

# 経験知に基づく画像表現からの CNN 特徴学習の検証

## Verification of CNN training from feature images based on worker's experience knowledge.

山川 泰司<sup>\*1</sup>  
Taishi Yamakawa

黒木 雅彦<sup>\*2</sup>  
Masahiko Kuroki

岡部 弘志<sup>\*2</sup>  
Hiroshi Okabe

<sup>\*1</sup> 東京電力ホールディングス株式会社 経営技術戦略研究所  
Tokyo Electric Power Company Holdings, Inc. TEPCO Research Institute

<sup>\*2</sup> 東京電力フュエル&パワー株式会社  
TEPCO Fuel & Power, Incorporated

For CNN training, abnormal images were drawn using original normal images based on skilled worker's experience knowledge. CNN could learn features of abnormality from those drawn images. Even in the case of rare abnormality like machine trouble, if abnormal images are visualized with skilled worker's know-how, CNN will be possible method to detect abnormality without real teacher data.

### 1. はじめに

Deep Learning (深層学習) は, GPU 等のハードウェア環境, オープンデータやツール等のソフトウェア環境, および戦略的な研究投資を背景に著しい発展を遂げている. 産業分野においても, 既存の情報資産や IoT 化で増大するデータを活用した新たな価値の創出が期待されている.

とりわけ, CNN (Convolutional Neural Network) による画像認識の精度は, 従来人間が担っていた異常判別や外観検査を代替できる水準に達している.

筆者らは, 異常データの替わりとして, CNN に熟練者の経験知に基づく画像を学習させ, 異常を判別させることを試みた. 試験結果と学習プロセスの考察を踏まえながら, 現場の経験知を CNN の学習に活用する可能性について検証した.

### 2. 低頻度の異常の学習

CNN による画像認識は, 画像の特徴量を自動的に, かつ精度よく学習できる利点をもつ一方, 学習に使用する教師データを多数用意する必要がある.

例えば ImageNet のような画像データベースを利用した大規模な多クラス分類のベンチマークでは, 1 クラスあたり数百から数千の画像を教師データとして学習する. 多様なクラスや複雑な特徴を精度よく認識させるためには, 相応数の教師データが必要とされる. 製造現場において CNN により製品の的外観検査を行う場合は, 教師データとなる正常時と異常時の画像データを予め準備しておく必要がある.

ところが, 産業の現場によっては, 教師データの無い異常が生じる場合がある. 例えば, 設備の損傷トラブルのような事象は低い頻度でしか発生しないため, 異常データの取得機会そのものが少ない. 教師データが十分に無いと, 深層学習で異常の特徴を学習させることが困難となる.

その一方で, 設備の損傷トラブルは経営的に大きなリスクを伴うことから, 現象特有の感知センサーや, 人間によるパトロール, ルール記述による異常検知システム等, トラブルを初期段階で発見するために多くのリソースを費やしている場合がある. このような低頻度の異常に対して画像認識による異常判別が適用出来れば, 産業の現場にもたらすメリットは大きい.

### 2.1 検証課題の概要

検証の対象とする設備は, チェーンにけん引される運搬板が, コンベア面に落下した石炭を掻いて移送する石炭回収コンベアである. 異常判別の対象は, 運搬板の変形である. 異常な変形は年に 1 回程度の頻度で生じ, パトロール時の目視により発見されている. 発見後は速やかに修繕作業を行うため, 教師データになりうる稼働中の異常を捉えた画像は取得されていない. 設備の概要, 稼働中の画像, 変形が生じた運搬板の写真を図 1, 図 2 に示す.

画像データを収集するため, 設備の監視窓に PC カメラを設置してインターバル撮影を行った. 照明とシャッター速度の性能上, 移動中の板は不鮮明に映るが, 変形が生じた場合は全体の形状から判別できる. 収集した画像から必要な領域をトリミングし, 縦 100 ピクセル, 横 160 ピクセルにリサイズした後, グレースケール画像に変換した. その後, 画像に映る運搬板の外観に応じて保留, 正常, 異常のクラスに分類し, 保留 4,308 枚, 正常 3,682 枚, 異常 0 枚の教師データを得た (図 3).

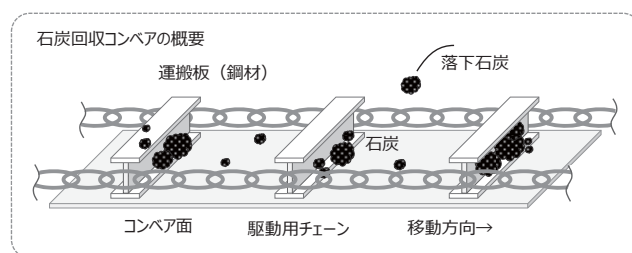


図 1: 検証する設備の概要

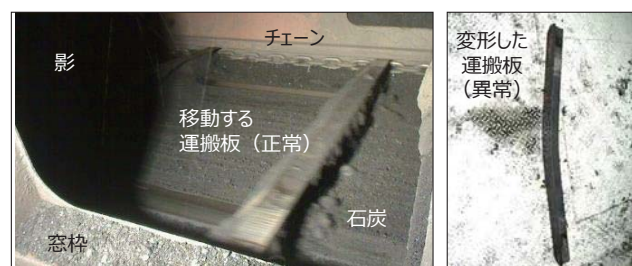


図 2: 稼働中の画像と運搬板の異常例

連絡先: 山川 泰司, 東京電力ホールディングス株式会社  
経営技術戦略研究所, yamakawa.taishi@tepcoco.jp

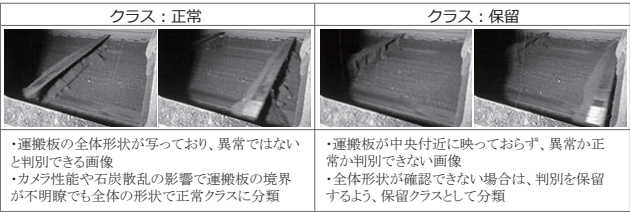


図 3: 取得画像のクラス分類

2.2 形式知に基づく学習の検討

学習に使用する教師データの不足は根本的な課題である。この課題に対する既往のアプローチは、できるだけ定量的な情報を収集し、明示的に特徴を記述するものであり、人間があいまいな情報からでも特徴を認識できることと比較すると、形式知を介した学習には多くのコストを必要とする。

(1) 異常の物理的再現による学習

実験等の手段により試験体に異常を生じさせ、その画像を教師データとすることで異常を学習する方法がある。現実の条件に即した異常の再現が可能であるが、試験体や設備、再現実験に要する大きなコストが課題となる。

(2) シミュレーション結果の学習

シミュレーションの結果を学習データとする方法がある。例えば、地震動に対する応答を大規模な数値計算で求め、その結果を深層学習の教師データに用いた研究がある[藤田 17]。今回のような異常に対しても、形状や物性、応力や条件をモデリングし、異常をシミュレーションで再現させて画像情報を取得することは不可能ではない。しかしながら、再現条件の取得、現象の正確な把握、精緻な数理解析が必須となり、現場の設備へ展開することは難しい。

(3) VAE, GAN による画像データ生成

人工知能による画像の生成手段、特に VAE (Variational Auto Encoder) や GAN (Generative Adversarial Network) では、現実感のある高品質な画像を自動生成できるようになっている。生成された画像を学習データとして活用する研究も行われており、例えば少数の教師画像から GAN を使用して類似の画像を生成し、CNN の教師データ数を増加させることにより精度向上を図る検証がおこなわれている[伊藤 17]。

他方、これら深層学習に基づき生成される画像は、教師画像を低次元化して得た特徴量をベースに再構成された画像であるため、画像に含まれる特徴に偏りが生じる可能性が示唆されており[Arora 17]、特徴の多様性をいかに充実するかが検討課題として残る。

(4) 教師なし学習による判別

異常クラスの画像が無い場合、正常クラスの画像のみを使用して教師なし学習が行われる場合がある。教師なし学習は、データ群の分布のばらつきを計算し、ばらつく範囲からの外れ度合によって学習したデータ群に属するか否かを判別する手法で、マハラノビス距離や One Class SVM 等が用いられる。

規格製品のような類似した画像を学習する場合は、外れ度合を基準にすることで異常が生じた製品を判別することができる。しかし、背景にばらつきのある画像を学習すると、相対的に小さな異常はばらつきの範囲内に入ってしまう、異常として認識されにくくなる特性がある。

このことから、教師なし学習で異常を判別する場合には、安定した画像が取得できるよう環境条件を整えたり、特徴を捉えるためにルール記述を併用する必要がある。

2.3 経験知に基づく学習の検証

本検証では、異常の特徴を CNN に学習させるため、経験知に基づいて異常を表現した画像を生成し、教師データとして学習に使用することを試みた。学習した CNN が、意図した特徴を正しく認識し、正常と異常を判別することが可能であるか実際の設備画像を用いて検証した。

(1) 教師データの生成

正常クラスの画像を加工編集し、異常が生じた運搬板の外観を模擬的に再現した。再現のため、異常な変形の具体的な状態を、報告書の内容と現場作業員のヒアリングに基づく経験知として得た。得られた経験知を画像として表現するために、市販の写真編集ソフトウェア Adobe Photoshop を使用した。このソフトウェアは、写真の加工や合成、色調変換や補間処理等、高度な画像編集機能を備えている。

異常を表現するための加工編集手順は、まず正常クラスの画像を編集ソフトウェアに読み込み、加工する範囲を指定する。次に編集手段：変形を選択し、プレビュー画面を確認しながら変形の方法や量をマウス操作により直観的に調整していく。いくつかの編集手順を組み合わせ、最終的に経験知と整合するように、異常の外観状態を再現する。再現出来たら、編集操作の履歴をマクロとして保存する。保存されたマクロは、正常クラスの画像から異常クラスの画像を生成するための加工フィルタとしての機能をもつ。本検証では、加工の領域や編集の量が異なる 3 種類のマクロを使用し、正常クラスの画像から同数の異常クラスの画像を生成した(図 4)。

加えて、写真を加工するための複数種のフィルタを重ねて適用することで、コントラストやノイズ分布が異なる画像を生成し、学習に使用する教師データ数を計 46,660 枚まで増加させた(図 5)。

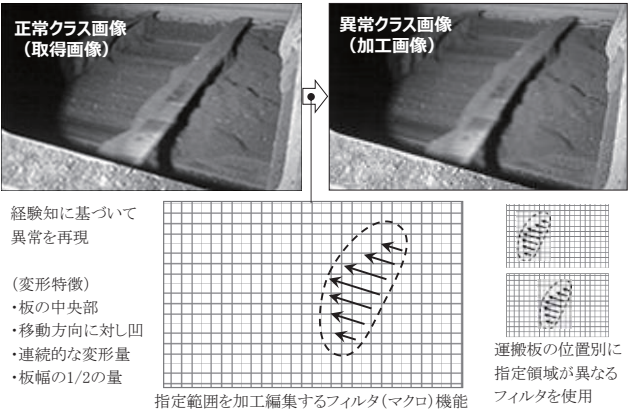


図 4: 経験知を反映した異常状態の生成

適用画像	加工フィルタ			
	コントラスト強調 加工フィルタ	ぼかし 加工フィルタ	ランダムノイズ 付加フィルタ	
正常クラス画像 (取得画像)	正常クラス + 輝度分布の鮮鋭化	正常クラス + 輝度分布の平滑化	正常クラス + 輝度分布の均質化	
異常クラス画像 (加工画像)	異常クラス + 輝度分布の鮮鋭化	異常クラス + 輝度分布の平滑化	異常クラス + 輝度分布の均質化	

図 5: 学習画像のデータ拡張



(2) CNN による判別結果

学習させる CNN は、畳み込み 2 層、プーリング 2 層、全結合 2 層を経て 3 クラスに分類する (図 6)。

試験用の画像は、学習用の画像から予め分割した。正常と異常のクラスはデータ拡張前の画像単位で分割し、保留のクラスでは生成した画像の中から無作為に分割した。

訓練後に試験を行った結果、異常クラスの画像に対して 99% 以上の精度で正しく判別するネットワークが学習された (表 1)。誤判別も少なく、経験知で表現された特徴を CNN が正しく認識していることが示唆された。

3. 経験知からの特徴学習に関する検証

判別試験の結果に基づき、CNN の学習過程を踏まえた特徴表現の留意点と、経験知の活用に向けた課題を検証した。

3.1 画像からの特徴抽出

CNN の畳み込みフィルタ層は、1 層目は画像中の数ピクセル四方の特徴を、2 層目以降は特徴の結合形状を抽出する機能を持つ。畳み込む回数が増えるほど複雑な形状の認識が可能になるが、学習の負荷は増加する。

ニューラルネットワークの全結合 (fc) 層は、畳み込み層で出力された反応マップと教師データのクラスの間を、ニューラルネットワークで関連付ける機能を持つ。fc のパラメータも認識性能と学習負荷のトレードオフになる。

判別精度を評価する際には、これらの学習パラメータが画像の特徴を正しく抽出したかを確認する必要がある。

(1) 過学習の抑制

学習に使用する教師データ数がパラメータの数に対して少ない場合、クラスの特徴ではなくデータ固有の特徴を学習する過学習が生じやすくなる。過学習は判別精度の低下要因となるため、本検証例のように教師データを増加させたり、ドロップアウトを適用して抑制する必要がある [岡谷 15]。

(2) 加工痕の学習の抑制

本手法では、画像を加工編集して異常状態を表現する。その際、意図しない加工編集の痕跡、例えば不自然な輝度境界やパターン等が残存した場合、その痕跡を異常の特徴として学習し、実際の異常を正しく認識しない可能性が生じる。このような誤った学習を抑止するためには、本来の特徴以外に偏った情報を持たない教師データで学習するのが望ましい。

本手法では、学習用の画像にコントラスト強調、ぼかし加工、ノイズ付加のフィルタを適用し、画像にわずかに変化を与えて枚数を増加させた。このような操作によるデータ拡張は CNN の 5x5 ピクセルの畳み込み領域における輝度分布に多様性をもたせ、意図しない誤った特徴の学習を抑止する。

CNN がどのような領域の特徴に反応してクラスを判別しているかを確認する手法として、学習時に逆伝播させる誤差の勾配を視覚化する手法が提案されている [Selvaraju 17]。今回は、畳み込みフィルタの 1 層目の学習結果と、異常画像の入力に対する反応マップを直接確認することで、加工痕が学習されていないことを確かめた。

図 7 上はフィルタの学習結果、図 7 下は異常クラス画像の反応マップの例 (64 チャンネル) を示す。CNN が異常の加工痕を学習した場合、反応マップに意図しない反応領域が生じると推察される。本検証例では、フィルタは輝度の濃淡やエッジを抽出するよう学習され、反応マップにも不自然な反応領域が無いことから、誤った特徴は学習されていないと評価した。

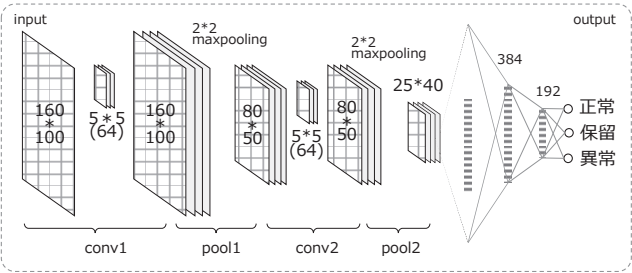
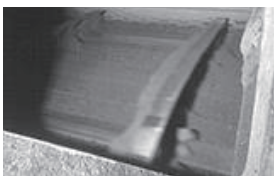
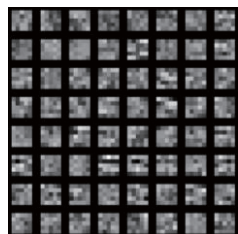


図 6: CNN の構成

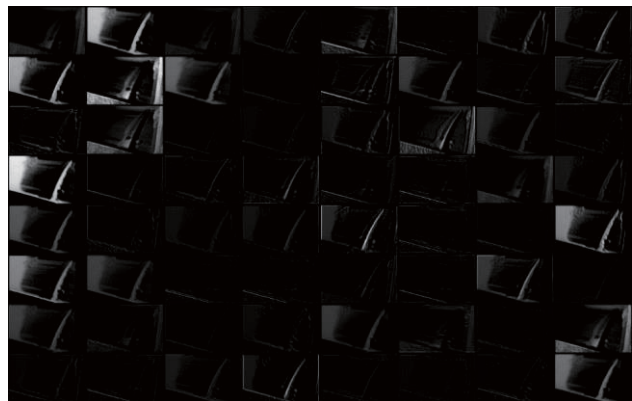
表 1: CNN の判別結果

訓練画像数	試験画像数	クラス	判別：正常	判別：異常	判別：保留
10,968	4,704	正常	96.7 %	0.4 %	3.0 %
10,968	4,704	異常	0.1 %	99.7 %	0.3 %
13,784	1,532	保留	2.3 %	0.0 %	97.7 %



異常クラス画像

畳み込みフィルタ1層目 (64チャンネル)  
輝度の濃淡やエッジ部に反応するよう学習されたと推察



反応マップ (64チャンネル): 加工編集の境界部位に反応は生じていない

図 7: 畳み込みフィルタ (1 層目: 上)  
フィルタの反応マップ (1 層目: 下)

(3) 特徴の選択的学習

本手法は、正常クラスの画像 (周辺部 A + 正常部 B) を加工編集して異常クラスの画像 (A + 異常部 B') を生成する (図 8)。よって、クラスの差は指定領域 (B' - B) にのみ現れる。例えば、強い特徴を有する背景 C や、外乱要素 D が含まれるような画像であっても、B' - B のみがクラスの差の特徴として学習される。すなわち、異常の領域を選択的に学習させる効果がある。今回収集した運搬板の画像群は、板の位置、コンベアの表面状態、石炭の量のばらつきに対し、変形による輝度の分布変化は相対的に小さい。それでも異常クラスを精度よく判別することから、本手法による選択的な領域の学習が、特徴の抽出精度の向上に有効に作用すると考えられる。

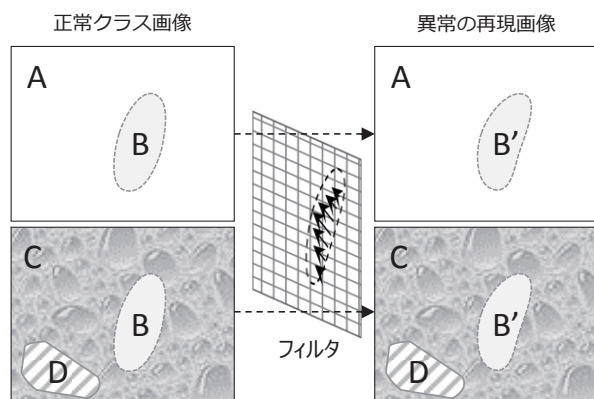


図 8: 選択的な特徴の学習

### 3.2 経験知の活用に向けた課題

#### (1) 他の異常状態への適用性

本手法は、外観によりクラスを判別する課題に広く適用することができると考えられる。必要な条件は、加工編集の元となる画像が収集でき、クラスの差を画像として表現できることである。例えば、破損、腐食、汚損、異物混入等の異常に対しても、各々の状態を表現する加工編集のフィルタを作成することにより学習できる可能性がある。また、他の異常判別手法や、画像生成手法と組み合わせて利用することもできる。

課題として、経験知の無い未知の事象や、外観で表現できない事象については学習できないことが挙げられる。

#### (2) 経験知の画像表現プロセス

最近の高機能な写真編集ソフトウェアは、人間の目には見分けられないほど緻密な加工写真を容易に生成できるようになっている。これにより、人間の経験知をリアルな画像として表現するアプローチを検討できるようになった。

実際に、視覚的に表現した特徴を、CNN アルゴリズムは画像から直接的に認識することができた。本例のような画像を介した特徴の伝達は、言語的な記述よりもリッチな情報を簡潔に伝達できる可能性があり、特徴学習のプロセスをより人間的な手段に近づけたと考えることもできる。

一方、経験知の深さや画像の表現能力には個人差が存在し、伝達される特徴の質や量も人によって異なる。経験知の学習品質を安定化させるためには、個人の経験知の表現能力の向上や、複数の経験知による均質化も課題として捉えてゆく必要がある。このような伝達プロセスが整備されると、経験知に基づく学習がより実用的になると考えられる。

#### (3) 現場での活用と継続的な学習

経験知により特徴を学習した CNN は、経験知に基づく異常は判別可能になる。しかし、現実の異常を認識できるかどうかは、現実の異常データを入力しなければわからない。そのため、本手法による学習で一定の評価を得た段階で、現場での試運用に移行することが望ましいと考えられる。

移行した後は、現場で収集される新たな教師データを使用して追加の学習や精度の確認、経験知の更新やシステムの改良を随時反映していくことで、現実の異常に対する判別の信頼度を高めることが可能だと考えている。

## 4. おわりに

本検証では、経験知に基づく異常の特徴を、画像表現を介して CNN に学習させることにより、高い精度で異常を判別できることを確認した。このことは、これまで深層学習による学習が困難であった学習データの無い低頻度の異常に対しても、画像認識により異常を検知するための有用な一手段となる可能性を示唆している。

産業の現場では、アナログ情報の時代から継承されている有益な経験知が多い。人間にとっては、理屈や因果関係から導かれるような単純な知識であっても、その内容を標準化したり、形式知に変えて伝達するには大変な労力を要する。労働力の減少が懸念される現状において、熟練者の経験知が引き継がれないまま失われていく課題も生じている。

社会課題に対する AI の活用に関しては、データの収集と深層学習そのものを生産のプロセスと捉える学習工場[松尾 17]のフレームワークが提案されているように、既存の経験知をより直観的に AI へ伝達させる学習が実現できれば、産業分野における AI 活用の裾野がさらに広がると考えている。

今後、本検証内容を実際の現場設備に適用して試運用を行い、経験知を活用した画像認識による異常判別の評価を進めたい。

## 参考文献

- [藤田 17] 藤田航平: 大規模有限要素解析と AI によるモンテカルロ地震動強度分類, session 2042, GTC Japan 2017, 2017.
- [伊藤 17] 伊藤秀将, 入本勇次, 榎本晋一: 画像認識精度を向上させるディープラーニング技術を用いた学習用画像の自動生成, 東芝レビュー Vol.72 No.4, 2017.
- [Arora 17] Sanjeev Arora, Yi Zhang: Do GANs actually learn the distribution? An empirical study. arXiv preprint arXiv: 1706.08224, 2017.
- [岡谷 15] 岡谷貴之: 深層学習, 講談社, 2015.
- [Selvaraju 17] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. arXiv preprint arXiv: 1610.02391, 2017.
- [松尾 16] 松尾豊: DL 産業論, 参考資料(プレゼン資料), 産業構造審議会新産業構造部会(第 11 回), 2016, [http://www.meti.go.jp/committee/sankoushin/shin\\_sangyoukouzou/011\\_haifu.html](http://www.meti.go.jp/committee/sankoushin/shin_sangyoukouzou/011_haifu.html).