T_i^M,\ f_{\mathrm{svm}}(\boldsymbol{\gamma}_i),\ f_{\mathrm{fknn}}(\boldsymbol{\gamma}_i)\}_{i=1\sim N}$ が与えられた際, 提案手法では, 変状 x についての条件付き確率 $P(x \mid t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})$ を下式により求める.

$$\begin{aligned} &P(x \mid t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}}) \\ &\quad = \frac{P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})}{P(t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})} \\ &\quad = \frac{P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})}{\sum_x P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})} \end{aligned} \tag{3.12}$$

ただし, $\sum_x P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\mathrm{svm}},\ x_{\mathrm{fknn}})$ は x に関して周辺化した確率である.

さらに, 結合確率 $P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\text{svm}},\ x_{\text{fknn}})$ は以下のように式変形する.

$$\begin{aligned} &P(x, t^1,\ t^2,\ ...,\ t^M,\ x_{\text{svm}},\ x_{\text{fknn}}) \\ &\quad = P(\ x_{\text{svm}} \mid x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\text{fknn}})\ P(x,\ t^1,\ t^2,\ ...,\ t^M,\ x_{\text{fknn}}) \\ &\quad = P(\ x_{\text{svm}} \mid x)\ P(\ x_{\text{fknn}} \mid x,\ t^1,\ t^2,\ ...,\ t^M)\ P(x,\ t^1,\ t^2,\ ...,\ t^M) \\ &\quad = P(\ x_{\text{svm}} \mid x)\ P(\ x_{\text{fknn}} \mid x)\ P(x \mid t^1,\ t^2,\ ...,\ t^M)\ P(t^1,\ t^2,\ ...,\ t^M) \end{aligned} \quad (3.13)$$

ただし, 上記の式変形においては, SVM と FK-NN との条件付き独立を仮定している. これは, 近年種々提案されている分類器の統合手法 [43, 44, 45, 46] においても同様に仮定していることから妥当な仮定である. さらに, $P(x \mid t^1,\ t^2,\ ...,\ t^M)$, $P(\ x_{\text{svm}} \mid x)$ と $P(\ x_{\text{knn}} \mid x)$ は条件付き確率表により求められる.

以上より, 式 (3.12) は下式と等価となる.

$$\begin{aligned} &P(x \mid t^1,\ t^2,\ ...,\ t^M,\ x_{\text{svm}},\ x_{\text{fknn}}) \\ &\quad = \frac{P(\ x_{\text{svm}} \mid x)\ P(\ x_{\text{fknn}} \mid x)\ P(x \mid t^1,\ t^2,\ ...,\ t^M)}{\sum_x P(\ x_{\text{svm}} \mid x)\ P(\ x_{\text{fknn}} \mid x)\ P(x \mid t^1,\ t^2,\ ...,\ t^M)} \end{aligned} \quad (3.14)$$

テストデータ $\{T^1,\ T^2,\ ...,\ T^M,\ f_{\text{svm}}(\boldsymbol{\gamma}),\ f_{\text{fknn}}(\boldsymbol{\gamma})\}$ が与えられた際, 未知の変状 X を条件付き確率 $P(x = X | t^1 = T^1, t^2 = T^2, ..., t^M = T^M, x_{\text{svm}} = f_{\text{svm}}(\boldsymbol{\gamma}),\ x_{\text{fknn}} = f_{\text{fknn}}(\boldsymbol{\gamma}))$ に基づき算出する. 以上より, $P(x = X | t^1 = T^1, t^2 = T^2, ..., t^M = T^M, x_{\text{svm}} = f_{\text{svm}}(\boldsymbol{\gamma}),\ x_{\text{fknn}} = f_{\text{fknn}}(\boldsymbol{\gamma}))$ が最大となる X が結果として出力される. 上記処理により, 各識別器の結果をベイジアンネットワークに導入することで, Multi-feature BN の構築が可能となる.

3.4　適応的なベイジアンネットワークモデルの選択

本節では, Text-based BN の信頼度に基づき, Text-based BN と Multi-feature BN を適応的に選択する方法について説明する. 画像から得られる特徴 $\boldsymbol{\gamma}$ とテキストデータ $\{T^1, T^2, ..., T^M\}$ を有するテストデータ $\mathcal{D} = \{\boldsymbol{\gamma}, T^1, T^2, ..., T^M\}$ が与えら

れた時, 3.2 節で算出した TAT に基づき, そのテストデータのテキストデータの組における信頼度 $A(t^1 = T^1, t^2 = T^2, ..., t^M = T^M)$ が算出される.

さらに, 算出された $A(t^1 = T^1, t^2 = T^2, ..., t^M = T^M)$ と閾値 λ を比較することで, テストデータを入力するベイジアンネットワークを決定する. 図 3.4 に概要を示す. 具体的に, 信頼度 $A(t^1 = T^1, t^2 = T^2, ..., t^M = T^M)$ が閾値 λ を下回る場合を考える. これは, Text-based BN では正確な分類が困難であるということを表すため, このような場合は, Multi-feature BN へテストデータを入力する. したがって, 本手法は, テキストデータの組み合わせに基づき, 最適なベイジアンネットワークを選択して最終分類結果を取得する. 結果的に, 出力結果 X^* は下式に示す条件付き確率 $P(x|\mathcal{D})$ に基づき算出される.

$$X^* = \arg \max_x P(x|\mathcal{D}) \tag{3.15}$$

$$P(x|\mathcal{D}) = \begin{cases} P(x|t^1 = T^1, t^2 = T^2, ..., t^M = T^M, x_{\text{svm}} = f_{\text{svm}}(\boldsymbol{\gamma}), x_{\text{fknn}} = f_{\text{fknn}}(\boldsymbol{\gamma})) \\ \qquad\qquad \text{if } A(t^1 = T^1, t^2 = T^2, ..., t^M = T^M) < \lambda \\ P(x|t^1 = T^1, t^2 = T^2, ..., t^M = T^M) \qquad \text{otherwise} \end{cases}$$

ここで, $P(x|t^1 = T^1, t^2 = T^2, ..., t^M = T^M)$ は 3.2 節で算出される. また, $f_{\text{svm}}(\boldsymbol{\gamma})$ と $f_{\text{fknn}}(\boldsymbol{\gamma})$ は, 3.3 節で構築する SVM と FK-NN の出力結果であり, $P(x|t^1 = T^1, t^2 = T^2, ..., t^M = T^M, x_{\text{svm}} = f_{\text{svm}}(\boldsymbol{\gamma}), x_{\text{fknn}} = f_{\text{fknn}}(\boldsymbol{\gamma}))$ も 3.3 節で求められる Multi-feature BN の結果である.

このように, 本手法では, 適切なベイジアンネットワークモデルの選択を行うことで高精度な変状分類を実現する. さらに, テストステップでは, 学習ステップにおいて予め構築された TAT を用いることで高速な変状分類が可能となる. したがって, 本手法は, 実応用の観点からも効果的な手法となる.

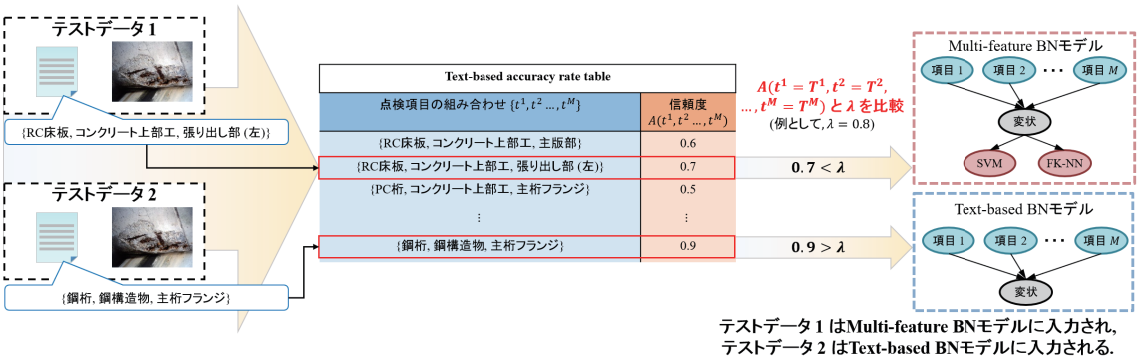


図 3.4: 本手法の有効点である適応的なベイジアンネットワークモデル選択の詳細な処理.

3.5　実験

本章では, 本手法の有効性を確認するための実験を行う. 以降, 3.5.1 で本実験で用いるデータなどの実験準備について説明し, 3.5.2 で従来の機械学習手法との比較を行う.

3.5.1　実験準備

本実験では, 表 3.3 に示す NEXCO 東日本より提供されたデータを用いる. 実際の点検時, 技術者は変状部位を撮影した画像と点検結果を記載したテキストデータを点検データとして保存する. そこで, 以降, テキストデータと変状画像について説明する. テキストデータには, 表 3.2 のように損傷した部位やその構造種別, 変状の程度などが記載されている. このように点検結果を記載したテキストには多数の点検項目が存在する. これらの点検項目における各内容の組み合わせによって点検の結果が記録されている. また, 変状画像は, 図 3.5 のように近景と遠景等, 2 パターン以上が撮影されている. ここで, 本研究では, 変状部分の様子を詳細に把握するため, 近景の画像を用いる. また, 図 3.6 において代表的な変状画像を示す.

腐食は, 塗装やメッキなどによる防食措置が施された普通鋼材では, 集中的に錆

表 3.3: 実験で用いる点検データ数.

変状の種類	$\Phi = 5$	$\Phi = 6$	$\Phi = 7$	$\Phi = 8$
ひびわれ	778	778	778	778
本体の損傷	745	745	745	745
遊離石灰	762	762	762	762
剥落	723	723	723	723
腐食	793	793	793	793
漏水	—	761	761	761
鉄筋露出	—	—	773	773
浮き	—	—	—	757
合計	3801	4562	5335	6092

が発生している状態, または錆が極度に進行し断面減少や腐食を生じている状態を指す. 耐候性鋼材の場合には, 安定錆が形成されず異常な錆が生じている場合や, 極度な錆の進行により断面減少が著しい状態を指す. また, 腐食が発生しやすい個所は漏水の多い桁端部, 水平材上面など滞水しやすい箇所, 支承部周辺, 通気性, 排水性の悪い連結部, 泥, ほこりの堆積しやすい下フランジの上面, 溶接部等である.

遊離石灰は, コンクリートの打継目絵やひびわれ部等から, 石灰分の滲出や漏出が生じている状態を指す. 遊離石灰は, ひびわれ, 浮き, 剥落など他のコンクリートの変状と共に発生する可能性が存在する.

鉄筋露出は, コンクリート部材の表面が剥落し, 剥落部で鉄筋が露出している状態を指す.

剥落は, コンクリート部材の表面が剥落している状態を指す. また, 剥落部で鉄筋が露出しているものについては鉄筋露出とする.

ひびわれは, コンクリート部材の表面にひびわれが生じている状態を指す. ひびわれは, 剥落や鉄筋露出などその他の変状が共に発生している可能性が存在する.

浮きは, コンクリート部材の表面付近が浮いた状態となるものを指す. コンクリート表面に生じる膨らみなどの変状から目視で判断できない場合にも, 打音検

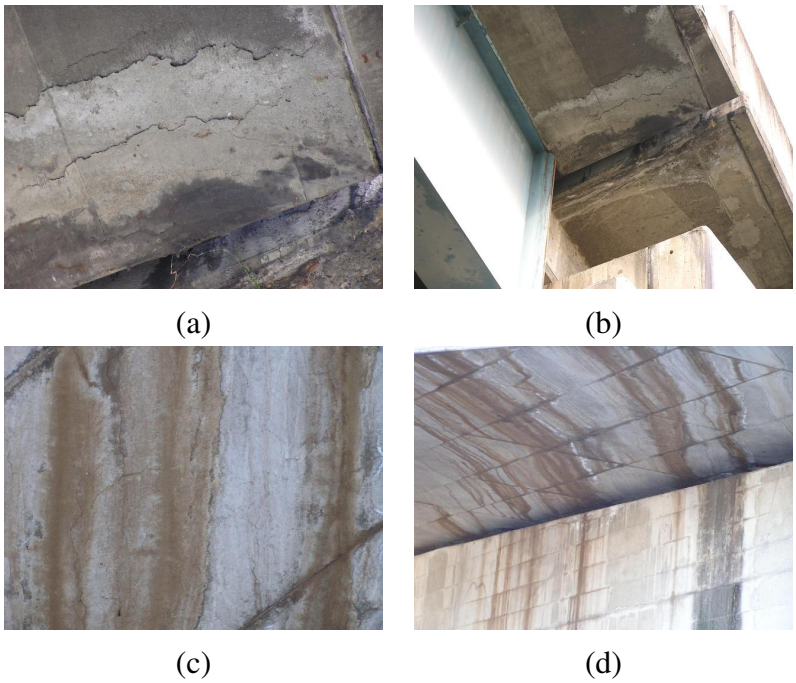


図 3.5: 変状画像の例. (a) および (c) は近景, (b) および (d) は遠景

査のいて濁音を生じることで検出できる場合が存在する. また, 浮いた部分のコンクリートが剥落した場合には, 剥落や鉄筋露出として扱う.

漏水は, コンクリートの打継目絵やひびわれ部等から, 水の滲出や漏出が生じている状態を指す. また, 伸縮装置, 排水施設等から雨水などが本来の排水機構によらず漏出している状態を指す. 漏水は, ひびわれ, 浮き, 剥落など他のコンクリートの変状と共に発生する可能性が存在する.

本体の損傷は, 道路の路肩などに発生することが多い変状である.

以上より, 変状の種類に応じて, 視覚的な特徴が異なっていることから, 変状画像から得られる特徴を用いて変状分類手法を構築することで高精度な変状分類の実現が期待される. また, 分類する変状数 Φ を変更 $(\Phi = 5, 6, 7, 8)$ して実験を行う. また, 閾値 λ は, 学習データセットに leave-one-out 法を適用することで最も精度の高くなった時の値を採用した. FK-NN の近傍数 K も同様の処理により求めた. さらに, mRMR アルゴリズムで選択する次元数 d は, 分類精度が最も高く

表 3.4: 1 つ目の実験で用いる比較手法.

手法	詳細
比較手法 1	SVM [23]
比較手法 2	FK-NN [36]
比較手法 3	Text-based BN
比較手法 4	Multi-feature BN [49]

なった時の値を採用した. また, SVM の正則化パラメータとカーネルパラメータは grid サーチ [47] により決定した.

検証方法は 5 分割交差検定であり, 評価指標は以下に示す Recall, Precision, F 値および正解率を用いた.

$$\text{Recall} = \frac{\text{正しく変状を推定されたサンプル数}}{\text{その変状のサンプル数}} \tag{3.16}$$

$$\text{Precision} = \frac{\text{正しく変状を推定されたサンプル数}}{\text{その変状と推定されたサンプル数}} \tag{3.17}$$

$$\text{F 値} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall + Precision}} \tag{3.18}$$

$$\text{正解率} = \frac{\text{正解数}}{\text{すべてのサンプル数}} \tag{3.19}$$

本実験では, 本手法の有効性の確認のため, 2 種類の実験を行った. 1 つ目の実験では, 表 3.4 に示す 4 つの比較手法を用いた. 比較手法 1, 2 は, Multi-feature BN の構築の際に使われた分類器であり, ベンチマーキング手法である. また, 比較手法 3, 4 はそれぞれ Text-based BN と Multi-feature BN である. 2 つ目の実験では, 表 3.5 に示す統合手法との比較を行う. ただし, 比較手法 B は, 従来より広く用いられてきた統合手法 [43] に基づく変状分類手法である. この手法は, 著者が以前提案した手法 [48] である. 以上の比較実験を行うことで, 本手法の有効性を確認する.

表 3.5: 2 つ目の実験で用いる比較手法.

手法	詳細
比較手法 A	Adaptive weighted fusion [50]
比較手法 B	Supervised learning from multiple experts [43] に基づく手法 [48]

3.5.2　実験結果および考察

まず, 1 つ目の実験結果を表 3.6 と表 3.7 と図 3.7 にそれぞれ示す. それぞれの評価指標の平均に注目すると, 本手法は全ての比較手法を上回っていることが確認できる. 特に, 比較手法 3 と 4 よりも精度が高いことから, Text-based BN と Multi-feature BN を適応的に選択することの有効性が確認された. 次に, 特定の変状に注目して考察を行う. “本体の損傷” に着目すると, 比較手法 1, 2, 4 に勝るものの, 比較手法 3 を下回る結果が得られた. これは, “本体の損傷” が “剥落” や “鉄筋露出” と視覚的に類似していることが原因であると考えられる. 続いて, 実際に本手法を適用することで正解した変状画像の例を図 3.8 に示す. 図 3.8 (a)-(h) は元画像であり, 図 3.8 (a’)-(h’) は領域が抽出された後の画像である. 図 3.8 (b), (c) および (d) では, 空や草などの不要な領域が含まれるものの, 図 3.8 (b’), (c’) および (d’) ではそれらが取り除かれていることが判る. したがって, 顕著性マップ [35] により変状の領域を抽出することの有効性が確認できる.



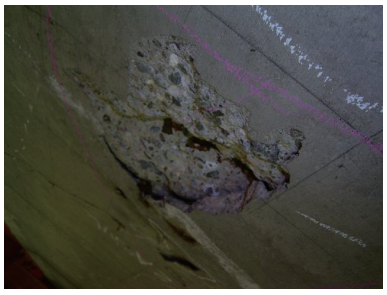

腐食の例




遊離石灰の例




鉄筋露出の例

剥落の例




ひびわれの例



浮きの例

漏水の例

本体の損傷の例

図 3.6: 変状画像の例.

表 3.6: 各手法の Recall, Precision および F 値 ($\Phi = 5$).

	比較手法 1			比較手法 2			比較手法 3			比較手法 4			本手法		
変状	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
ひびわれ	0.543	0.739	0.624	0.818	0.792	0.805	0.813	**0.882**	0.846	0.828	0.802	0.815	**0.905**	0.874	**0.889**
本体の損傷	0.803	0.747	0.765	0.959	0.949	0.954	**1.000**	**1.000**	**1.000**	0.971	0.964	0.967	**1.000**	0.977	0.988
遊離石灰	0.627	**0.809**	0.701	0.695	0.778	0.733	0.779	0.776	0.777	0.721	0.778	0.747	**0.787**	0.807	**0.796**
剥落	0.558	0.732	0.631	0.745	0.707	0.724	0.779	0.722	0.749	0.753	0.726	0.738	**0.799**	**0.828**	**0.813**
腐食	0.736	0.448	0.551	0.962	0.953	0.958	**1.000**	**1.000**	**1.000**	0.967	0.972	0.969	**1.000**	**1.000**	**1.000**
平均	0.653	0.695	0.654	0.836	0.836	0.835	0.874	0.876	0.874	0.849	0.848	0.847	**0.898**	**0.897**	**0.897**

表 3.7: 各手法の Recall, Precision および F 値 ($\Phi = 7$).

	比較手法 1			比較手法 2			比較手法 3			比較手法 4			本手法		
変状	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
ひびわれ	0.740	0.500	0.587	0.760	0.786	0.772	0.778	**0.822**	0.799	0.729	0.793	0.759	**0.859**	0.805	**0.830**
本体の損傷	0.750	0.822	0.778	0.864	0.769	0.814	**0.955**	0.857	**0.904**	0.885	0.798	0.839	0.935	**0.870**	0.901
遊離石灰	0.596	0.628	0.637	0.628	0.660	0.643	0.722	0.610	0.661	0.700	**0.689**	0.693	**0.736**	0.672	**0.701**
剥落	0.496	0.624	0.546	0.624	0.598	0.610	0.598	0.598	0.598	**0.680**	0.613	0.645	0.615	**0.686**	**0.646**
腐食	0.774	0.972	0.702	0.972	0.964	0.968	**1.000**	**1.000**	**1.000**	0.969	0.974	0.971	**1.000**	0.998	0.999
漏水	0.566	0.694	0.659	0.694	0.825	0.754	0.709	**0.905**	0.795	0.733	0.853	0.788	**0.744**	0.879	**0.805**
鉄筋露出	0.561	0.632	0.597	0.632	0.592	0.611	0.640	0.654	0.647	0.618	0.611	0.614	**0.663**	**0.664**	**0.662**
平均	0.641	0.674	0.644	0.739	0.742	0.739	0.771	0.778	0.772	0.759	0.762	0.758	**0.793**	**0.796**	**0.792**

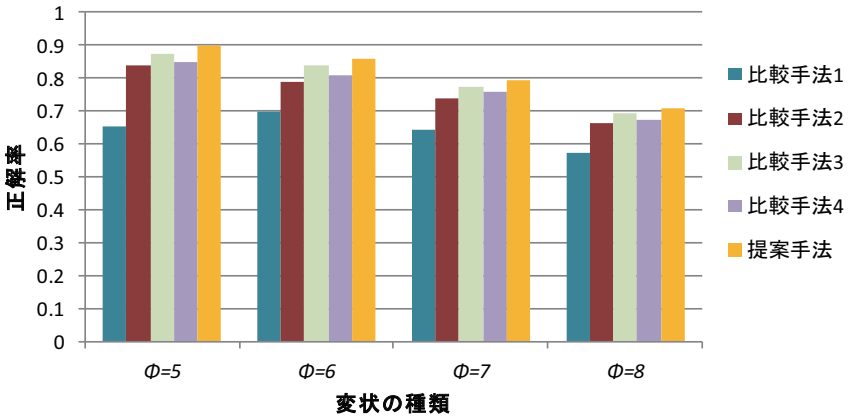


図 3.7: 比較手法 1 から 4 および提案手法の正解率の比較.

次に, 2 つ目の実験結果を表 3.8 と表 3.9 と図 3.9 にそれぞれ示す. 実験結果より, 本手法により多くの変状の分類精度の向上が確認された. したがって, 異種の特徴を重み付けしながら統合する手法 [50] や複数の分類器を統合する手法 [43] よりも高精度となることから, 本手法により高精度な変状分類が実現したと言える. また, 文献 [50] の手法の精度が極めて低いことから, テキストデータから得られる特徴と変状画像から得られる特徴が有する性質は互いに異なるものであり, 本手法ではそれらを協調的に用いることが可能であることも確認できる.

以上より, 本手法の有効性は確認された. さらに, 提案手法は, モデル選択の際, 予め学習データセットで構築した TAT を参照するため, 高速な変状分類が可能となる. 表 3.10 にテストデータを分類する際に要する時間を示す. 本実験で用いた計算機のスペックは, 3.40GHz Intel Core i7-3770 であり, メモリは 16GB である. 結果より, 本手法が絶対的に高速であることが確認できる. 以上より, 本手法は高精度かつ高速に変状を分類可能となることから, 実際の点検業務を支援するために有効な手法であることが示された.

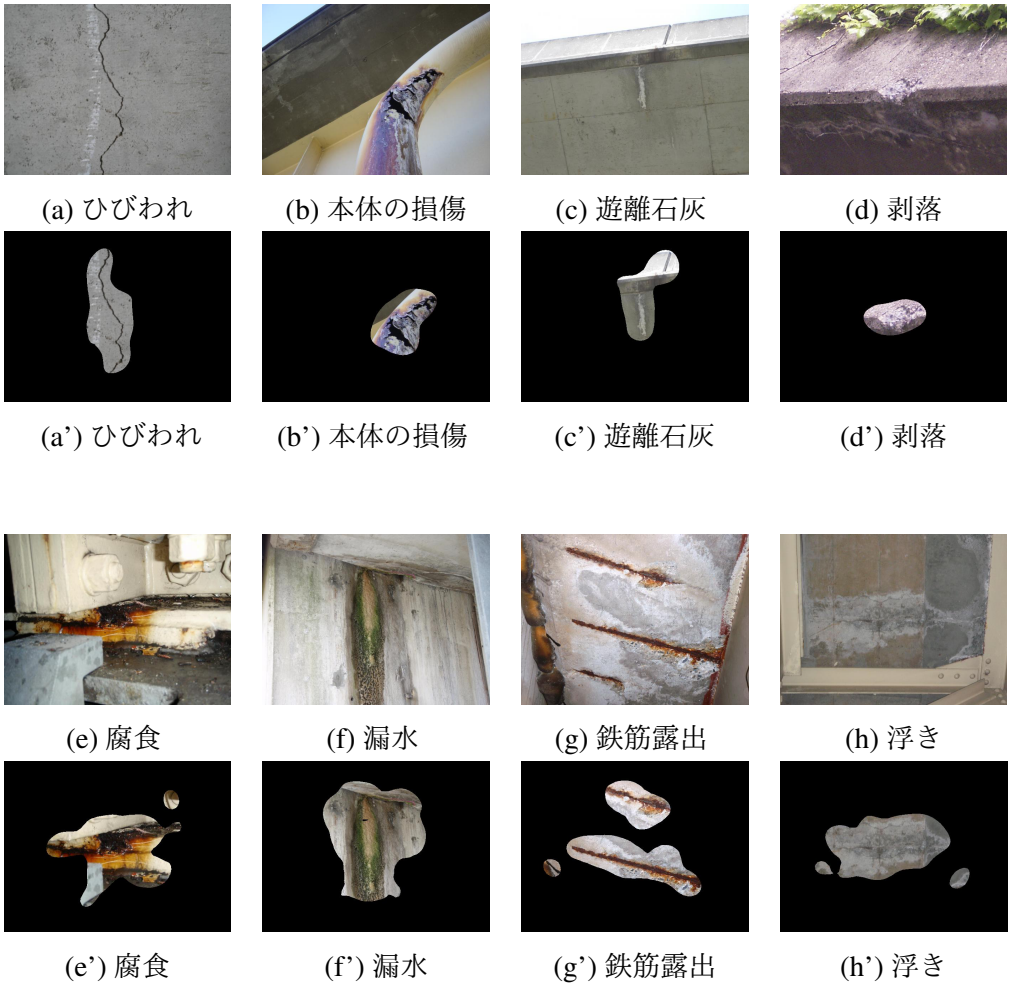


図 3.8: 本手法により正しく分類された変状画像: (a) – (h) は元画像であり, (a') – (h') は領域が抽出された後の画像である.

表 3.8: 各手法の Recall, Precision および F 値 ($\Phi = 5$).

	比較手法 A			比較手法 B			本手法		
変状	R	P	F	R	P	F	R	P	F
ひびわれ	0.060	0.181	0.076	0.846	0.854	0.848	**0.905**	**0.874**	**0.889**
本体の損傷	0.325	0.562	0.386	0.992	**0.988**	**0.990**	**1.000**	0.977	0.988
遊離石灰	0.379	0.300	0.328	0.781	**0.807**	0.793	**0.787**	**0.807**	**0.796**
剥落	0.579	0.281	0.373	**0.801**	0.783	0.791	0.799	**0.828**	**0.813**
腐食	0.710	0.675	0.681	0.981	0.982	0.981	**1.000**	**1.000**	**1.000**
平均	0.411	0.400	0.369	0.880	0.882	0.880	**0.898**	**0.897**	**0.897**

表 3.9: 各手法の Recall, Precision および F 値 ($\Phi = 7$).

	比較手法 A			比較手法 B			本手法		
変状	R	P	F	R	P	F	R	P	F
ひびわれ	0.113	0.189	0.138	0.814	**0.812**	0.813	**0.859**	0.805	**0.830**
本体の損傷	0.137	0.090	0.093	0.933	0.863	0.897	**0.935**	**0.870**	**0.901**
遊離石灰	0.110	0.237	0.143	0.685	**0.685**	0.684	**0.736**	0.672	**0.701**
剥落	0.392	0.155	0.210	**0.672**	0.601	0.633	0.615	**0.686**	**0.646**
腐食	0.233	0.268	0.241	0.975	0.988	0.981	**1.0**	**0.998**	**0.999**
漏水	0.358	0.317	0.321	0.707	0.863	0.776	**0.744**	**0.879**	**0.805**
鉄筋露出	0.147	0.281	0.185	**0.667**	**0.673**	**0.667**	0.663	0.664	0.662
平均	0.213	0.220	0.190	0.779	0.783	0.778	**0.793**	**0.796**	**0.792**

表 3.10: 本手法における計算時間.

処理	計算時間 (sec)
特徴抽出	3.29
SVM と FK-NN による分類	1.00×10^{-3}
モデル選択による分類	1.36×10^{-4}

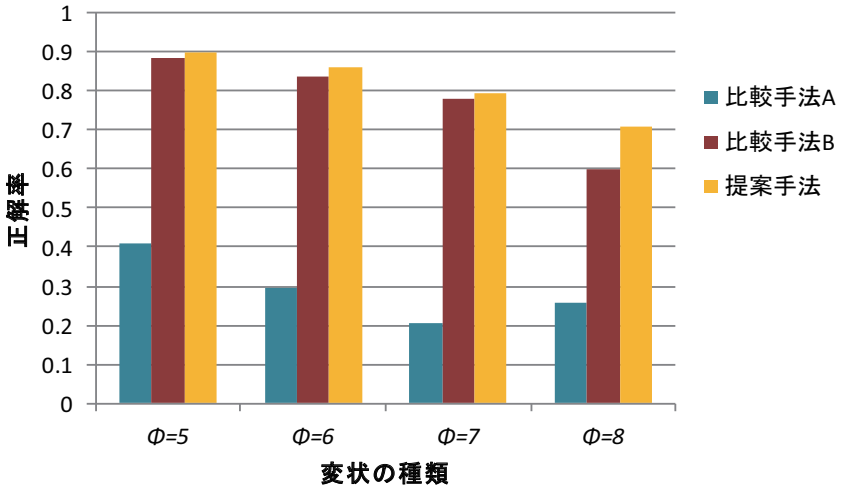


図 3.9: 比較手法 A, B および提案手法の正解率の比較.

3.6　まとめ

本章では, 2 種類のベイジアンネットワークをテキストデータの組み合わせによって算出される信頼度に基づいて, 分類器として用いるベイジアンネットワークを適応的に選択する手法を提案した. 具体的に, テキストデータを用いて Text-based BN を構築し, テキストデータと変状画像を用いて Multi-feature BN を構築する. 次に, Text-based BN から信頼度を算出する. 最後に, テストデータの持つテキストデータの組み合わせに基づいて得られる信頼度を用いて, 入力するベイジアンネットワークを適応的に選択した. 以上より, 実際の維持管理業務のプロセスに基づき, 異種データを用いて構築した手法により, 変状分類の高精度化が実現したことから, 変状分類において異種特徴間の関連性の存在が明らかとなった.

第4章　異種特徴の正準相関を考慮した変状分類

4.1　はじめに

本章では, 異種特徴に正準相関分析 [21] を適用することで変状分類に有効な特徴へ変換可能な射影を導出し, それらの特徴を用いた変状分類手法を構築する. 異種特徴間の正準相関を最大化する射影によって変換される特徴は, 分類の精度向上に寄与するという報告 [19] に基づき, これらの相関関係を考慮可能な正準相関分析を導入することで, 特徴間の相関最大化を実現する射影を導出する. これにより, 画像特徴とテキスト特徴の両者を互いに高相関な空間へ変換可能となり, 変換後の特徴を用いて変状分類を行うことで変換前の特徴を用いるよりも高精度な変状分類を実現する. また, 本手法は, 図 4.1 に示すように, 相関最大化のための射影をニューラルネットワークの隠れ層に用いることで Correlation-maximizing Weighted Extreme Learning Machine (以降, CMWELM) を構築する. まず, 変状画像から画像特徴, テキストデータからテキスト特徴を算出する. その後, それらの特徴に正準相関分析を適用することで, 射影を算出する. これをニューラルネットワークの一種である Weighted Extreme Learning Machine (以降, WELM) [51] の隠れ層間の重みとして設定し特徴変換および分類を行う.

以降, 4.2 節で, 画像特徴とテキスト特徴の算出を行い, 4.3 節で, 異種特徴間の相関関係に基づく射影の導出を行う. つづいて, 4.4 節で, 特徴変換された特徴を用いた変状分類について説明し, 4.5 節で, 実験を行い, 最後に 4.6 節で, まとめとする.

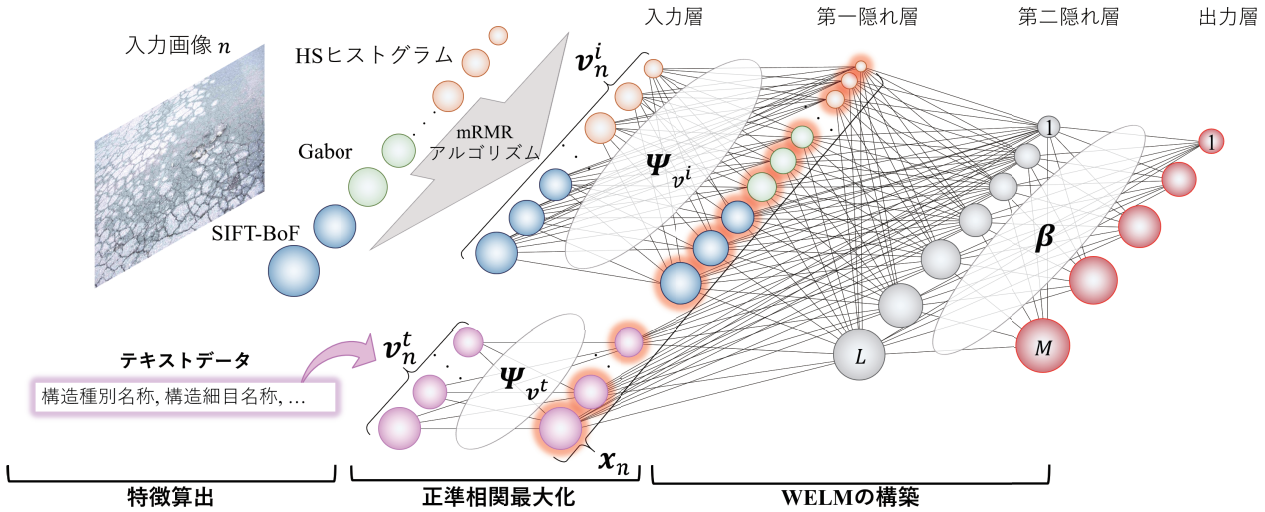


図 4.1: 提案手法の概要図. 本手法は 3 つのステップから成る. 1 つ目は, 特徴算出. 2 つ目は画像特徴とテキスト特徴間の正準相関を最大化する変換を行う隠れ層の構築. 最後に, WELM に基づく最終層の重みの導出を行う.

表 4.1: 本手法で用いる画像特徴の詳細とその次元.

特徴の詳細	次元
SIFT-BOF 特徴 [39]	1000
HS histogram に基づく特徴 [38]	128
Gabor wavelet に基づく特徴 [37]	48
テキスト特徴	216

4.2　画像特徴の算出およびテキスト特徴の算出

学習画像 n ($n = 1, 2, ..., N$; N は学習画像数) が与えられた際, 変状画像から画像特徴を抽出する. 具体的に, 表 4.1 に示す通り, Bag of Features (BoF) 表現された Scale-Invariant Feature Transform (SIFT) 特徴 [39], HS ヒストグラムに基づく特徴 [38] および Gabor wavelet に基づく特徴 [37] を算出する. ただし, “H” と “S” は色相と彩度を表す. 提案手法で用いるこれらの画像特徴は, 画像認識分野における様々なタスクで用いられており, 変状分類においても, その有効性が示されている [52, 48].

画像特徴の算出およびその次元の削減のために, 提案手法では, 相互情報量に基づく次元削減手法である minimum-Redundancy-Maximum-Relevance (mRMR) アルゴリズム [40] を採用する. mRMR アルゴリズムにより, 特徴間の冗長性を低く, 特徴とクラスの関連性を高くなるような画像特徴ベクトル $\boldsymbol{v}_n^i \in \mathbb{R}^{d^i}$ を算出する. ただし, d^i は画像特徴の次元数を表す. さらに, 我々は, 学習画像 n に対応するテキストデータ (表 4.2) からテキスト特徴ベクトル $\boldsymbol{v}_n^t \in \mathbb{R}^{d^t}$ を算出する. ただし, d^t はテキスト特徴ベクトルの次元数である. 具体的に, テキスト特徴ベクトルはバイナリ値によって構成される. テキスト特徴は, 各点検項目 $t(t = 1,\ 2,\ ...,\ T)$ が取りうる点検結果の種類を d_t とし, $\Sigma_{t=1}^{T} S_t$ 次元のバイナリベクトルとする. 対応する値を 1, そうでない場合を 0 とする.

表 4.2: 点検データ中の点検項目の例. 表中の ID は, 図 4.2 の変状画像の例と対応している.

		点検項目			
ID	変状	構造種別名称	構造細目名称	···	変状部位名称
1	ひびわれ	地覆	地覆・高欄	···	左・外・下
2	遊離石灰	RC 床板	コンクリート上部工	···	主版部
3	腐食	鋼桁	鋼構造物	···	主桁フランジ
4	剥落	PC 桁	コンクリート上部工	···	横桁
5	鉄筋露出	PC 桁	コンクリート上部工	···	主桁フランジ

4.3 相関関係を考慮することによる変状に関連のある特徴の算出

本節では, CMWELM において, 画像特徴とテキスト特徴間の正準相関を最大化する変換を行う第一隠れ層の構築について説明する. これらの特徴間の関係を考慮可能な重み行列を算出するために, 画像特徴行列 $\boldsymbol{V}^i = [\boldsymbol{v}_1^i, \boldsymbol{v}_2^i, ..., \boldsymbol{v}_N^i] \in \mathbb{R}^{d^i \times N}$ とテキスト特徴行列 $\boldsymbol{V}^t = [\boldsymbol{v}_1^t, \boldsymbol{v}_2^t, ..., \boldsymbol{v}_N^t] \in \mathbb{R}^{d^t \times N}$ に正準相関分析を適用することで, 射影行列 $\boldsymbol{\Psi}_{\boldsymbol{v}^i}$ と $\boldsymbol{\Psi}_{\boldsymbol{v}^t}$ を算出する. CCA を計算するために, 以下の目的関数を最大化する.

$$(\hat{\boldsymbol{\psi}}_{\boldsymbol{v}^i}, \hat{\boldsymbol{\psi}}_{\boldsymbol{v}^t}) = \arg \max_{\boldsymbol{\psi}_{\boldsymbol{v}^i}, \boldsymbol{\psi}_{\boldsymbol{v}^t}} \frac{\boldsymbol{\psi}_{\boldsymbol{v}^i}^{\mathrm{T}} \boldsymbol{C}_{V^i V^t} \boldsymbol{\psi}_{\boldsymbol{v}^t}}{\sqrt{\boldsymbol{\psi}_{\boldsymbol{v}^i}^{\mathrm{T}} \boldsymbol{C}_{V^i V^i} \boldsymbol{\psi}_{\boldsymbol{v}^i}} \sqrt{\boldsymbol{\psi}_{\boldsymbol{v}^t}^{\mathrm{T}} \boldsymbol{C}_{V^t V^t} \boldsymbol{\psi}_{\boldsymbol{v}^t}}} \tag{4.1}$$

ただし, $\boldsymbol{\psi}_{\boldsymbol{v}^i}$ と $\boldsymbol{\psi}_{\boldsymbol{v}^t}$ は, 射影ベクトルであり,

$$\boldsymbol{C}_{V^i V^t} = \boldsymbol{V}^i \boldsymbol{V}^{t\top} \tag{4.2}$$

$$\boldsymbol{C}_{V^i V^i} = \boldsymbol{V}^i \boldsymbol{V}^{i\top} \tag{4.3}$$

$$\boldsymbol{C}_{V^t V^t} = \boldsymbol{V}^t \boldsymbol{V}^{t\top} \tag{4.4}$$



(a) ひびわれ (ID 1)



(b) 遊離石灰 (ID 2)



(c) 腐食 (ID 3)



(d) 剥落 (ID 4)



(e) 鉄筋露出 (ID 5)

図 4.2: 変状画像の例. ただし, 表 4.2 に示す ID に対応している.

となる. 式 (4.1) を最大化するために, 以下の式を解く.

$$\begin{aligned}\mathcal{L}(\boldsymbol{\psi}_{\boldsymbol{v}^i}, \boldsymbol{\psi}_{\boldsymbol{v}^t}) = \boldsymbol{\psi}_{\boldsymbol{v}^i}^\top C_{V^iV^t}\boldsymbol{\psi}_{\boldsymbol{v}^t} &- \frac{\lambda_{\boldsymbol{v}^i}}{2}(\boldsymbol{\psi}_{\boldsymbol{v}^i}^\top C_{V^iV^i}\boldsymbol{\psi}_{\boldsymbol{v}^i} - 1) \\ &- \frac{\lambda_{\boldsymbol{v}^t}}{2}(\boldsymbol{\psi}_{\boldsymbol{v}^t}^\top C_{V^tV^t}\boldsymbol{\psi}_{\boldsymbol{v}^t} - 1)\end{aligned} \tag{4.5}$$

ただし, $\lambda_{\boldsymbol{v}^i} = \lambda_{\boldsymbol{v}^t}(= \lambda)$ であり, λ は以下の問題の固有値に対応している.

$$C_{V^iV^i}^{-1}C_{V^iV^t}C_{V^tV^t}^{-1}C_{V^iV^t}^\top\boldsymbol{\psi}_{\boldsymbol{v}^i} = \lambda^2\boldsymbol{\psi}_{\boldsymbol{v}^i} \tag{4.6}$$

$$C_{V^tV^t}^{-1}C_{V^iV^t}^\top C_{V^iV^i}^{-1}C_{V^iV^t}\boldsymbol{\psi}_{\boldsymbol{v}^t} = \lambda^2\boldsymbol{\psi}_{\boldsymbol{v}^t} \tag{4.7}$$

複数の固有値 λ とそれらに対応する固有ベクトル $\boldsymbol{\psi}_{\boldsymbol{v}^i}$ と $\boldsymbol{\psi}_{\boldsymbol{v}^t}$ が算出されるので, d^{cca} 次元の係数ベクトルを並べることで, 射影行列 $\boldsymbol{\Psi}_{\boldsymbol{v}^i} = [\boldsymbol{\psi}^1_{\boldsymbol{v}^i}, \boldsymbol{\psi}^2_{\boldsymbol{v}^i}, ..., \boldsymbol{\psi}^{d^{cca}}_{\boldsymbol{v}^i}]^\top \in \mathbb{R}^{d^{cca} \times d^i}$ および $\boldsymbol{\Psi}_{\boldsymbol{v}^t} = [\boldsymbol{\psi}^1_{\boldsymbol{v}^t}, \boldsymbol{\psi}^2_{\boldsymbol{v}^t}, ..., \boldsymbol{\psi}^{d^{cca}}_{\boldsymbol{v}^t}]^\top \in \mathbb{R}^{d^{cca} \times d^t}$ を算出する. CMWELM の第一隠れ層に射影行列を用いることで, 以下の関係を考慮可能な射影ベクトル $\boldsymbol{u}^i_n \in \mathbb{R}^{d^{cca}}$ と $\boldsymbol{u}^t_n \in \mathbb{R}^{d^{cca}}$ を算出する.

$$\boldsymbol{u}^i_n = \boldsymbol{\Psi}_{\boldsymbol{v}^i} \boldsymbol{v}^i_n \tag{4.8}$$

$$\boldsymbol{u}^t_n = \boldsymbol{\Psi}_{\boldsymbol{v}^t} \boldsymbol{v}^t_n \tag{4.9}$$

多くの深層学習では, 複数の隠れ層を使用することによって, 画像特徴から意味的な特徴へ変換する [53]. これらの手法では, 決定すべき隠れ層のパラメータの数が非常に多いので, 多数の学習データが必要となる [8, 24, 54]. 一方, 第一隠れ層の重みパラメータとして, 画像特徴とテキスト特徴との間の正準相関を考慮可能な CCA により得られた射影行列を用いることで, 画像特徴から意味的な特徴に直接変換することが可能となる. そこで, 本手法では一つの隠れ層のみを用いて変換が可能となるため, 少量の学習データを用いてネットワークを構築することができる.

4.4　得られた特徴を用いた変状分類

特徴ベクトル $\boldsymbol{u}^i_n$ と $\boldsymbol{u}^t_n$ が与えられたとき, $\boldsymbol{x}_n = [\boldsymbol{u}^{i\top}_n, \boldsymbol{u}^{t\top}_n]^\top \in \mathbb{R}^{d^x}$ $(n = 1, 2, ..., N)$ を WELM へ入力する. ただし, $d^x = 2d^{cca}$ である. ここで, ELM は, 以前より提案されている Random Vector Functional Link Network [55, 56] の構造を模倣した手法である. WELM は図 4.1 に示す通り, 3 層から構成される. WELM の学習の目的は, 第二隠れ層と出力層間の重み行列 $\boldsymbol{\beta} \in \mathbb{R}^{M \times L}$ (M クラス分類問題における出力層のノード数であり, L は第二隠れ層のノード数である) を算出することにある. 重み行列 $\boldsymbol{\beta}$ を算出するために, 学習誤差 $\boldsymbol{\xi}_n = [\xi_{n,1}, \xi_{n,2}, ..., \xi_{n,M}]^\top$

を最小化するように最適化を行う.

$$\begin{aligned} \min_{\boldsymbol{\beta}\in\mathbb{R}^{M\times L}} R \quad &= \frac{1}{2}\|\boldsymbol{\beta}\|_F^2 + \frac{C}{2}\sum_{n=1}^{N} W_{nn}\|\boldsymbol{\xi}_n\|^2 \\ \text{s.t.} \qquad &\boldsymbol{\beta}\boldsymbol{h}(\boldsymbol{x}_n) = \boldsymbol{t}_n - \boldsymbol{\xi}_n \end{aligned} \tag{4.10}$$

ただし, $\boldsymbol{t}_n = [t_{n,1}, t_{n,2}, ..., t_{n,M}]^\top$ はクラスを表すベクトルであり, サンプルがクラス m に属する場合, それに対応する要素が 1 となり, それ以外が 0 となる. さらに, C は正則化パラメータである. $\boldsymbol{W} = diag\{W_{nn}\}$ はクラス不均衡を解決するために導入する重み行列である. さらに, $\boldsymbol{h}(\boldsymbol{x}_n) \in \mathbb{R}^L$ は第二隠れ層の出力ベクトルである.

$$\begin{aligned} \boldsymbol{h}(\boldsymbol{x}_n) &= [h_1(\boldsymbol{x}_n), h_2(\boldsymbol{x}_n), ..., h_L(\boldsymbol{x}_n)]^\top \\ &= [G(\boldsymbol{a}_1, b_1, \boldsymbol{x}_n), ..., G(\boldsymbol{a}_L, b_L, \boldsymbol{x}_n)]^\top \end{aligned} \tag{4.11}$$

ただし, $\boldsymbol{h}(\boldsymbol{x}_n)$ は d^x 次元の入力空間から L 次元の隠れ層の空間へマッピングする射影であり, G は活性化関数である. さらに, $\boldsymbol{a}_l$ $(l = 1, 2, ..., L)$ と b_l $(l = 1, 2, ..., L)$ は活性化関数のパラメータである. 文献 [57] に基づき, 式 (4.10) は Karush-Kuhn-Tucker 条件 [58] に基づく以下の最適化問題を解くことと等価である.

$$\begin{aligned} \min_{\boldsymbol{\beta}\in\mathbb{R}^{M\times L}} \tilde{R} = &\frac{1}{2}\|\boldsymbol{\beta}\|_F^2 + \frac{C}{2}\sum_{n=1}^{N} W_{nn}\|\boldsymbol{\xi}_n\|^2 \\ &- \sum_{n=1}^{N}\sum_{m=1}^{M} \alpha_{n,m}\Big\{(\boldsymbol{\gamma}_m)^\top \boldsymbol{h}(\boldsymbol{x}_n) - t_{n,m} + \xi_{n,m}\Big\} \end{aligned} \tag{4.12}$$

ただし, $\boldsymbol{\gamma}_m = [\boldsymbol{\beta}_{1,m}, \boldsymbol{\beta}_{2,m}, ..., \boldsymbol{\beta}_{L,m}]^\top$ は, m 番目の出力ノードと隠れ層をつなぐ重みベクトルである. 最適な重み $\boldsymbol{\beta}$ を以下の式で求める.

$$\boldsymbol{\beta} = \boldsymbol{T}\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{W}\right)^{-1}\boldsymbol{H}^\top \tag{4.13}$$

ただし,

$$\begin{aligned} \boldsymbol{H} &= [\boldsymbol{h}(\boldsymbol{x}_1), \boldsymbol{h}(\boldsymbol{x}_2), ..., \boldsymbol{h}(\boldsymbol{x}_N)] \\ &= \begin{bmatrix} G(\boldsymbol{a}_1, b_1, \boldsymbol{x}_1) & \cdots & G(\boldsymbol{a}_1, b_1, \boldsymbol{x}_N) \\ \vdots & \ddots & \vdots \\ G(\boldsymbol{a}_L, b_L, \boldsymbol{x}_1) & \cdots & G(\boldsymbol{a}_L, b_L, \boldsymbol{x}_N) \end{bmatrix} \end{aligned} \tag{4.14}$$

さらに, $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, ..., \boldsymbol{t}_N]$ である. WELM では, クラス数の不均衡を考慮するために, 以下に示すクラス情報に基づき, 重み行列 $\boldsymbol{W}$ を定義する.

$$W_{nn} = \frac{1}{P_n} \tag{4.15}$$

ただし, P_n は画像 n に属するクラスの画像数である. 各クラスの学習画像数で割ることで, 式 (4.10) の学習誤差の項における少数クラスに属するデータへの影響を大きくする. 式 (4.10) によると, WELM は学習誤差を最小化するので, 画像数が少ないほど, その影響は大きくなる. また, 提案手法は活性化関数として以下に示すガウス関数を用いている.

$$\begin{aligned} G(\boldsymbol{a}_l, b_l, \boldsymbol{x}_n) &= \exp(-b_l \|\boldsymbol{x}_n - \boldsymbol{a}_l\|^2) \\ &= K(\boldsymbol{x}_n, \boldsymbol{a}_l) \end{aligned} \tag{4.16}$$

ただし, $K(\cdot, \cdot)$ はカーネルを表す. Mercer の定理を用いることで, カーネル行列を以下の通り算出する.

$$\boldsymbol{\Omega} = \boldsymbol{H}^T \boldsymbol{H}, \tag{4.17}$$

$$\begin{aligned} \Omega_{n,j} &= \boldsymbol{h}(\boldsymbol{x}_n)^\top \cdot \boldsymbol{h}(\boldsymbol{x}_j) \\ &= K(\boldsymbol{x}_n, \boldsymbol{x}_j) \end{aligned} \tag{4.18}$$

式 (4.13) は, 式 (4.18) に基づき, 簡潔に書き直すことができる.

$$\boldsymbol{\beta} = \boldsymbol{T}\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{\Omega}\boldsymbol{W}\right)^{-1}\boldsymbol{H}^{\top} \tag{4.19}$$

以上より, ノード数 L はあらかじめ与えられる必要はない.

新たなテストデータ $\boldsymbol{h}(\boldsymbol{x})$ が与えられたとき, 得られた $\boldsymbol{\beta}$ を用いることで, 出力ベクトル $\boldsymbol{f} = [f_1,\ f_2,\ ...,\ f_M]^{\top}$ が定義される.

$$\boldsymbol{f} = \boldsymbol{T}\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{\Omega}\boldsymbol{W}\right)^{-1}\boldsymbol{H}^{\top}\boldsymbol{h}(\boldsymbol{x}) \tag{4.20}$$

さらに, カーネル関数を用いることで式 (4.20) は以下の通り書き直される.

$$\boldsymbol{f} = \boldsymbol{T}\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{\Omega}\boldsymbol{W}\right)^{-1}\begin{bmatrix} K(\boldsymbol{x}, \boldsymbol{x}_1) \\ \vdots \\ K(\boldsymbol{x}, \boldsymbol{x}_N) \end{bmatrix} \tag{4.21}$$

最終結果は $class = \arg\max_{m\in\{1,...,M\}} f_m$ として算出される. 以上のように, CMWELM に基づく変状分類が実現する. 結果的に, 効果的な特徴変換とクラス不均衡な点検データに対処可能な分類が可能となる.

4.5　実験

本節では, 提案手法の有効性を示すための実験を行う. 実験条件についてを 4.5.1 で示し, 提案手法の精度検証についてを 4.5.2 で示す.

4.5.1　実験準備

本実験では, 東日本高速道路株式会社 (NEXCO 東日本) よりご提供頂いた, 点検データを用いた. 本実験で用いる点検データの数を表 4.3 に示す. この表より, これらの学習データはクラス不均衡である. 本実験では, 提案手法の有効性を証

表 4.3: 本実験で用いる点検データの数.

変状	学習データ数	テストデータ数	交差検証用のデータ数
ひびわれ	2,893	100	100
遊離石灰	3,537	100	100
腐食	1,262	100	100
剥落	736	100	100
鉄筋露出	874	100	100
合計	9,302	500	500

表 4.4: パラメータの探索範囲.

パラメータ	範囲
画像特徴の次元 d^i	$\{10,\ 40,\ 70,\ ...,\ 220,\ 250,\ 280\}$
正則化パラメータ C	$\{2^{-10},\ 2^{-5},\ 2^{-1},\ 2^{1},\ 2^{5},\ 2^{10},\ 2^{20},\ 2^{50}\}$
カーネルパラメータ b_l	$\{2^{-50},\ 2^{-20},\ 2^{-10},\ 2^{-5},\ 2^{-1},\ 2^{1},\ 2^{5},\ 2^{10}\}$

表 4.5: 提案手法で用いられるパラメータの詳細.

詳細	パラメータ	値
次元数	d^i	280
	d^t	216
	d^x	378
正則化パラメータ	C	2^{10}
カーネルパラメータ	b_l	2^{-10}
クラス数	M	5

明するために, 5 クラス分類問題を解く. 我々は, 表 4.3 に示すように, 学習データ, テストデータおよび交差検証用データの 3 種類のデータを用いた. パラメータ設定においては, 交差検証用データを用い, それぞれの分類器において最も精度が良

表 4.6: 正解のクラスと推定クラスとの間の関係.

		真のクラス	
		対象のクラス	対象ではないクラス
推定クラス	対象のクラス	TP	FP
	対象ではないクラス	FN	TN

くなる時の値を採用した. ただし, 表 4.4 に示す範囲でパラメータ探索を行った. CMWELM で用いるパラメータの詳細は表 4.5 に示されている. 本実験では, 以下に示す Recall, Precision および F 値を用いて精度検証を行う.

$$\begin{aligned} \text{Recall} &= \frac{\text{正しく分類されたサンプルの数}}{\text{そのクラスのサンプル数}} \\ &= \frac{\text{TP}}{\text{TP} + \text{FN}} \end{aligned} \tag{4.22}$$

$$\begin{aligned} \text{Precision} &= \frac{\text{正しく分類されたサンプルの数}}{\text{手法によりそのクラスに分類されたサンプル数}} \\ &= \frac{\text{TP}}{\text{TP} + \text{FP}} \end{aligned} \tag{4.23}$$

$$\text{F 値} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{4.24}$$

ここで, 正しく分類されたサンプルとは, ある対象クラスに分類されたサンプルを示す (true positive: TP). 上記の式の要素間の関係を表 4.6 に示す. Recall, Precision および F 値は機械学習分野における一般的な評価指標の一つであるので, 採用した. Recall と Precision は誤分類と過検出の比率を表す. しかしながら, Recall と Precision はトレードオフの関係を持つため, Recall と Precision の調和平均である F 値を総合性能の評価のために用いる.

提案手法の有効性を示すために, 一般的な ELM [59], WELM [60], Multi Layer WELM (以降, MLWELM) [61] および CNN に基づくアプローチの一つである Fine-tuned VGG16 [24] を用いる. WELM は文献 [51] に基づいて構築した我々の従来手法 [60] である. WELM の改良版である MLWELM は WELM を多層化した深層

学習手法である. 比較手法の詳細を以下に示す. VGG16 を除くすべての比較手法の構築においては, 簡便に画像特徴ベクトルとテキスト特徴ベクトルを並べた特徴ベクトル $\boldsymbol{x}_n = [\boldsymbol{v}^{i\top}, \boldsymbol{v}^{t\top}]^\top$ を入力として用いる.

- ELM

  ELM は 3 種類の層 (入力層, 隠れ層および出力層) を用いることで構築する. ELM の学習の目的は, ノード数 L を持つ隠れ層とノード数 M を持つ出力層間の重み行列 $\boldsymbol{\beta} \in \mathbb{R}^{M\times L}$ を算出することである. 以下の式を満たすような $\boldsymbol{\beta}$ を求める.

  $$\boldsymbol{\beta}\boldsymbol{H} = \boldsymbol{T} \tag{4.25}$$

  ただし, $\boldsymbol{H}$ と $\boldsymbol{T}$ は CMWELM と同様に求めることができる. しかしながら, ELM の活性化関数としてシグモイド関数を用いる.

  $$G(\boldsymbol{a}_l, b_l, \boldsymbol{x}_n) = \frac{1}{1 + \exp(-(\boldsymbol{a}_l^\top \boldsymbol{x}_n + b_l))} \tag{4.26}$$

  ただし, $\boldsymbol{a}_l$ と b_l はランダムに決定される. 最小二乗問題を解くために, 式 (4.25) は下式で解くことができる.

  $$\boldsymbol{\beta} = \boldsymbol{T}\left(\boldsymbol{H}^\top \boldsymbol{H}\right)^{-1} \boldsymbol{H}^\top \tag{4.27}$$

- WELM

  我々の従来手法である WELM に基づく変状分類は 4.4 節と同じ手順で構築可能である.

- MLWELM

  特徴ベクトルを $\boldsymbol{z}_n^0 = \boldsymbol{x}_n$ と定義することで, 我々は MLWELM を学習する. MLWELM は入力層, K 個の隠れ層および出力層から構成される. つまり,

MLWELM は $K+2$ 層を有する. MLWELM の学習の目的は, 隠れ層と出力層間の重み行列を計算することである．MLWELM は特徴変換 **(I)** と分類 **(II)** の二つのステップから構成される. 具体的に, k が $K+1$ よりも小さい時, WELM に基づく Auto Encoder (以降, WELM-AE) により特徴変換される. また, $k=K+1$ の時, WELM により分類される.

(I) $k=1,2,...,K$

MLWELM では, k 番目の隠れ層の出力行列 $\boldsymbol{Z}^k=[\boldsymbol{z}_1^k,\boldsymbol{z}_2^k,...,\boldsymbol{z}_N^k]$ がシグモイド関数 G と $\boldsymbol{z}_n^k=G(\boldsymbol{\beta}^k\boldsymbol{z}_n^{k-1})$ に基づく $\boldsymbol{Z}^{k-1}$ によって計算される．ただし, $\boldsymbol{\beta}^k\in\mathbb{R}^{L^k\times L^{k-1}}$ は k 番目と $(k-1)$ 番目の隠れ層間の重み行列．行列 $\boldsymbol{\beta}^k$ は, 3 層で構成される WELM-AE により構成される．これらの層は, 入力ノード数は L^{k-1}, 隠れノード数は L^k および出力ノード数は L^{k-1} である．入力ベクトル $\boldsymbol{z}_n^{k-1}\in\mathbb{R}^{L^{k-1}}$ が与えられたとき, WELM-AE の隠れ層の出力 $\boldsymbol{h}_n^k\in\mathbb{R}^{L^k}$ は $\boldsymbol{h}_n^k=G(\boldsymbol{A}^k\boldsymbol{z}_n^{k-1}+\boldsymbol{b}^k)$ として算出される．ただし, $\boldsymbol{A}^k(\boldsymbol{A}^k)^\top=\boldsymbol{I}$ であり, $(\boldsymbol{b}^k)^\top\boldsymbol{b}^k=1$ である．WELM-AE はランダム値に基づく直交する重み $\boldsymbol{A}^k=[\boldsymbol{a}_1^k,\boldsymbol{a}_2^k,...,\boldsymbol{a}_{L^k}^k]^\top$ とランダム値に基づくバイアス $\boldsymbol{b}^k=[b_1^k,b_2^k,...,b_{L^k}^k]^\top$ を持つ. WELM-AE の入力行列 $\boldsymbol{Z}^{k-1}$ と隠れ層の出力行列 $\boldsymbol{H}^k=[\boldsymbol{h}_1^k,\boldsymbol{h}_2^k,...,\boldsymbol{h}_N^k]^\top\in\mathbb{R}^{N\times L^k}$ を用いることで, 重み行列 $\boldsymbol{\beta}^k$ は以下の通り得られる.

$$\boldsymbol{\beta}^k=(\boldsymbol{Z}^{k-1})^\top\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C_1}\sum_{l^k=1}^{L^k}\mathrm{KL}(\rho||\hat{\rho_{l^k}})+(\boldsymbol{H}^k)^\top\boldsymbol{H}^k\boldsymbol{W}\right)^{-1}(\boldsymbol{H}^k)^\top \quad (4.28)$$

ただし, C_1 は正則化パラメータであり, $\boldsymbol{W}$ は CMWELM と同様に求められる. $\mathrm{KL}(\rho||\hat{\rho}_{l^k})=\rho\log(\rho/\hat{\rho}_{l^k})+(1-\rho)\log((1-\rho)/(1-\hat{\rho}_{l^k}))$ は, スパース正則化のための KL ダイバージェンスである. KL ダイバージェンスは, しばしば活性化しているニューロンの数を制御するためのオートエンコーダの隠れ層で使われる. ρ は, 予め決定されているスパースのためのパラメータであり, この値は低ければ低いほどスパース性が高くなることを表す. ただ

し, 本実験において ρ は 0.05 である. さらに, $\hat{\rho}_{l^k}$ は k 番目の隠れ層における l^k 番目のノードの平均活性を表している.

(II) $k = K + 1$
最終隠れ層と出力層間の出力重み行列 $\boldsymbol{\beta}^{K+1} \in \mathbb{R}^{M \times L^K}$ は WELM によって算出される. 以下の式において, 学習誤差 $\boldsymbol{\xi}_n = [\xi_{n,1}, \xi_{n,2}, ..., \xi_{n,M}]^\top$ を最小化する.

$$\min_{\boldsymbol{\beta}^{K+1}} \frac{1}{2}\|\boldsymbol{\beta}^{K+1}\|_F^2 + \frac{C_2}{2}\sum_{n=1}^{N} W_{nn}\|\boldsymbol{\xi}_n\|^2 \quad \text{s.t. } \boldsymbol{\beta}^{K+1}\boldsymbol{z}_n^K = \boldsymbol{t}_n - \boldsymbol{\xi}_n \tag{4.29}$$

ただし, C_2 は, 正則化パラメータである. $\boldsymbol{t}_n$ と $\boldsymbol{T}$ は CMWELM と同様に得られる. 最後に, 最適な重み $\boldsymbol{\beta}^{K+1}$ は以下の式で得られる.

$$\boldsymbol{\beta}^{K+1} = \boldsymbol{T}\boldsymbol{W}\left(\frac{\boldsymbol{I}}{C} + (\boldsymbol{Z}^{K+1})^\top \boldsymbol{Z}^{K+1}\boldsymbol{W}\right)^{-1}(\boldsymbol{Z}^{K+1})^\top \tag{4.30}$$

次に, 各比較手法と提案手法との比較により以下の内容が示される.

- CMWELM vs ELM
  CMWELM がベースライン手法よりも効果的であることを確認する.
- CMWELM vs WELM
  CCA により得られる射影特徴を利用することの有効性を確認する.
- CMWELM vs MLELM
  WELM-AE による教師無し特徴変換よりも正準相関を考慮した特徴変換が有効であることを確認する.
- CMWELM vs VGG16
  CNN に基づく手法よりも有効であることを確認する.

さらに, ELM と WELM を比較することで, クラス不均衡を考慮することが変状分類に有効であることを確認する.

表 4.7: 全ての手法における Recall, Precision および F 値.

変状	CMWELM (Ours)			ELM			WELM			MLWELM			VGG16		
	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
ひびわれ	0.860	**0.843**	**0.851**	0.830	0.658	0.734	0.840	0.736	0.785	0.860	0.826	0.843	**0.880**	0.578	0.698
遊離石灰	0.730	**0.608**	0.663	0.680	0.531	0.596	0.750	0.585	0.657	0.750	0.600	**0.666**	**0.800**	0.529	0.637
腐食	**1.000**	**1.000**	**1.000**	0.990	**1.000**	0.994	0.990	**1.000**	0.994	0.990	0.980	0.985	0.920	0.920	0.920
剥落	0.410	**0.773**	0.535	0.310	0.738	0.436	0.330	0.702	0.448	**0.420**	0.750	**0.538**	0.340	0.618	0.438
鉄筋露出	**0.750**	0.600	**0.666**	0.590	0.561	0.575	0.700	0.625	0.660	0.690	0.605	0.644	0.280	**0.666**	0.394
平均	**0.750**	**0.765**	**0.743**	0.680	0.698	0.667	0.722	0.729	0.709	0.742	0.752	0.735	0.644	0.662	0.617

表 4.8: クラス均衡とクラス不均衡なデータにおける CMWELM の精度比較.

	クラス不均衡			クラス均衡		
変状	R	P	F	R	P	F
ひびわれ	**0.860**	**0.843**	**0.851**	**0.860**	0.811	0.834
遊離石灰	0.730	**0.608**	0.663	**0.750**	0.600	**0.666**
腐食	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
剥落	**0.410**	**0.773**	**0.535**	0.380	0.703	0.493
鉄筋露出	**0.750**	0.600	**0.666**	0.710	**0.612**	0.657
平均	**0.750**	**0.765**	**0.743**	0.738	0.745	0.729

4.5.2　実験結果および考察

全ての手法の Recall, Precision および F 値を表 4.7 に示す. この表より, 提案手法の Recall, Precision および F 値の平均が state-of-the-art である MLWELM を含む他の手法の精度よりも高いことが確認できる. MLWELM と提案手法を比較すると, 効果的な特徴変換を行うためには, 異種特徴間の正準相関を考慮することが多数の隠れ層を構築するよりも有効であることが示された. さらに, WELM と提案手法とを比較することで, Recall と Precision の調和平均である F 値においては, 正準相関を最大化する第一隠れ層を構築することの有効性が示された. 特に, 学習画像の少ない剥落の精度に注目すると, 提案手法は WELM を大きく上回ることも確認できる. さらに, CMWELM はほぼすべての評価指標において, VGG16 を上回るため, CMWELM は CNN に基づく手法よりも少量の学習データで効果的な学習が可能であることを確認できる. また, WELM の精度が ELM よりも高いことから, クラス不均衡を考慮することも有効であることが示された. 最後に, 正準相関係数を図 4.3 に示す. これにより, 画像特徴とテキスト特徴間を高相関とする特徴が算出されていることが確認できる.

さらに, CMWELM を用いることによるクラス不均衡データへの効果について検証する. 表 4.3 に示すように, 剥落のサンプル数である 736 サンプルに揃えた

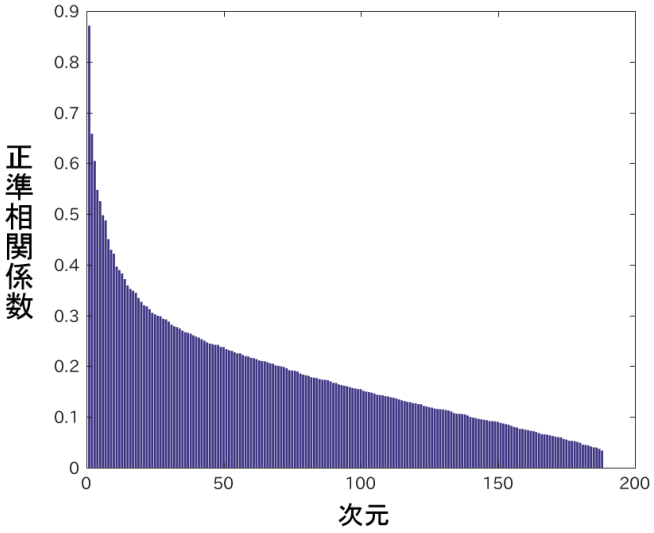


図 4.3: 正準特徴の各次元に対応する正準相関係数.

データを用いて実験を行う. 実験結果を表 4.8 に示す. クラス不均衡のデータと均衡のデータを比較することで, クラス不均衡のデータを用いた場合の精度が高いことから, クラス不均衡なデータであったとしても, 可能な限りデータを用いることが重要であることが確認できる.

実験結果より, 提案手法の精度向上は state-of-the-art である MLWELM と比較して, その差が 1%である. しかしながら, 以下に示すように, この差は大きな意味をもつ. 実際の維持管理業務におけるアプリケーションとして CMWELM を使用するためには, これまでに蓄積された大量のデータで学習する必要がある. NEXCO は約 1 億 5000 万のサンプルを所有しているので,「1%」の差は 150 万サンプルの差に相当する. したがって, この差が学習能力に大きな影響を与えることは明らかである. ここでは, NEXCO には大量のデータが保存されてるが, これらのデータには数百種類の変状が含まれており, サンプル数が不均衡なクラスが複数存在する. したがって, 表 4.3 に示すように, 不均衡が大きいために訓練データを十分に準備することが困難な場合が多い. 本手法は, このような不均衡なデータを扱う

表 4.9: 提案手法の混同行列.

	Predicted class				
True class	ひびわれ	遊離石灰	腐食	剥落	鉄筋露出
ひびわれ	0.86	0.03	0	0.05	0.06
遊離石灰	0.04	0.73	0	0.02	0.21
腐食	0	0	1.00	0	0
剥落	0.05	0.31	0	0.41	0.23
鉄筋露出	0.07	0.13	0	0.05	0.75

表 4.10: WELM の混同行列.

	Predicted class				
True class	ひびわれ	遊離石灰	腐食	剥落	鉄筋露出
ひびわれ	0.84	0.03	0	0.07	0.06
遊離石灰	0.11	0.75	0	0.02	0.12
腐食	0	0	0.99	0.01	0
剥落	0.11	0.32	0	0.33	0.24
鉄筋露出	0.08	0.18	0	0.04	0.70

ことができる. さらに, 将来劣化した構造物の数が増加すると, 誤分類されたデータの数が増加する. 実際, NEXCO は現在, 3,800 キロメートル以上の高速道路を維持管理しており, 将来的には保守点検によるデータ数の増加が明白である. 以上のことから, 提案手法は実用化に有効であることがわかる.

提案手法により正しく分類され, WELM によって誤分類された画像の例を図 4.4 に示す. 提案手法では, 図 4.4 (d), (f), (h) に示すように, 不要な領域を含む画像を正しく分類することができる. つまり, 提案手法では, 十分に高品質でない変状画像を正しく分類可能であり, また, 撮影された画像毎に異なる質やアングルを考慮して画像を扱うことが可能である. ただし, 本手法で使用されている変状画像は, 図 4.5 のように異なる角度で撮影されているため, 分類性能を十分に確保するためには, 撮影条件を標準化することが望ましい.

表 4.11: CMWELM と VGG16 の計算時間 (sec).

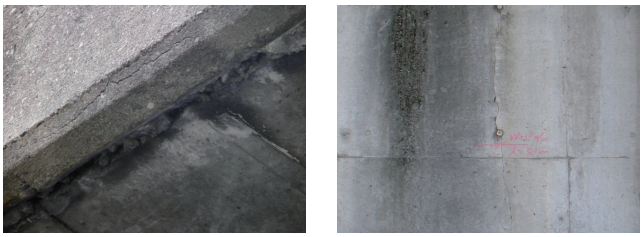
処理	CMWELM	VGG16
特徴算出	2.97×10^{-1}	-
正準相関最大化	7.37×10^{-4}	-
変状分類	2.59×10^{-1}	1.52 (特徴算出を含む)
Sum	5.56×10^{-1}	1.52

さらに, 相関最大化アプローチの有効性を検証するために, 提案手法と WELM の混同行列を表 4.9 と表 4.10 にそれぞれ, 示す. 行列の対角要素は, 表 4.7 に示す Recall の値に対応する. これらの表から, 剥落の画像をひびわれとみなす誤分類の割合は, 提案手法により減少した. 本実験で用いた点検データにおいては, アスファルトまたはコンクリートにはしばしばひびわれが生じる. また, 剥落もコンクリートでも起こる変状である. これらの変状の両方が同じ材料で発生するため, WELM は単に画像特徴およびテキスト特徴を使用し, これらの変状を誤分類した. 一方, 画像特徴とテキスト特徴の関係を考慮することにより, 提案手法はこれらの画像を正しく分類することができる. 以上からも, 正準相関を考慮することの有効性が確認された.

最後に, 提案手法の計算コストについて検証する. 各処理の計算コストを, 表 4.11 に示す. CMWELM は, 3.40GHz Intel Core i7-3770 と 16.0GB のメモリを有する計算機を使用して計算した. 一方, 一般的な深層学習方法である VGG16 の Fine-tuning は, 4.00GHz Intel Core i7-4790 と 32GB のメモリと GPU GeForce 1080Ti を使用して計算した. VGG16 モデルは GPU を用いて行ったが, CPU を用いて計算した CMWELM が VGG16 より速いことが確認されたことから, 提案手法は一般的な深層学習手法と比較して, 高速な分類を実現する. したがって, 提案手法は, 未知の対象物の変状をリアルタイムに分類することができるため, 実際の維持管理を支援する上で効果的である.

4.6　まとめ

本章では, 異種特徴間の相関を最大化する射影に基づく特徴を用いた変状分類手法を提案した. 実験では, この射影を用いずに多数の隠れ層により特徴変換を行う深層学習手法と比較して, 本手法の精度が高いことが確認できた. したがって, 多数の隠れ層による特徴変換よりも, 正準相関の考慮により得られる射影を用いた特徴変換を用いることが有効であると確認できる.



(a) ひびわれ　(b) ひびわれ



(c) 遊離石灰　(d) 遊離石灰



(e) 遊離石灰　(f) 腐食



(g) 剥落　(h) 鉄筋露出



(i) 鉄筋露出　(j) 鉄筋露出

図 4.4: 提案手法によって正しく正解された変状画像の例.



(a)　(b)

図 4.5: 異なるアングルから撮影された腐食画像の例.

第5章　異種特徴間の関連性を考慮可能な深層学習に基づく変状分類

5.1　はじめに

本章では, 異種特徴の相関関係を考慮可能な順伝搬型の深層学習を構築することで, 高精度な変状分類を実現する手法を提案する. 具体的に, 4 章で得られた射影を隠れ層に導入した順伝搬型の深層学習として新たに Correlation-maximizing Deep Extreme Learning Machine-local Receptive Field (以降, CMDELM-LRF) を構築する. 順伝搬型の深層学習の隠れ層の重みとして異種特徴の相関関係を考慮して算出した射影を用いることで, 少ない層で分類に有効な特徴への変換を可能とすることから, 最適に決定すべきパラメータ数の削減が可能となる. 以上により, 異種特徴間の関連性に基づいて, 高い弁別性を有する特徴への変換の枠組みを導入した深層学習に基づく, 高精度な変状分類を実現する. 本手法は, 図 5.1 に示す通り, 3 つのステップにより実現する. 一つ目のステップとして Local Receptive Field (以降, LRF) [62] に基づく画像特徴の算出, 二つ目のステップとして正準相関最大化を行う層の構築, 最後に順伝搬型の深層学習の構築を行う. これらの内容を以下の章に分けて説明する. まず, 5.2 節で, LRF に基づく画像特徴とテキスト特徴の算出について説明し, 5.3 節で, 異種特徴間の相関関係に基づく射影の導出を行う. つづいて, 5.4 節で, 特徴変換された特徴を用いた順伝搬型の深層学習の構築について説明し, 5.5 節で, 実験を行い, 最後に 5.6 節で, まとめとする.

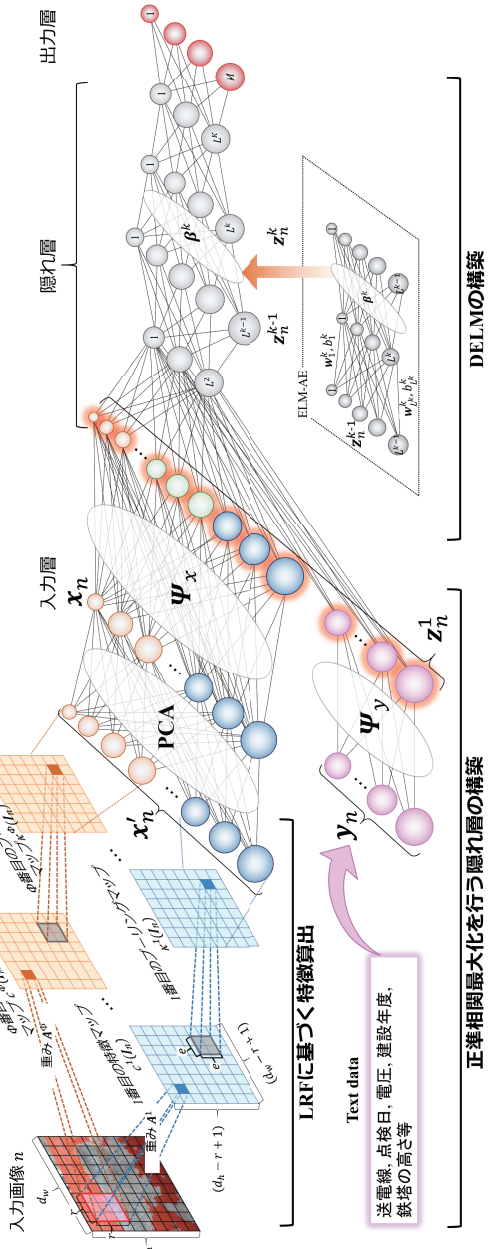


図 5.1: 本手法の概要図を示す. まず, LRF を用いて画像特徴を算出する. 次に, 画像特徴およびテキスト特徴間の正準相関を最大化する層を構築する. 最後に, 順伝搬型の深層学習である Deep Extreme Learning Machine (以降, DELM) を構築する.

5.2　LRF に基づく特徴抽出

学習画像 n ($n = 1, 2, ..., N$; N は学習画像の数) が与えられた時, 画像 n に対応する入力行列 $\boldsymbol{I}_n \in \mathbb{R}^{d_h \times d_w}$ から画像特徴を算出する. カラーチャネルを考慮するために, 文献 [62] と同じ処理を行う. 図 5.1 に示す通り, 画像特徴を算出するためには, 提案手法は特徴マップ構築とプーリングマップの構築の二つのステップを行う.

初めに, 重み行列 $\hat{\boldsymbol{A}} \in \mathbb{R}^{r^2 \times \Phi}$ の初期値をランダムに設定する. ただし, $r \times r$ は局所受容野のサイズを表し, Φ は特徴マップの数を表す. 特異値分解 (singular value decomposition: SVD) を用いることで, 重み行列の初期値 $\hat{\boldsymbol{A}}$ を直交化し, 直交ベクトル $\hat{\boldsymbol{a}}^{\phi} \in \mathbb{R}^{r^2}$ $(\phi = 1, 2, ..., \Phi)$ を算出する. ただし, $\hat{\boldsymbol{A}} = [\hat{\boldsymbol{a}}^1, \hat{\boldsymbol{a}}^2, ..., \hat{\boldsymbol{a}}^{\Phi}]$ である. $r^2 < \Phi$ の時, SVD により, $(\hat{\boldsymbol{A}})^{\top}$ を直交化する. そして, 文献 [62] でのプロセスと同様に, その行列の転置をとる. したがって, ϕ 番目の特徴マップごとに列方向に $\hat{\boldsymbol{a}}^{\phi}$ を持つ, 入力重み行列 $\boldsymbol{A}^{\phi} \in \mathbb{R}^{r \times r}$ を算出する. 算出した入力重み行列 $\boldsymbol{A}^{\phi}$ を用いることによって, ϕ 番目の特徴マップ $c^{\phi}(\boldsymbol{I}_n)$ は以下の式に基づき算出される.

$$
\begin{aligned}
c^{\phi}_{i,j}(\boldsymbol{I}_n) &= \sum_{s=1}^{r}\sum_{u=1}^{r} \boldsymbol{I}_n^{(i+r-s, j+r-u)} \times \boldsymbol{A}^{\phi}_{s,u} \\
i &= 1,\ 2,\ ...\ , (d_h - r + 1) \\
j &= 1,\ 2,\ ...\ , (d_w - r + 1)
\end{aligned}
\tag{5.1}
$$

ただし, $c^{\phi}_{i,j}(\boldsymbol{I}_n)$, $\boldsymbol{I}_n^{(i+r-s, j+r-u)}$ と $\boldsymbol{A}^{\phi}_{s,u}$ は, $c^{\phi}(\boldsymbol{I}_n)$ の (i, j) 番目の要素, $\boldsymbol{I}_n$ の $(i + r - s, j + r - u)$ 番目の要素, $\boldsymbol{A}^{\phi}$ の (s, u) 番目の要素を表す. 結果的に, サイズが $(d_h - r + 1) \times (d_w - r + 1)$ である特徴マップ $c^{\phi}(\boldsymbol{I}_n)$ を算出する.

第二に, 算出された特徴マップを用いることでプーリングマップを算出する. 提案手法で用いるプーリングサイズは $e \times e$ であり, プーリングマップのサイズは特徴マップのサイズである $(d_h - r + 1) \times (d_w - r + 1)$ と同じである. 平方/平方根プーリングを施すことで, ϕ 番目の特徴マップ $\kappa^{\phi}(\boldsymbol{I}_n)$ は以下の式により算出さ

表 5.1: テキストデータの一例およびその次元数を示す. 実際には, 送電線には名前が記録されており, 塩害区分は塩害の影響を受ける地域を表す.

項目	内容	次元数
送電線	C 線, D 線, ...	D_1
鉄塔種別	E 鉄塔, F 鉄塔, ...	D_2
塩害区分	A', B', C', ...	D_3
地域区分	G 地域, H 地域, ...	D_4
診断年月日	2015/03/10, 2014/03/03, ...	1
電圧 (kV)	66, 275, 154, ...	1
建設年度	1966, 1977, 1983, ...	1
鉄塔高さ (m)	56.4, 73.5, 72, ...	1
塗装年度	2004, 2001, ...	1
合計	-	$d_y = \sum_{t=1}^{4} D_t + 5$

れる.

$$
\begin{aligned}
\kappa^{\phi}_{p,q}(\boldsymbol{I}_n) &= \sqrt{\sum_{i=p-e}^{p+e}\sum_{j=q-e}^{q+e}\left\{c^{\phi}_{i,j}(\boldsymbol{I}_n)\right\}^2} \\
p &= 1,\ 2,\ \ldots,(d_h - r + 1) \\
q &= 1,\ 2,\ \ldots,(d_w - r + 1)
\end{aligned}
\tag{5.2}
$$

ただし, $\kappa^{\phi}_{p,q}(\boldsymbol{I}_n)$ は $\kappa^{\phi}(\boldsymbol{I}_n)$ のの (p, q) 番目の要素を表す. $c^{\phi}_{i,j}(\boldsymbol{I}_n)$ はゼロパディングされた特徴マップである. 平方/平方根プーリングは文献 [63] においても用いられており, その有効性が示されている. プーリングマップのそれぞれのピクセル値を並べることによって特徴ベクトル $\boldsymbol{x}'_n \in \mathbb{R}^{(d_h-r+1)\dot{(}d_w-r+1)\dot{\Phi}}$ を算出する. 最後に, 画像特徴ベクトルの次元 $\boldsymbol{x}'_n$ が学習画像の数より多い場合, 次節の処理である CCA の学習において過学習が発生する可能性があるため, Principal Component Analysis (PCA) [64] に基づき次元削減を行い, 特徴ベクトル $\boldsymbol{x}_n \in \mathbb{R}^{d_x}$ を算出する.



図 5.2: テキスト特徴の算出方法の一例. 図における画像 ID は各画像を表し, それぞれにテキストデータが対応している.

5.3　正準相関を最大化する隠れ層の構築

まず, テキストデータからテキスト特徴を算出する. テキストデータの例を表 5.1 に示す. 技術者は, 送電線, 鉄塔種別, 塩害区分, 地域区分, 診断年月日, 電圧, 建設年度, 鉄塔高さおよび塗装年度をテキストデータとして記録する. そこで, 本手法では, これらのテキストデータからテキスト特徴を算出する. 算出方法の具体例

を図 5.2 に示す. 図 5.2 の赤枠に示す通り, 送電線, 鉄塔種別, 塩害区分, 地域区分に関しては, 対応する次元に 1, それ以外を 0 とした特徴として算出する. 一方で, それ以外のテキストデータについては, 図 5.2 の青枠に示す通り, 対応する値を特徴として算出する. 最終的なテキスト特徴は, これらを一列に並べることで算出される. つまり, 表 5.1 に示すように, 送電線, 鉄塔種別, 塩害区分, 地域区分から $\sum_{t=1}^{4} D_t$ 次元の特徴を算出し, 他の 5 つのデータから 5 次元の特徴を算出することで, テキスト特徴ベクトル $\boldsymbol{y}_n \in \mathbb{R}^{d_y}$ $(d_y = \sum_{t=1}^{4} D_t + 5)$ を算出する.

次に, 画像特徴とテキスト特徴間の正準相関を最大化する隠れ層を構築する. これらの特徴間の関連性を考慮可能な隠れ層の重み行列を計算するために, 画像特徴行列 $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N] \in \mathbb{R}^{d_x \times N}$ とテキスト特徴行列 $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N] \in \mathbb{R}^{d_y \times N}$ に正準相関分析 [21] を適用することで射影行列 $\boldsymbol{\Psi}_x$ と $\boldsymbol{\Psi}_y$ を算出する. 具体的に, 以下の目的関数の最大化を行う.

$$(\hat{\boldsymbol{\psi}}_x, \hat{\boldsymbol{\psi}}_y) = \arg\max_{\boldsymbol{\psi}_x, \boldsymbol{\psi}_y} \frac{\boldsymbol{\psi}_x^{\mathrm{T}} \boldsymbol{C}_{XY} \boldsymbol{\psi}_y}{\sqrt{\boldsymbol{\psi}_x^{\mathrm{T}} \boldsymbol{C}_{XX} \boldsymbol{\psi}_x} \sqrt{\boldsymbol{\psi}_y^{\mathrm{T}} \boldsymbol{C}_{YY} \boldsymbol{\psi}_y}} \tag{5.3}$$

ただし, $\boldsymbol{\psi}_x$ と $\boldsymbol{\psi}_y$ は射影ベクトルである. また,

$$\boldsymbol{C}_{XY} = \boldsymbol{X}\boldsymbol{Y}^{\top} \tag{5.4}$$

$$\boldsymbol{C}_{XX} = \boldsymbol{X}\boldsymbol{X}^{\top} \tag{5.5}$$

$$\boldsymbol{C}_{YY} = \boldsymbol{Y}\boldsymbol{Y}^{\top} \tag{5.6}$$

である. 式 (5.3) を最大化するために, ラグランジュの未定乗数法により, 以下の式を解く.

$$\begin{aligned} \mathcal{L}(\boldsymbol{\psi}_x, \boldsymbol{\psi}_y) = \boldsymbol{\psi}_x^{\top} C_{XX} \boldsymbol{\psi}_y - \frac{\lambda_x}{2}(\boldsymbol{\psi}_x^{\top} C_{XX} \boldsymbol{\psi}_x - 1) \\ - \frac{\lambda_y}{2}(\boldsymbol{\psi}_y^{\top} C_{YY} \boldsymbol{\psi}_y - 1) \end{aligned} \tag{5.7}$$

ただし, $\lambda_x = \lambda_y\ (= \lambda)$ であり, λ は以下に示す. そして, 以下の固有値問題を解く.

$$C_{XX}^{-1} C_{XY} C_{YY}^{-1} C_{XY}^{\top} \boldsymbol{\psi}_x = \lambda^2 \boldsymbol{\psi}_x \tag{5.8}$$

$$C_{YY}^{-1} C_{XY}^{\top} C_{XX}^{-1} C_{XY} \boldsymbol{\psi}_y = \lambda^2 \boldsymbol{\psi}_y \tag{5.9}$$

ただし, λ は上記問題における固有値を表す. 上式を解くことで, 複数の固有値 λ とそれに対応する固有ベクトル $\boldsymbol{\psi}_x$ および $\boldsymbol{\psi}_y$ を算出されるため, d_{cca} 次元の固有ベクトルを並べることで, 射影行列 $\boldsymbol{\Psi}_x = [\boldsymbol{\psi}_x^1, \boldsymbol{\psi}_x^2, ..., \boldsymbol{\psi}_x^{d_{cca}}]^{\top} \in \mathbb{R}^{d_{cca} \times d_x}$ と $\boldsymbol{\Psi}_y = [\boldsymbol{\psi}_y^1, \boldsymbol{\psi}_y^2, ..., \boldsymbol{\psi}_y^{d_{cca}}]^{\top} \in \mathbb{R}^{d_{cca} \times d_y}$ を算出する. 図 5.1 に示すように, LRF と DELM の間に射影行列を挿入することで, 以下の式で定義される射影特徴 $\hat{\boldsymbol{x}}_n \in \mathbb{R}^{d_{cca}}$ と $\hat{\boldsymbol{y}}_n \in \mathbb{R}^{d_{cca}}$ を算出する.

$$\hat{\boldsymbol{x}}_n = \boldsymbol{\Psi}_x \boldsymbol{x}_n \tag{5.10}$$

$$\hat{\boldsymbol{y}}_n = \boldsymbol{\Psi}_y \boldsymbol{y}_n \tag{5.11}$$

多くの深層学習では, 多くの隠れ層を使用することによって画像から得られた特徴を意味的な特徴への変換を実現している. これらの深層学習では, 設定すべきパラメータの数が多いので, オーバーフィッティングを避けるために, 大量の学習データを用意する必要がある. 一方, 画像特徴を意味的な特徴への直接変換は, 第 4 章において, 画像特徴とテキスト特徴間の正準相関を考慮することで実現可能であることが示されている. したがって, 最適に決定すべきパラメータの減少可能であるため, 少量の学習データを用いた深層学習を構築することが可能となる.

提案手法では, 画像特徴とテキスト特徴を使用する. 変状画像は背景などの不要な領域を含むので, 画像特徴は雑音となる情報を含むことになる.　一方, テキストの特徴は技術者が記録した実際の維持管理業務に基づいてデータから算出されるため, 質の高い特徴となる. 画像特徴とテキスト特徴を使用して表現能力の高い特徴を取得するためには, これらの異種特徴を互いの相関が高くなるような空間に投影する必要がある.　したがって, 我々は CCA に焦点を当てる. CCA は

2 つの異なる特徴を比較可能な特徴空間に射影し, それらの相関最大化を可能にする最も一般的な方法の 1 つである. さらに, CCA で得られた正準相関を最大にする射影行列を深層学習の重みとして用いることができるため, 特徴変換と分類の両方の end-to-end な学習が可能となる.

5.4　DELM の構築

本節では, DELM による変状分類について説明する. まず, DELM のモデルを図 5.1 に示す. DELM は, 入力層 (1 層), 隠れ層 (K 層), 出力層 (1 層) の計 $K+2$ 層で構成されており, 第 k 層と第 $k-1$ 層間の重み行列 (ただし, $k = 2, 3, ..., K+1$) を逐次算出することで学習が行われる ($k = 1$ の時は入力層である) . 具体的に, **(I)** $k = 2, 3..., K$ までの重み行列は, Cambria らにより提案された Extreme Learning Machine-Auto Encoder (以降, ELM-AE) [65] を各層ごとに構築することで算出される. 一方, **(II)** $k = K+1$ における重み行列は, 通常の Extreme Learning Machine (以降, ELM) [59] と同様に求められる. したがって, 以降では, 上述した重み行列の算出方法について説明する.

(I) $k = 2, 3, ..., K$

DELM では, k 番目の隠れ層の出力行列 $\boldsymbol{Z}^k = [\boldsymbol{z}_1^k, \boldsymbol{z}_2^k, ..., \boldsymbol{z}_N^k]$ と $(k-1)$ 番目の隠れ層の出力行列 $\boldsymbol{Z}^{k-1}$ 間の関係は以下の式で表される.

$$\boldsymbol{z}_n^k = G(\boldsymbol{\beta}^k \boldsymbol{z}_n^{k-1}) \tag{5.12}$$

ただし, $\boldsymbol{Z}^k$ は学習データから得られる行列であり, $\boldsymbol{\beta}^k \in \mathbb{R}^{L^k \times L^{k-1}}$ は, k 番目と $(k-1)$ 番目の隠れ層間の重み行列である. さらに, G は活性化関数として用いたシグモイド関数を表す.

重み行列 $\boldsymbol{\beta}^k$ を計算するために, 各層ごとに ELM-AE を構築する. 入力層,　隠れ層および出力層の三層で構成される ELM-AE モデルは図 5.1 の右下に示される. これらの層は L^{k-1} 個の入力ノード, L^k 個の隠れノード, そして L^{k-1} 個の出

力ノードをもつ. ELM-AE はオートエンコーダであるため, 入力特徴が出力特徴と等しくなるように隠れ層を学習する. 入力ベクトル $\boldsymbol{z}_n^{k-1} \in \mathbb{R}^{L^{k-1}}$ が与えられたとき, ELM-AE の隠れ層の出力 $\boldsymbol{h}_n^k \in \mathbb{R}^{L^k}$ は以下の式で得られる.

$$\boldsymbol{h}_n^k = G(\boldsymbol{W}^k \boldsymbol{z}_n^{k-1} + \boldsymbol{b}^k) \tag{5.13}$$

$$\boldsymbol{W}^k (\boldsymbol{W}^k)^\top = \boldsymbol{I} \tag{5.14}$$

$$(\boldsymbol{b}^k)^\top \boldsymbol{b}^k = 1 \tag{5.15}$$

ELM-AE はランダム値に基づく直交重み $\boldsymbol{W}^k = [\boldsymbol{w}_1^k, \boldsymbol{w}_2^k, ..., \boldsymbol{w}_{L^k}^k]^\top$ とランダムバイアス $\boldsymbol{b}^k = [b_1^k, b_2^k, ..., b_{L^k}^k]^\top$ をもつ. 以上より, ELM-AE への入力行列 $\boldsymbol{Z}^{k-1}$ と ELM-AE の隠れ層における出力行列 $\boldsymbol{H}^k = [\boldsymbol{h}_1^k, \boldsymbol{h}_2^k, ..., \boldsymbol{h}_N^k]^\top \in \mathbb{R}^{N \times L^k}$ を用いることで, 出力重み $\boldsymbol{\beta}^k$ が以下の式で算出される.

$$\boldsymbol{\beta}^k = \left(\frac{\boldsymbol{I}}{C_1} \sum_{l^k=1}^{L^k} \mathrm{KL}(\rho || \hat{\rho_{l^k}}) + (\boldsymbol{H}^k)^\top \boldsymbol{H}^k \right)^{-1} (\boldsymbol{H}^k)^\top (\boldsymbol{Z}^{k-1})^\top \tag{5.16}$$

ただし, $\mathrm{KL}(\rho || \hat{\rho}_{l^k}) = \rho \log \frac{\rho}{\hat{\rho}_{l^k}} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_{l^k}}$ は KL ダイバージェンスを表し, C_1 は正則化パラメータを表す. 隠れ層の表現をスパースとする制約を設けることで, オートエンコーダは高い表現能力を持つ特徴の算出が実現すると報告されている [66]. KL ダイバージェンスは, しばしばオートエンコーダのスパース表現のための制約として用いられる. パラメータ $\hat{\rho}_{l^k}$ は k 番目の隠れ層における l^k 番目の隠れノードにおける全学習データの平均活性値を表す. 予め設定したスパース性を決定するパラメータ ρ に基づき, KL ダイバージェンスにおける平均活性値 $\hat{\rho}_{l^k}$ を小さくするように隠れ層が構築される. KL ダイバージェンスは隠れ層のノードを除去するのではなく,　ノードの活性化を制御するため, 入力画像に応じて活性化されるノードが変化する. どのノードが重要であるかを抽出できるため, 表現能力が向上する [66].

結果的に, DELM におけるそれぞれの層の出力重み $\boldsymbol{\beta}^k$ $(k = 1, 2, ..., K)$ は ELM-AE の出力重み $\boldsymbol{\beta}^k$ を用いることで算出可能となる. 一般的な深層学習手法では,

誤差逆伝搬法により膨大な数のパラメータを設定するために, 大量の学習画像と膨大な計算時間を必要とするが, 提案手法では各層ごとに教師無し学習である ELM-AE を構築することでパラメータを決定する. 加えて, パラメータ $\boldsymbol{W}^k$ と $\boldsymbol{b}^k$ はランダム値に基づき算出される. したがって, 提案手法では, 少量の学習データによる学習が可能となり, 計算コストの削減も可能となる.

(II) $k = K + 1$

K 番目の隠れ層と出力層間の重み行列 $\boldsymbol{\beta}^{K+1}$ は教師あり学習である通常の ELM と同様に求めることが可能である具体的には, 学習誤差 $\boldsymbol{\xi}_n = [\xi_{n,1}, \xi_{n,2}, ..., \xi_{n,M}]^\top$ (M は M クラス分類問題における出力ノード数を表す) と重み行列の大きさをを最小化するように最適化を行う.

$$\begin{aligned} \min_{\boldsymbol{\beta}^{K+1} \in \mathbb{R}^{M \times L^{K+1}}} R &= \frac{1}{2}\|\boldsymbol{\beta}^{K+1}\|_F^2 + \frac{C_2}{2}\sum_{n=1}^{N}\|\boldsymbol{\xi}_n\|^2 \\ \text{s.t. } \boldsymbol{\beta}^{K+1}\boldsymbol{z}_n^{K+1} &= \boldsymbol{t}_n - \boldsymbol{\xi}_n \end{aligned} \tag{5.17}$$

ただし, $\boldsymbol{t}_n = [t_{n,1}, t_{n,2}, ..., t_{n,M}]^\top$ は, 真のクラスが m の時に m 番目の要素が 1, 他の要素が 0 となるベクトルである. さらに, C_2 は正則化パラメータである. 文献 [57] によると, 式 (5.17) は Karush-Kuhn-Tucker 条件に基づき, 以下の式を解くことと等価となる.

$$\begin{aligned} \min_{\boldsymbol{\beta}^{K+1} \in \mathbb{R}^{M \times L^{K+1}}} \tilde{R} = &\frac{1}{2}\|\boldsymbol{\beta}^{K+1}\|_F^2 + \frac{C_2}{2}\sum_{n=1}^{N}\|\boldsymbol{\xi}_n\|^2 \\ &- \sum_{n=1}^{N}\sum_{m=1}^{M}\alpha_{n,m}\left\{(\boldsymbol{\gamma}_m^{K+1})^\top \boldsymbol{z}_n^{K+1} - t_{n,m} + \xi_{n,m}\right\} \end{aligned} \tag{5.18}$$

ただし, $\boldsymbol{\gamma}_m^{K+1} = [\boldsymbol{\beta}_{1,m}^{K+1}, \boldsymbol{\beta}_{2,m}^{K+1}, ..., \boldsymbol{\beta}_{L^{K+1},m}^{K+1}]^\top$ は m 番目の出力ノードと隠れ層をつなげる重みベクトルである. $\boldsymbol{\gamma}_m^{K+1}$, $\boldsymbol{\xi}_n$ および $\boldsymbol{\alpha}_n$ を用いることで, 最適な重み $\boldsymbol{\beta}^{K+1}$ は下式により得られる. ただし, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_N]^\top$ であり, $\boldsymbol{\alpha}_n =$

$[\alpha_{n,1}, \alpha_{n,2}, ..., \alpha_{n,M}]^\top$ である.

$$\boldsymbol{\beta}^{K+1} = \boldsymbol{T}(\boldsymbol{Z}^{K+1})^\top \left(\frac{\boldsymbol{I}}{C_2} + \boldsymbol{Z}^{K+1}(\boldsymbol{Z}^{K+1})^\top \right)^{-1} \tag{5.19}$$

ただし, $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, ..., \boldsymbol{t}_N]$ である. 結果的に, 重み行列 $\boldsymbol{\beta}^{K+1}$ の算出が実現する.

テストデータ $\boldsymbol{z}^{K+1}$ が与えられたとき, $\boldsymbol{\beta}^{K+1}$ を用いることで, 出力値 $\boldsymbol{v} = [v_1,\ v_2,\ ...,\ v_M]^\top$ は下式により得られる.

$$\boldsymbol{v} = \boldsymbol{\beta}^{K+1}\boldsymbol{z}^{K+1} \tag{5.20}$$

さらに, 最終結果は $class = \arg\max_{m \in \{1,...,M\}} v_m$ として算出される.

結果的に, 画像特徴とテキスト特徴間の正準相関を考慮可能な CMDELM-LRF により高精度な変状分類が実現する.

5.5　実験

本節では, 提案手法の有効性を確認するための実験を行う. 以降, 5.5.1 において実験準備を説明し, 5.5.2 において実験結果と考察を示す.

5.5.1　実験準備

提案手法の有効性を検証するために, 東京電力ホールディングス株式会社からご提供頂いたデータを用いる. このデータセットには, 送電鉄塔における実際の維持管理業務において技術者が撮影した変状画像を記録したテキストデータが含まれている. データセットには, クラス A, クラス B, クラス C の 3 つのレベルがある. ただし, クラス C は腐食が著しいものである. 本実験で用いる画像数の詳細は, 表 5.2 に示す. 実験では, 層数 K, 正則化パラメータ C_1 と C_2, および PCA を使用して選択された次元数である次元 d_x は, 検証データセットにおいて最良のパフォーマンスを出力した際の値を採用した. ELM に基づく手法では, シグモイド

表 5.2: 本実験で用いる画像数.

	画像数
クラス A	1,044
クラス B	1,351
クラス C	748
合計	3,107

関数のパラメータはランダムな値に設定される. パラメータの詳細は, 表 5.3 に示す. 検証方法は, 5 分割交差検定を用いた. 以下に示される, Recall, Precision および F 値を用いることで提案手法の有効性を検証する.

$$\text{Recall} = \frac{\text{正しく分類されたサンプルの数}}{\text{そのクラスのサンプル数}}$$

$$\text{Precision} = \frac{\text{正しく分類されたサンプルの数}}{\text{手法によりそのクラスに分類されたサンプル数}}$$

$$\text{F 値} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

表 5.3: 提案手法で用いるパラメータの詳細.

詳細	パラメータ	値
入力画像サイズ	d_w	50
	d_h	50
LRF の窓サイズ	r	5
LRF のプーリングサイズ	e	3
次元数	d_x	471
	d_y	364
正則化パラメータ	C_1	2^{15}
	C_2	2^{15}
スパース性を決めるパラメータ	ρ	0.05
LRF で算出するチャネル数	$\mathbf{\Phi}$	35
隠れ層の数	K	4
隠れ層のノード数	L^1	394
	L^2	591
	L^3	472
	L^4	377
	L^5	301
クラス数	M	3

表 5.4: 本実験で用いる手法の詳細.

		元の特徴		射影後の特徴	
手法	詳細	画像	テキスト	画像	テキスト
提案手法	CMDELM-LRF	-	-	✓	✓
比較手法 1	CMDELM-LRF	-	-	✓	-
比較手法 2	CMDELM-LRF	-	-	-	✓
比較手法 3	DELM-LRF [67]	✓	-	-	-
比較手法 4	DELM-LRF	-	✓	-	-
比較手法 5	DELM-LRF	✓	✓	-	-
比較手法 6	CCA + ELM-LRF	-	-	✓	✓
比較手法 7	CCA + ELM-LRF	-	-	✓	-
比較手法 8	CCA + ELM-LRF	-	-	-	✓
比較手法 9	ELM-LRF [62]	✓	-	-	-
比較手法 10	DELM [65]	✓	-	-	-
比較手法 11	KELM [68]	✓	-	-	-
比較手法 12	ELM [59]	✓	-	-	-
比較手法 13	SVM [23]	✓	-	-	-
比較手法 14	CaffeNet-CNN [69]	✓	-	-	-
比較手法 15	VGG16-CNN [24]	✓	-	-	-
比較手法 16	多層パーセプトロン	-	-	✓	✓

表 5.5: 本実験で用いる手法のネットワーク構成を示す．“conv”, “norm”, “pool”は, 畳み込み層, 局所応答正規化層, プーリング層を表し, “pca”, “cca”は, PCA, CCA を表す．“fc-elm”, “fc”は ELM に基づく全結合層, 一般的な全結合層を表す．

	提案と比較 1, 2	比較手法 3-5	比較手法 6-8	比較手法 9	比較手法 14	比較手法 15	比較手法 16
	conv	conv	conv	conv	conv	conv×2	conv
	pool	pool	pool	pool	pool	pool	pool
	pca	pca	pca	pca	norm	conv×2	pca
	cca	fc-elm×4	cca	fc-elm×1	conv	pool	fc×4
	fc-elm×4		fc-elm×1		pool	conv×3	
					norm	pool	
					conv ×3	conv×3	
					pool	pool	
					fc×3	conv×3	
						pool	
						fc×3	
層数	8	7	5	4	13	21	7

また, この実験では, 表 5.4 に示すように, 16 の比較方法を使用した. DELM-LRF は我々が, 以前に報告した [67] の方法である. その方法は, 変状画像およびテキスト特徴を直接深層学習に入力することで構築されている. DELM, KELM, ELM, SVM は, Caffe-Net [69] から抽出された画像特徴に基づいて構築される. さらに, 提案手法の有効性を検証するために, CNN に基づく方法と比較した. 送電鉄塔では, 十分な数の学習画像を作成することが難しいため, 比較方法として ImageNet で学習済の CNN に対し, 本手法で用いるデータで Fine-tuning を行う手法を採用する. 具体的には, Caffe-Net モデル [69] と VGG16 モデル [24] を採用した. 特に, VGG16 モデルは一般的で強力な深層学習方法の 1 つである. さらに、我々は, 最も一般的な深層学習手法の一つである多層パーセプトロンに基づく深層学習を使用した. さらに, 提案手法および比較手法の隠れ層の詳細は, 表 5.5 に示されている. この表から, CMDELM-LRF は, 深層学習でベンチマークな比較手法 14, 15 と比較して, 隠れ層が少ないことが確認できる.

表 5.6: 全ての手法における Recall, Precision および F 値.

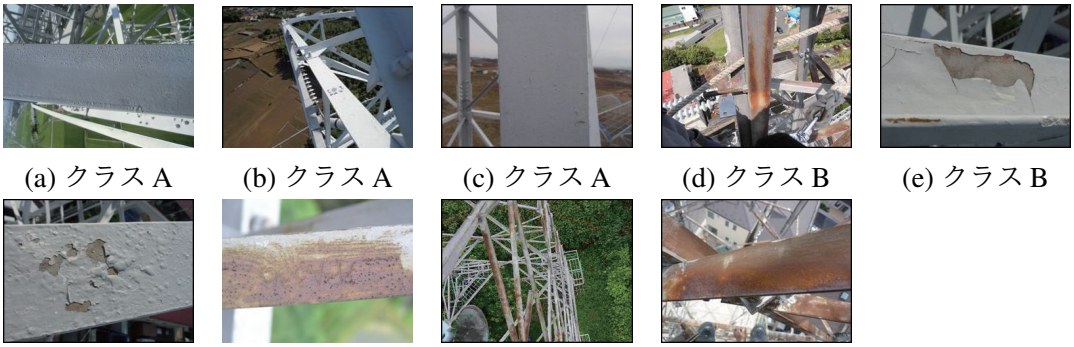
	クラス A			クラス B			クラス C			平均		
手法	R	P	F	R	P	F	R	P	F	R	P	F
提案手法	0.985	0.902	**0.939**	0.799	**0.956**	**0.863**	**0.922**	**0.819**	**0.863**	**0.902**	**0.892**	**0.888**
比較手法 1	0.978	0.881	0.923	0.751	0.937	0.826	0.900	0.780	0.834	0.877	0.866	0.861
比較手法 2	0.496	0.511	0.498	0.424	0.581	0.488	0.595	0.403	0.479	0.505	0.499	0.488
比較手法 3	**0.987**	0.890	0.933	0.593	0.952	0.721	0.912	0.621	0.733	0.831	0.821	0.796
比較手法 4	0.765	0.594	0.668	0.512	0.792	0.620	0.703	0.591	0.637	0.660	0.659	0.642
比較手法 5	0.851	0.647	0.735	0.609	0.854	0.708	0.763	0.734	0.742	0.741	0.745	0.728
比較手法 6	0.970	**0.913**	0.937	0.783	0.931	0.840	0.896	0.793	0.830	0.883	0.879	0.869
比較手法 7	0.959	**0.913**	0.931	0.810	0.920	0.847	0.870	0.805	0.823	0.880	0.879	0.867
比較手法 8	0.561	0.700	0.619	0.620	0.714	0.662	0.770	0.519	0.620	0.651	0.644	0.634
比較手法 9	0.880	0.856	0.862	0.688	0.857	0.745	0.814	0.651	0.714	0.794	0.788	0.774
比較手法 10	0.498	0.427	0.460	0.331	0.522	0.405	0.511	0.362	0.424	0.447	0.437	0.430
比較手法 11	0.464	0.401	0.430	0.347	0.463	0.397	0.430	0.354	0.388	0.414	0.406	0.405
比較手法 12	0.485	0.414	0.446	0.321	0.476	0.383	0.448	0.337	0.384	0.418	0.409	0.405
比較手法 13	0.127	0.546	0.204	**0.909**	0.432	0.585	0.034	0.288	0.060	0.357	0.422	0.283
比較手法 14	0.517	0.362	0.426	0.363	0.464	0.407	0.163	0.207	0.182	0.347	0.344	0.338
比較手法 15	0.835	0.707	0.765	0.611	0.763	0.679	0.762	0.695	0.727	0.736	0.722	0.724
比較手法 16	0.944	0.818	0.872	0.592	0.896	0.708	0.920	0.693	0.788	0.819	0.802	0.789

5.5.2　実験結果および考察

Recall, Precision, F 値は, 表 5.6 に示されている. 表 5.6 により, 提案手法が比較手法 5 よりも優れており, 比較手法 7 が比較手法 9 よりも優れていることが確認できる. したがって, 画像特徴とテキスト特徴との間の正準相関を最大にする CCA による特徴変換の有効性が示された. 加えて, 提案手法は比較手法 1 および 2 に勝ることから, 射影された特徴が本分類問題に有効であることがわかる. 提案手法と比較手法 6 の精度が比較手法 9 と 11-13 を含む他の手法の精度よりも高い. したがって, 深層学習に基づく手法の有効性が示された. 提案手法が比較手法 6 を上回ることから, 隠れ層を追加することが精度向上に貢献していることがわかる. 以上により, CCA による特徴変換は, 深層学習手法である DELM の隠れ層による特徴変換よりも有効であることが確認できる. さらに, CMDELM-LRF と比較手法 14-16 を比較することで, 多数の隠れ層を持つ他の深層学習手法よりも優れていることが確認される. 特に, 表 5.5 に示す通り, 比較手法 15 は提案手法と比較して多くの隠れ層を持つが, 提案手法はその比較手法よりも高いことがわかる. したがって, CCA に基づく特徴変換の有効性が確認された.

表 5.6 より, 比較手法 6 と 7 がクラス A において有効であることがわかる. 比較手法 6 と 7 は ELM-LRF と CCA を組み合わせた手法である. 本実験では, クラス A に属する画像は変状を含まない画像であるが, クラス B および C の画像は変状を含んでいる. そこで, クラス B やクラス C のような隣接するクラスを分類するためには, その小さな差を判別できる表現能力の高い特徴量を算出する必要がある. しかしながら, 比較手法 6 と 7 は浅いニューラルネットワークであり, CMDELM-LRF のような深層学習に比べて表現能力が低い. これにより, クラス B およびクラス C の分類性能が低下する. さらに, この実験では, F 値の平均値が高くなるように隠れ層のパラメータが決定されるため, 分類困難なクラス B とクラス C の精度よりも, クラス A と他のクラスとを区別するように調整される. したがって, 比較手法 6 および 7 はクラス A で高い性能を達成すると考えられる.

提案手法では, ランダム値に基づく $\boldsymbol{\beta}^k$ を用いるため, 分類精度を比較することで間接的に $\boldsymbol{\beta}^k$ の正しさを評価する. 具体的に, CMDELM-LRF と誤差逆伝搬法



(a) クラス A　(b) クラス A　(c) クラス A　(d) クラス B　(e) クラス B

(f) クラス B　(g) クラス C　(h) クラス C　(i) クラス C

図 5.3: 提案手法によって正しく推定された画像の例. ただし, 本実験では, 画像中央から 50×50 $(= d_w \times d_h)$ ピクセルを用いている.

表 5.7: それぞれの手法における計算コストを示す.

処理	計算量	CMDELM-LRF	比較手法 14	比較手法 15
畳み込み	$\mathcal{O}(\Phi^{k-1} d_h^k d_w^k \Phi^k r^{k^2})$	✓	✓	✓
プーリング	$\mathcal{O}(\Phi^{k-1} d_h^k d_w^k e^{k^2})$	✓	✓	✓
PCA	$\mathcal{O}(d_{in}^2 N)$	✓	-	-
CCA	$\mathcal{O}(d_{in}^2 N)$	✓	-	-
正規化	$\mathcal{O}(\Phi^{k-1} d_h^k d_w^k)$	-	✓	-
ELM に基づく全結合層	$\mathcal{O}(L^{k^2} L^{k-1})$	✓	-	-
一般的な全結合層	$\mathcal{O}(d_{in} d_{out})$	-	✓	✓
誤差逆伝搬法	$\mathcal{O}(E_p N \sum_k^K L^{k-1} L^k)$	-	✓	✓

に基づき構築される多層パーセプトロンに基づく深層学習手法 (比較手法 16) と提案手法とを比較する. 誤差逆伝搬法は最適なパラメータ $\boldsymbol{\beta}^k$ を算出することが可能であるが, 最適解を得るためには, 大量の学習データが必要である. したがって, 本実験で用いられる学習データ数が少ないため, 分類精度向上に限界がある. 表 5.6 に示す通り, CMDELM-LRF の精度が比較手法 16 の精度よりも高いことが確認できる. つまり, $\boldsymbol{\beta}^k$ の算出のためには, 学習データ数が限られている状況において, ELM-AE が誤差逆伝搬法よりも有効であることを意味している.

提案手法により正しく推定された画像の例 [70] を図 5.3 に示す. 図 5.3 から, 送電鉄塔の画像には多くのバリエーションがあることがわかる. 図 5.3 (b), (d) およ

表 5.8: 提案手法, Caffe-Net および VGG16 の学習時間 (sec).

処理	提案手法	Caffe-Net	VGG16
特徴算出	5.01×10^{1}	-	-
正準相関最大化	2.62×10^{-1}	-	-
分類器の構築	4.49×10^{-1}	1.44×10^{4}	2.16×10^{3}
合計	5.08×10^{1}	1.44×10^{4}	2.16×10^{3}

表 5.9: 提案手法, Caffe-Net および VGG16 のテスト時間 (sec).

手法	計算時間
提案手法	$\mathbf{7.35 \times 10^{-1}}$
Caffe-Net	1.31
VGG16	2.06

び (h) は遠景画像であり, それ以外は近景画像を表す. また, 図 5.3 (a) と (c) に示すように, 対象物体の角度は異なる. このように, 提案手法は, 様々な種類の画像を正しく分類することができる. 一般的な分類手法 [6] では, そのようなバリエーションに対処するために, 画像中の対象物体が表す領域を手動で切り取ることで分類を行うため, これらは半自動的な手法である. 一方で, 提案手法は元画像を CMDELM-LRF に直接入力することができるため, 完全に自動分類可能な手法である. したがって, 実際のアプリケーションに用いることを考慮しても, 提案手法が有効であることが示された.

さらに, 提案手法と比較手法 14 と 15 の計算時間を比較する. 隠れ層の構成は表 5.5 に示されており, 各プロセスの計算コストを表 5.7 に示す. 表 5.7 において, k は各層のインデックスを表し, K は層の総数を表す. Φ^k は k 番目の層のフィルタ数を表す. Φ^{k-1} は $(k-1)$ 番目の層の入力チャネルの数を表す. d_h^k と d_w^k は出力する特徴マップのサイズである. r^k と e^k は畳み込みフィルタとプーリングフィルタのサイズをそれぞれ表す. d_{in} と d_{out} は入力特徴と出力特徴の次元数をそれぞれ表す.

さらに, L^k は k 番目の層におけるノード数を表す. N は画像数を表し, E_p はネットワークの学習エポック数を示す. この表から, 学習ステップにおいて, 比較手法 14 と 15 では, 学習のために多くの計算コストを必要とする誤差逆伝搬法を採用しているため, これらの計算コストは $\mathcal{O}(\sum_{k_c}^{K_c} \Phi^{k_c-1} d_h^{k_c} d_w^{k_c} \Phi^{k_c} {r^{k_c}}^2 + E_p N \sum_k^K L^k L^{k-1})$ である. ただし, K_c は畳み込み層の数を表し, k_c は畳み込み層のインデックスを表す. 表 5.7 に示すように, 畳み込みや誤差逆伝搬法は他のプロセスよりも複雑であるため, 計算オーダーの算出においては, 他のプロセスを無視することができる. 効果的な深層学習手法を学習するために, 比較手法 14 と 15 は多くのエポック数 E_p を必要とする. さらに, 表 5.5 に示す通り, 隠れ層の数は比較的多いので, これらの手法の計算コストは高くなる. 一方で, 提案手法の計算コストは, $\mathcal{O}(\sum_{k_c}^{K_c} \Phi^{k_c-1} d_h^{k_c} d_w^{k_c} \Phi^{k_c} {r^{k_c}}^2)$ であり, 畳み込み層が一層なので, $K_c = 1$ となる. そのため, 提案手法では, 入力画像のサイズ $d_h^{k_c} \times d_w^{k_c}$ とフィルタサイズ r^{k_c} は小さくなる. 以上より, 学習ステップにおいて, 提案手法は, 比較手法 14 と 15 よりも高速に学習可能であることが示された. また, 実測時間を表 5.8 に示す. 実際に, 表 5.8 に示すように, 提案手法の実測時間は, 比較手法 14 と 15 にような一般的な深層学習手法よりもはるかに低いことが確認できる. 具体的に, 提案手法, CaffeNet および VG16 の計算時間は, 5.08×10^1 秒, 1.44×10^4 秒および 2.16×10^3 秒である. 提案手法は CPU 演算を行っており, 動作させた計算機のスペックは, 3.40GHz Intel Core i7-3770 であり, メモリは 16GB である. CaffeNet-CNN は GPU 演算を行っており, 計算機のスペックは, 2.30GHz Intel Xeon E5-2699 v3 であり, メモリは 512GB である. また, GPU は Nvidia Tesla K80 である. この表より, DELM の学習コストは CPU 演算であるにもかかわらず高速であることがわかる. この理由として, ELM-AE が各層で構築される教師無し学習であり, ランダム値でパラメータが設定されることが挙げられる.

さらに, テストステップでは, 提案手法と比較手法 14 と 15 の計算コストは畳み込み層の数に依存するため, 計算コストの違いはわずかである. しかしながら, 提案手法は一つの畳み込み層で構成されており, 比較手法 14 と 15 は多くの畳み込み層から構成される. したがって, 表 5.9 に示す通り, 提案手法の実測時間は低い

ことから, 計算コストの削減を実現した. 以上の議論により, CMDELM-LRF が変状分類に有効であることが示された.

最後に, CCA によって得られた射影特徴が分類性能の向上に寄与する理由について考察する. 一般的な分類問題においては, ラベルと特徴間の相関関係が高くなればなるほど, 弁別性の高い特徴が算出可能であることが報告されている [71]. したがって, 予測性能と射影された特徴または元の特徴との間の関係を評価するために, 学習データにおける特徴とラベルとの間のピアソンの相関係数を計算した. 具体的には, クラスラベルと各特徴の各次元の間の相関係数を計算し, 図 5.4 と図 5.5 のようにヒストグラムを作成した. ただし, それぞれのヒストグラムのビンの値は対象特徴の次元数によって正規化されている. 例えば, 図 5.4(a) の場合, 元の画像特徴の次元 d_x は 471 であるため, 471 個の係数の頻度が表示される. 図 5.4 は元の特徴とラベルの間の相関係数のヒストグラムを示す. 図 5.4 (a), (b) および (c) は, 比較手法 3, 4, 5 の結果に対応する. 図 5.5 は, CCA により射影された特徴とラベルとの間の相関係数のヒストグラムを示す. 図 5.5 (a), (b) および (c) は, 提案手法, 比較手法 1 および 2 に対応している. これらの画像は, 横軸において ± 1 に近い値の数が多いほど, ラベルとの相関が高いことを意味する. さらに, これらの図の結果を定量的に比較するために, 相関係数 (Var) の分散と絶対係数値 (AveA) の合計の平均を計算した. これらの値が大きいほど, 特徴とラベルの間の相関が高くなることを意味する.

- 図 5.4 (a) と 図 5.5 (a) の比較

  “Var” と “AveA” の値とこれらの図に示される通り, 射影された画像特徴とラベルとの間の相関がより高いことを確認することができる. また, 比較手法 1 と 3 の実験結果から明らかなように, 射影された画像特徴は, 元の画像特徴よりも効果的であることがわかる.

- 図 5.4 (b) と図 5.5 (b) の比較

  一方で, 図 5.4 (b) と図 5.5 (b) に示す通り, 射影されたテキスト特徴とラベルとの間の相関が低いことが確認できる. さらに, 上述の関係と同様に, 射

影されたテキスト特徴を用いて構築された比較手法 2 もまた低いことが確認された.

上記の二つの点より, 各特徴間の相関関係と分類精度と間に関係が存在することが確認された.

- 図 5.4 (c) と図 5.5 (c) の比較

  図 5.4 (c) は図 5.4 (a) と図 5.4 (b) を単純に合成することで得られるヒストグラムである. 同様に, 図 5.5 (c) は図 5.5 (a) と図 5.5 (b) を単純に合成することで得られるヒストグラムである. 図 5.4 (c) と図 5.5 (c) に示す通り, 射影された画像特徴と射影されたテキスト特徴を並べた特徴がラベルと高い相関があることが確認できる. さらに, 射影特徴の “Var” と “AveA” の値は元の特徴の値よりも高い.　加えて, 提案手法は比較手法 5 よりも高い精度である. CCA は, 2 種類の特徴間の相関が最大化されるように射影を計算する方法であり, 特徴とラベルとの間の相関を高くするプロセスを含まない. それにも関わらず, CCA により得られた射影された特徴 $[\hat{\boldsymbol{x}_n}^\top, \hat{\boldsymbol{y}_n}^\top]^\top$ がラベルとの相関が高いことから, CCA により得られる射影は, 画像特徴とテキスト特徴間の正準相関の最大化のみならず, 弁別性の高い特徴の算出が可能となることがわかる. 結果的に, 提案手法の有効性が確認された.

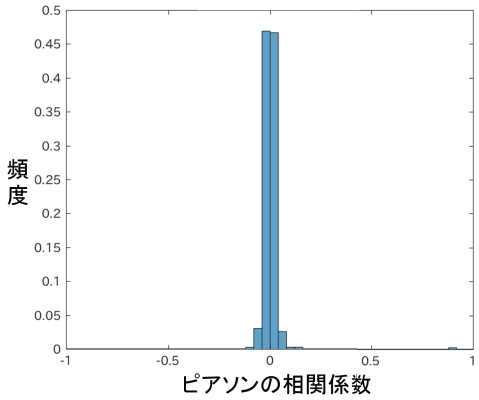
上記の議論により, “Var” と “AveA” が新たに, 元画像と CCA により射影された特徴の選択のための指標になりえることが確認された. 言い換えると, これらの指標を用いることで, 分類精度向上に寄与する特徴を選択することが可能となることも期待できる.

以上より, 異種特徴間の関連性を考慮可能な深層学習により, 少量の画像を用いた場合であっても, 高精度な変状分類が実現する.

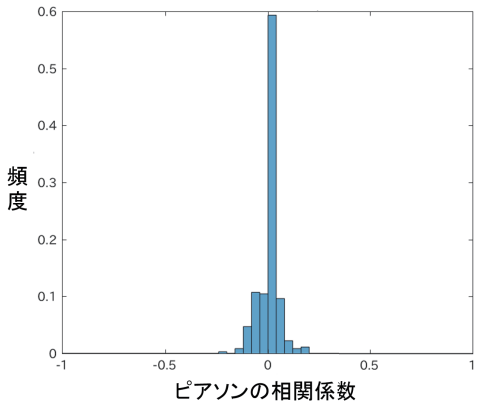
5.6　まとめ

本章では, 異種特徴の相関関係を考慮可能な順伝搬型の深層学習を構築することで, 高精度な変状分類を実現した. 本手法では, 相関最大化を行う射影を順伝搬

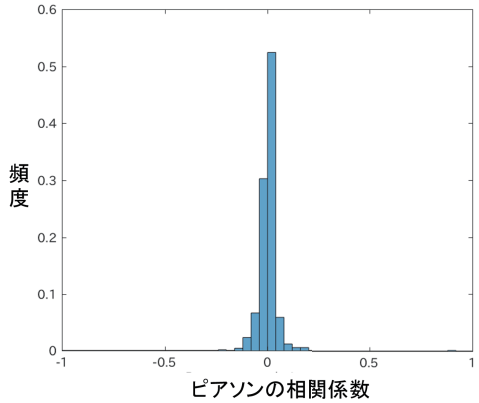
型の深層学習の隠れ層に導入することで, 弁別性の高い特徴への変換と同時に最適に決定すべきパラメータ数の削減が可能とした. 以上より, 少量の学習画像を用いた変状分類の高精度化が実現された.



(a) 画像特徴 $\boldsymbol{x}_n$ とクラスラベルとの相関
Var : 2.25×10^{-3}, AveA : 1.73×10^{-2}

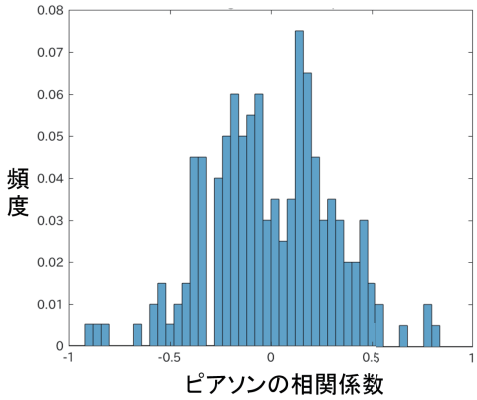


(b) テキスト特徴 $\boldsymbol{y}_n$ とクラスラベルの相関
Var : 2.22×10^{-3}, AveA : 2.67×10^{-2}

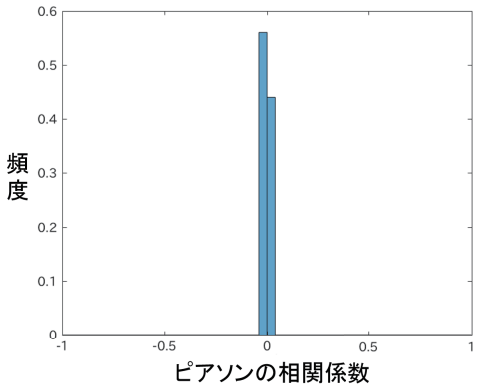


(c) 画像特徴とテキスト特徴を並べた特徴 $[\boldsymbol{x}_n{}^\top, \boldsymbol{y}_n{}^\top]^\top$ とクラスラベルとの相関
Var : 2.24×10^{-3}, AveA : 2.16×10^{-2}

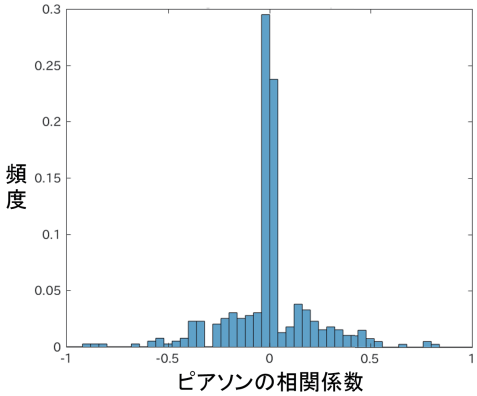
図 5.4: 元特徴とクラスラベルとの間のピアソンの相関係数を示す. また, “Var” は係数の分散を表し, “AveA” は係数の絶対値の和の平均を表す. (a), (b) および (c) は比較手法 3, 4 および 5 にそれぞれ対応している.



(a) 正準画像特徴 $\hat{\boldsymbol{x}}_n$ とクラスラベルとの相関
Var : 9.40×10^{-2}, AveA : 2.49×10^{-1}



(b) 正準テキスト特徴 $\hat{\boldsymbol{y}}_n$ とクラスラベルとの相関
Var : 5.27×10^{-5}, AveA : 7.27×10^{-3}



(c) 正準画像特徴と正準テキスト特徴を並べた特徴 $[\hat{\boldsymbol{x}}_n^\top, \hat{\boldsymbol{y}}_n^\top]^\top$ とクラスラベルとの相関
Var : 4.70×10^{-2}, AveA : 1.28×10^{-1}

図 5.5: 正準特徴とクラスラベルとの間のピアソンの相関係数を示す. また, “Var” は係数の分散を表し, “AveA” は係数の絶対値の和の平均を表す. (a), (b) および (c) は比較手法 1, 2 および提案手法にそれぞれ対応している.

第6章　結論

6.1　本研究の総括

本論文では, 実際の維持管理業務で撮影される変状画像と記録されるテキストデータの異種データから算出される異種特徴を用いて, 分類に有効な弁別性の高い特徴へ変換可能な枠組みを導入した深層学習による変状分類手法を提案した. 以下に, 本論文における各章の概要について述べる.

第 2 章では, 構造物の維持管理支援を目的とした, 機械学習に基づく従来研究について紹介し, 本論文で解決すべき課題を明らかにした. 第 3 章では, 異種特徴と変状を変数とみなすことで, それらの因果関係を確率値として表現可能なベイジアンネットワークに基づく変状分類手法を構築した. 具体的に, 実際の維持管理業務において, 変状画像とテキストデータが技術者の判断に有効であることに着想を得て, 維持管理業務を模擬した枠組みにより変状分類を行った. 実験により, 異種特徴間の関連性の存在を明らかにした. 第 4 章では, 異種特徴に正準相関分析を適用することで変状分類に有効な特徴を算出し, それらの特徴を用いた変状分類手法を構築した. 実験により, 正準相関分析を適用した後の特徴を用いることで変状分類の精度向上が確認できたことから, 異種特徴の相関関係に基づいて算出した射影を用いることの有効性が確認された. 第 5 章では, 得られた射影を隠れ層に導入した順伝搬型の深層学習として新たに Correlation-maximizing Deep Extreme Learning Machine-local Receptive Field を構築した. 順伝搬型の深層学習の隠れ層の重みとして異種特徴の相関関係を考慮して算出した射影を用いることで, 少ない層で分類に有効な特徴への変換を可能とした.　以上により, 最適に決定すべきパラメータ数の削減も可能となることから, 少量の学習画像を用いた高精度な変状分類手法の構築が実現した.

6.2　本研究の今後の課題

本論文では, 実際の維持管理業務で撮影される変状画像と記録されるテキストデータの異種データから算出される異種特徴を用いて, 分類に有効な弁別性の高い特徴へ変換可能な枠組みを導入した深層学習による変状分類手法を構築することで, 少量の学習画像を用いた高精度な変状分類を実現した. 近年の, 構造物の維持管理支援を目的とした研究では, 変状分類だけではなく変状の発生している領域 (以降, 変状領域) も推定可能な手法が提案されているため, 今後は, 異種特徴を用いた変状領域推定への手法拡張が求められる. また, 近年の画像分類では, 人間から得られる視線や脳波などの生体情報を活用することで, 分類精度の向上が実現することが報告されているため [72], 維持管理業務の際の技術者の視線情報を取得することで, 変状分類の更なる高精度化や, 変状領域推定の実現が可能になると考えられる. 以上より, 視線などの生体情報を導入した変状分類の高精度化および変状領域推定の実現が本研究の今後の課題としてまとめられる.

謝辞

本研究は, 著者が北海道大学に在学中に 2014 年 4 月から現在に至る約 5 年間にわたって行ったものである. この 5 年の間に多くの方々の御支援をいただいた.

本研究に関して, 研究の遂行のみならず, 終始御指導, 御鞭撻を頂きました長谷山美紀教授に心より感謝致します. また, 国内・国外問わず, 多くの学会への参加や実際の維持管理業務の現場視察等, 有益な機会を多数頂けたことに対しても, 深くお礼申し上げます.

本論文をまとめるにあたり, 御助言をいただき, 副査をお引き受けいただいた北海道大学大学院情報科学研究科 情報メディア環境学研究室 山本強特任教授, 北海道大学大学院情報科学研究科 言語メディア学研究室 荒木健治教授, 北海道大学大学院情報科学研究科 メディア創生学研究室 坂本雄児教授, ならびに北海道大学大学院情報科学研究科 メディアダイナミクス研究室 小川貴弘准教授に深謝致します.

小川准教授におきましては, 日頃から研究の本質を考えるようご指導とご鞭撻を賜りました. 研究の進め方, 考え方, まとめ方など研究の基礎を丁寧にご教授していただきました. また, 日頃の学生生活においても, 日々の様子を気にかけていただき, 公私にわたり多大なる御助力を頂きましたこと, 心より感謝いたします. さらに, 北海道大学工学研究院, 高橋翔准教授におきましても, 研究に対する姿勢, 研究者としての考え方などをご教授頂きました. また, 研究のみならず学生生活におけるご助言を頂けたこと, どんなに忙しい時でも真摯に対応して頂きましたこと, 心より感謝致します.

また, 本研究のデータをご提供いただいた東日本高速道路株式会社および東京電力ホールディングス株式会社に深く感謝致します. 著者の研究期間中, 御協力頂いた北海道大学大学院情報科学研究科メディアネットワーク専攻メディアダイ

ナミクス研究室諸先輩, 同輩ならびに後輩学生の皆様に感謝致します. 皆様と共に, 切磋琢磨し合うことで, 約 5 年にわたる私の研究が無事に遂行できました.

最後に, これまで, 自分が決断した道を見守り, 応援していただいた家族に深い感謝の意を表し謝辞とさせていただきます.

備考

本論文における, 図表は著者の研究業績 [52, 73, 22] から引用している. 具体的に, 図 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9 と表 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10 は文献 [52] から引用した. 上記図表等は, 個人的な目的でのみダウンロードすることができる. その他の使用に関しては, 米国土木学会 ASCE の事前許可が必要である. また, 文献 [52] は, https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29CP.1943-5487.0000686 でダウンロード可能である.

また, 図 4.1, 4.2, 4.3, 4.4, 4.5 と表 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11 は文献 [73] から引用した. 文献 [73] は, https://doi.org/10.1016/j.aei.2018.04.014 でダウンロード可能である.

さらに, 図 5.1, 5.3, 5.4, 5.5 と表 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 は文献 [22] から引用した. 文献 [22] の DOI は 10.1109/$JSTSP$.2018.2849593 であり, https://ieeexplore.ieee.org/abstract/document/8391700 からダウンロード可能である.

参考文献

[1] MLIT, “White paper on land, infrastructure, transport and tourism in japan 2018.” <http://www.mlit.go.jp/hakusyo/mlit/h29/hakusho/h30/pdf/np101200.pdf>, 2018 (11/26/2018 access).

[2] MLIT. <http://www.mlit.go.jp/common/001158222.pdf>, 2018 (11/26/2018 access).

[3] A. Tedeschi and F. Benedetto, “A real-time automatic pavement crack and pothole recognition system for mobile android-based devices,” *Advanced Engineering Informatics*, vol.32, pp.11–25, 2017.

[4] E. Zalama, G.B.J. Gómez, R. Medina, and J. Llamas, “Road crack detection using visual features extracted by gabor filters,” *Computer-Aided Civil and Infrastructure Engineering*, vol.29, no.5, pp.342–358, 2014.

[5] S. Li, C. Yuan, D. Liu, and H. Cai, “Integrated processing of image and gpr data for automated pothole detection,” *Journal of Computing in Civil Engineering*, p.04016015, 2016.

[6] F. Tsutsumi, H. Murata, T. Onoda, O. Oguri, and H. Tanaka, “Automatic corrosion estimation using galvanized steel images on power transmission towers,” in *Proc. IEEE Transmission & Distribution Conference & Exposition: Asia and Pacific*, pp.1–4, 2009.

[7] B. Yan, S. Goto, A. Miyamoto, and H. Zhao, “Imaging-based rating for corrosion states of weathering steel using wavelet transform and pso-svm techniques,” *Journal of Computing in Civil Engineering*, vol.28, no.3, p.04014008, 2013.

[8] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proc. Advances in Neural Information Processing Systems*, pp.1097–1105, 2012.

[9] L. Zhang, F. Yang, Y.D. Zhang, and Y.J. Zhu, “Road crack detection using deep convolutional neural network,” in *Proc. IEEE Int. Conf. Image Processing*, pp.3708–3712, 2016.

[10] F.C. Chen and M.R. Jahanshahi, “Nb-cnn: deep learning-based crack detection using convolutional neural network and naive bayes data fusion,” *IEEE Trans. Industrial Electronics*, vol.65, no.5, pp.4392–4400, 2018.

[11] X. Gibert, V.M. Patel, and R. Chellappa, “Deep multitask learning for railway track inspection,” *IEEE Trans. Intelligent Transportation Systems*, vol.18, no.1, pp.153–164, 2017.

[12] Y.z. Lin, Z.h. Nie, and H.w. Ma, “Structural damage detection with automatic feature-extraction through deep learning,” *Computer-Aided Civil and Infrastructure Engineering*, vol.32, no.12, pp.1025–1046, 2017.

[13] Y.J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk, “Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types,” *Computer-Aided Civil and Infrastructure Engineering*, vol.33, no.9, pp.731–747, 2018.

[14] R. Li, Y. Yuan, W. Zhang, and Y. Yuan, “Unified vision-based methodology for simultaneous concrete defect detection and geolocalization,” *Computer-Aided Civil and Infrastructure Engineering*, vol.33, pp.527–544, 2018.

[15] K. Gopalakrishnan, H. Gholami, A. Vidyadharan, A. Choudhary, and A. Agrawal, “Crack damage detection in unmanned aerial vehicle images of civil infrastructure using pre-trained deep learning model,” *Int. Journal for Traffic and Transport Engineering*, vol.8, p.1, 2018.

[16] Y. Gao and K.M. Mosalam, “Deep transfer learning for image-based structural damage recognition,” *Computer-Aided Civil and Infrastructure Engineering*, vol.33, no.9, pp.748–768, 2018.

[17] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, and A. Agrawal, “Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection,” *Construction and Building Materials*, vol.157, pp.322–330, 2017.

[18] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, “Decaf: A deep convolutional activation feature for generic visual recognition.,” in *Proc. IEEE Int. Conf. Machine Learning*, pp.647–655, 2014.

[19] Y.R. Yeh, C.H. Huang, and Y.C.F. Wang, “Heterogeneous domain adaptation and classification by exploiting the correlation subspace,” *IEEE Trans. Image Processing*, vol.23, no.5, pp.2009–2018, 2014.

[20] N. Friedman, D. Geiger, and M. Goldszmidt, “Bayesian network classifiers,” *Machine learning*, vol.29, no.2-3, pp.131–163, 1997.

[21] H. Hotelling, “Relations between two sets of variates,” *Biometrika*, vol.28, no.3/4, pp.321–377, 1936.

[22] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Estimation of deterioration levels of transmission towers via deep learning maximizing canonical correlation between heterogeneous features,” *IEEE Journal of Selected Topics in Signal Processing*, vol.12, no.4, pp.633–644, 2018.

[23] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol.20, pp.273–297, 1995.

[24] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[25] C. Szegedy, S. Ioffe, V. Vanhoucke, and A.A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning.,” *AAAI*, p.12, 2017.

[26] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp.770–778, 2016.

[27] Y.J. Cha, W. Choi, and O. Büyüköztürk, “Deep learning-based crack damage detection using convolutional neural networks,” *Computer-Aided Civil and Infrastructure Engineering*, vol.32, no.5, pp.361–378, 2017.

[28] Y. Freund, R. Schapire, and N. Abe, “A short introduction to boosting,” *Journal-Japanese Society For Artificial Intelligence*, vol.14, no.771-780, pp.1–14, 1999.

[29] R. Wu, S. Yan, Y. Shan, Q. Dang, and G. Sun, “Deep image: Scaling up image recognition,” *arXiv preprint arXiv:1501.02876*, 2015.

[30] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in *Proc. Advances in neural information processing systems*, pp.91–99, 2015.

[31] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp.2818–2826, 2016.

[32] G. Huang, Z. Liu, L. Van Der Maaten, and K.Q. Weinberger, “Densely connected convolutional networks.,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, p.3, 2017.

[33] A. Liaw, M. Wiener, *et al.*, “Classification and regression by randomforest,” *R news*, vol.2, no.3, pp.18–22, 2002.

[34] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan Kaufmann, 2014.

[35] S. Takahashi, T. Ogawa, and M. Haseyama, “Feature extraction for road structure assessment based on visual saliency of distress image,” in *Proc. Int. Workshop on Advanced Image Technology*, pp.P. 2–B. 6, 2016.

[36] J.M. Keller, M.R. Gray, and J.A. Givens, “A fuzzy k-nearest neighbor algorithm,” *IEEE Trans. Systems, Man and Cybernetics*, no.4, pp.580–585, 1985.

[37] J.P. Jones and L.A. Palmer, “An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex,” *Journal of Neurophysiology*, vol.58, no.6, pp.1233–1258, 1987.

[38] A.R. Smith, “Color gamut transform pairs,” *ACM Siggraph Computer Graphics*, vol.12, no.3, pp.12–19, 1978.

[39] D.G. Lowe, “Object recognition from local scale-invariant features,” in *Proc. IEEE Int. Conf. Computer Vision*, pp.1150–1157, 1999.

[40] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.27, no.8, pp.1226–1238, 2005.

[41] T.M. Cover and P.E. Hart, “Nearest neighbor pattern classification,” in *Proc. IEEE Trans. Information Theory*, vol.13, no.1, pp.21–27, 1967.

[42] R. Rifkin and A. Klautau, “In defense of one-vs-all classification,” *Journal of Machine Learning Research*, vol.5, pp.101–141, 2004.

[43] V.C. Raykar, S. Yu, L.H. Zhao, A. Jerebko, C. Florin, G.H. Valadez, L. Bogoni, and L. Moy, “Supervised learning from multiple experts: whom to trust when everyone lies a bit,” in *Proc. Annu. Int. Conf. Machine Learning*, pp.889–896, 2009.

[44] Y. Yan, R. Rosales, G. Fung, M.W. Schmidt, G.H. Valadez, L. Bogoni, L. Moy, and J.G. Dy, “Modeling annotator expertise: Learning when everybody knows a bit of something,” *AISTATS*, pp.932–939, 2010.

[45] F. Rodrigues, F. Pereira, and B. Ribeiro, “Learning from multiple annotators: Distinguishing good from random labelers,” *Pattern Recognition Letters*, vol.34, no.12, pp.1428–1436, 2013.

[46] J. Goldberger, “Combining soft decisions of several unreliable experts,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp.2334–2338, 2016.

[47] C.W. Hsu, C.C. Chang, C.J. Lin, *et al.*, “A practical guide to support vector classification,” 2003.

[48] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Distress classification of road structures via decision level fusion,” in *Proc. IEEE Int. Conf. Digital Signal Processing*, pp.589–593, 2016.

[49] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Distress classification of road structures via multiple classifier-based bayesian network,” in *Proc. Int. Workshop on Advanced Image Technology*, pp.P. 1–B. 11, 2016.

[50] Y. Xu and Y. Lu, “Adaptive weighted fusion: A novel fusion approach for image classification,” *Neurocomputing*, vol.168, pp.566–574, 2015.

[51] W. Zong, G.B. Huang, and Y. Chen, “Weighted extreme learning machine for imbalance learning,” *Neurocomputing*, vol.101, pp.229–242, 2013.

[52] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Distress classification of road structures via adaptive Bayesian network model selection,” *Journal of Computing in Civil Engineering*, vol.31, no.5, p.04017044, 2017.

[53] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, “Decaf: A deep convolutional activation feature for generic visual recognition,” in *Proc. Int. Conf. Machine Learning*, pp.647–655, 2014.

[54] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pp.2818–2826, 2016.

[55] W.F. Schmidt, M.A. Kraaijveld, and R.P. Duin, “Feedforward neural networks with random weights,” in *Proc. IAPR Int. Conf. Pattern Recognition*, pp.1–4, 1992.

[56] Y.H. Pao, G.H. Park, and D.J. Sobajic, “Learning and generalization characteristics of the random vector functional-link net,” *Neurocomputing*, vol.6, no.2, pp.163–180, 1994.

[57] G.B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Trans. Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol.42, no.2, pp.513–529, 2012.

[58] R. Fletcher, Practical methods of optimization, 2013.

[59] G.B. Huang, Q.Y. Zhu, and C.K. Siew, “Extreme learning machine: a new learning scheme of feedforward neural networks,” in *Proc. IEEE Int. Joint Conf. Neural Networks*, pp.985–990, IEEE, 2004.

[60] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Distress classification of class imbalanced data for maintenance inspection of road structures in express

way," in *Proc. Int. Conf. Civil and Building Engineering Informatics*, pp.182–185, 2017.

[61] B. Mirza, S. Kok, Z. Lin, Y.K. Yeo, X. Lai, J. Cao, and J. Sepulveda, "Efficient representation learning for high-dimensional imbalance data," in *Proc. IEEE Int. Conf. Digital Signal Processing*, pp.511–515, IEEE, 2016.

[62] G.B. Huang, Z. Bai, L.L.C. Kasun, and C.M. Vong, "Local receptive fields based extreme learning machine," *IEEE Computational Intelligence Magazine*, vol.10, no.2, pp.18–29, 2015.

[63] A. Coates, H. Lee, and A.Y. Ng, "An analysis of single-layer networks in unsupervised feature learning," *Ann Arbor*, vol.15, pp.215–223, 2010.

[64] I.T. Jolliffe, "Principal component analysis," 2002.

[65] E. Cambria, G.B. Huang, L.L.C. Kasun, H. Zhou, C.M. Vong, J. Lin, J. Yin, Z. Cai, Q. Liu, and K.L. et al., "Extreme learning machines [trends & controversies]," *IEEE Intelligent Systems*, vol.28, no.6, pp.30–59, 2013.

[66] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Proc. Advances in Neural Information Processing Systems*, pp.341–349, 2012.

[67] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Automatic estimation of deterioration level on transmission towers via deep extreme learning machine based on local receptive field," in *Proc. IEEE Int. Conf. Image Processing*, pp.2379–2383, 2017.

[68] G.B. Huang, Q.Y. Zhu, and C.K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol.70, no.1, pp.489–501, 2006.

[69] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in *Proc. ACM Int. Conf. Multimedia*, pp.675–678, ACM, 2014.

[70] H. Hamada *et al.* Tokyo Electronic Power Company Holdings, Inc. (*private communication*).

[71] S.B. Chen, Y. Zhang, C.H. Ding, Z.L. Zhou, and B. Luo, “A discriminative multiclass feature selection method via weighted l2, 1-norm and extended elastic net,” *Neurocomputing*, vol.275, pp.1140–1149, 2018.

[72] N. Karessli, Z. Akata, B. Schiele, A. Bulling, *et al.*, “Gaze embeddings for zero-shot image classification,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2017.

[73] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Distress classification of class-imbalanced inspection data via correlation-maximizing weighted extreme learning machine,” *Advanced Engineering Informatics*, vol.37, pp.79–87, 2018.

著者の研究業績

論 文

(A) 学会誌

[**A-1**] K. Maeda, T. Ogawa, and M. Haseyama, "Automatic Martian Dust Storm Detection from Multiple Wavelength Data Based on Decision Level Fusion," *IPSJ Transactions on Computer Vision and Applications*, vol. 7, pp. 79–83, 2015.

[**A-2**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Distress Classification of Road Structures via Adaptive Bayesian Network Model Selection," *ASCE Journal of Computing in Civil Engineering*, vol. 31, no. 5, pp. 04017044-1–04017044-14 (2017 IF:1.798).

[**A-3**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Distress Classification of Class-imbalanced Inspection Data via Correlation Maximizing Weighted Extreme Learning Machine," *Elsevier Journal of Advanced Engineering Informatics*, vol. 37, pp. 79–87, 2018. (2017 IF:3.358).

[**A-4**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Estimation of Deterioration Levels of Transmission Towers via Deep Learning Maximizing Canonical Correlation between Heterogeneous Features," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 633–644, 2018. (2017 IF:4.361).

[**A-5**] T. Ogawa, Y. Sasaka, K. Maeda and M. Haseyama, "Favorite Video Classification Based on Multimodal Bidirectional LSTM," *IEEE Access* (2017 IF:3.557) (Accepted for publication).

(B) 国際学会

[**B-1**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Bayesian Network-based Distress Estimation Using Image Features in Road Structure Assessment," in *Proc. IEEE Global Conf. Consumer Electronics*, 2014, pp. 169–170.

[**B-2**] K. Maeda, T. Ogawa, and M. Haseyama, "Automatic Detection of Martian Dust Storms from Heterogeneous Data Based on Decision Level Fusion," in *Proc. IEEE Int. Conf. Image Processing*, 2015, pp. 2246–2250.

[**B-3**] K. Maeda, T. Ogawa, and M. Haseyama, "Distress Classification of Road Structures via Multiple Classifier-based Bayesian Network," in *Proc. Int. Workshop on Advanced Image Technology*, 2016, pp. P.1–B.11.

[**B-4**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Distress Classification of Road Structures via Decision Level Fusion," in *Proc. IEEE Int. Conf. Digital Signal Processing*, 2016, pp. 589–593.

[**B-5**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Distress Classification of Class Imbalanced Data for Maintenance Inspection of Road Structures in Express Way," in *Proc. Int. Conf. Civil and Building Engineering Informatics in conjunction with Conference on Computer Applications in Civil and Hydraulic Engineering*, 2017, pp. 182–185.

[**B-6**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Deterioration Level Estimation on Transmission Towers via Extreme Learning Machine based on Combination Use of Local Receptive Field and Principal Component Analysis," in *Proc. Int. Tech. Conf. Circuits/Systems, Computers and Communications*, 2017, pp. 457–458.

[**B-7**] K. Horii, K. Maeda, T. Ogawa, and M. Haseyama, "Image Classification for Trend Prediction Based on Integration of fNIRS and Visual Features," in *Proc. Int. Global Conf. Consumer Electronics*, 2017, pp. 301–302.

[**B-8**] K. Maeda, T. Ogawa, and M. Haseyama, "Automatic Martian Dust Storm Detection via Decision Level Fusion Based on Deep Extreme Learning Machine," in *Proc. IEEE Int. Conf. Image Processing*, 2017, pp. 435–439.

[**B-9**] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, "Automatic Estimation of Deterioration Level on Transmission Towers via Deep Extreme Learning Machine based on Local Receptive Fields," in *Proc. IEEE Int. Conf. Image Processing*, 2017, pp. 2379–2383.

[**B-10**] K. Horii, K. Maeda, T. Ogawa, and M. Haseyama, "A Human-centered Neural Network Model with Discriminative Locality Preserving Canonical Correlation Analysis for Image Classification," in *Proc. IEEE Int. Conf. Image Processing*, 2018, pp. 2366–2370.

[**B-11**] Y. Moroto, K. Maeda, T. Ogawa, and M. Haseyama, "User-centric Visual Attention Estimation Based on Relationship Between Image and Eye Gaze Data," in *Proc. Int. Global Conf. Consumer Electronics*, 2018, pp. 44–45.

(C) 技術報告

[**C-1**] K. Maeda, T. Ogawa, and M. Haseyama, "Automatic Martian Dust Storm Detection from Multiple Wavelength Data Based on Decision Level Fusion," in *Proc. 18 th Meeting on Image Recognition and Understanding*, 2015, pp. OS4-5.

[**C-2**] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, "道路構造物の維持管理における高精度な変状分類に関する検討 - タグデータと撮影画像に基づく識別結果の統合 -," 映像情報メディア学会技術報告, vol. 40, no. 6, pp. 181-184, 2015.

[**C-3**] 前田 圭介, 小川 貴弘, 長谷山 美紀, "火星に発生する dust storm の自動検出に関する検討 - 不均衡データを考慮した識別器の構築 -," 映像情報メディア学会技術報告, vol. 40, no. 28, pp. 47-48, 2016.

[C-4] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “深層学習を用いた道路構造物の維持管理における変状分類の高精度化に関する検討,” 映像情報メディア学会技術報告, vol. 41, no. 5, pp. 51-54, 2017.

[C-5] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “正準相関最大化を導入した深層学習に基づく送電鉄塔の劣化レベル分類に関する検討,” 映像情報メディア学会技術報告, vol. 41, no. 29, pp. 11–14, 2017.

[C-6] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “道路構造物に発生する変状の自動分類の高精度化に向けた Convolutional Sparse Coding の導入に関する検討,” 映像情報メディア学会技術報告, vol. 42, no. 4, pp. 189-194, 2018.

[C-7] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “Convolutional Sparse Coding を導入した深層学習に基づく変状分類手法の構築,” 第 21 回 画像の認識・理解シンポジウム, pp. PS2-47, 2018.

(D) 講演発表

[D-1] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “複数の画像特徴を用いたベイジアンネットワークに基づく構造物の変状の推定の高精度化に関する検討,” 電気・情報関係学会北海道支部連合大会 講演論文集, p. 140, 2014.

[D-2] 前田 圭介, 高橋 翔, 小川 貴弘, 長谷山 美紀, “個々の道路構造物に関する点検項目の導入による道路構造物の変状推定の高精度化に関する検討,” 電気・情報関係学会北海道支部連合大会 講演論文集, p. 133, 2015.

[D-3] 堀井 風葉, 前田 圭介, 小川 貴弘, 長谷山 美紀, “画像特徴量と fNIRS 特徴量の関連性に注目した画像分類の高精度化に関する検討,” 電気・情報関係学会北海道支部連合大会 講演論文集, pp. 114–115, 2017.

[D-4] 諸戸 祐哉, 前田 圭介, 小川 貴弘, 長谷山 美紀, “画像注視時の注視領域の時間変化を考慮したテンソル解析に基づく感情推定に関する検討,” 電気・情報関係学会北海道支部連合大会 講演論文集, pp. 137–138, 2018.

[D-5] 斉藤 僚汰, 前田 圭介, 小川 貴弘, 長谷山 美紀, “地下鉄トンネルの点検における視線データを用いた熟練度の分析に関する一考察,” 電気・情報関係学会北海道支部連合大会 講演論文集, pp. 14–15, 2018.

(E) 受賞

[E-1] 平成 27 年度電気・情報関係学会北海道支部連合大会優秀論文発表賞 (2016 年 1 月)

[E-2] The 2015 IEEE Sapporo Section Encouragement Award (2016 年 2 月)

[E-3] 2017 IBM Ph. D. Fellowship Award (2017 年 3 月)

[E-4] 精密工学会画像応用技術専門委員会・映像情報メディア学会メディア工学研究委員会合同サマーセミナー 優秀発表賞 (2017 年 9 月)

[E-5] The 2017 IEEE Sapporo Section Encouragement Award (2018 年 1 月)