

ハンディカメラを用いた列車前方映像からの設備の劣化度推定

○ [電] 向嶋 宏記 [電] 長峯 望 [電] 合田 航

(公益財団法人鉄道総合技術研究所)

Deterioration Estimation of Facilities from Train Driver's View Images
using One Handy Camera

○Hiroki Mukojima, Nozomi Nagamine, Wataru Goda

(Railway Technical Research Institute)

Signalling and communications facilities in the railway are installed not in one place but scattered out along track-side between adjacent stations. For this reason, a great deal of labour is currently required in maintenance work for performing individual inspections, and in facility management work for ascertaining the installed positions and their types. In order to reduce the workload in maintenance, there is a requirement for a system that can automatically recognize and inspect the equipment without going to the site. Therefore, we are developing a system that supports the maintenance work of signal equipment using only a handy camera. Our system is that assists ledger management by recognizing signal and communication equipment from the video obtained by the handy camera and estimating the location and deterioration of the equipment. This paper describes the outline of our system and the estimation results of a study on a deterioration of the equipment by deep learning.

キーワード: ハンディカメラ, 深層学習, 劣化度推定, 設備保全, 列車前方映像

Key Words: Handy camera, Deep learning, Deterioration estimation, Maintenance, Train driver's view

1. はじめに

鉄道において、線路沿線の信号通信関係の設備は駅および駅間に点在して設置されている。そのため、線区全体の設備の位置、種別の把握、並びに設備の個別検査といった設備管理に多大な労力を要している。

このような沿線設備のメンテナンス業務において、作業負担を軽減するようなシステムが注目されている。特にこれまでは Time Based Maintenance (TBM)の考え方が一般的であったが、近年は Condition Based Maintenance (CBM)に向けて機器のモニタリング等の取り組みがなされている。

そこで我々は市販のハンディカメラのみを用いて信号設備のメンテナンスを支援するシステムを開発している¹⁾。具体的には、ハンディカメラから得られた列車前方の映像から、設備等を自動で認識することでメンテナンス業務を支援するシステムである。本システムの構成品はビデオカメラと三脚等のカメラ雲台のため、導入の初期コストを抑

えられるというメリットがある。

これまでに提案したシステムでは、設備の劣化度合い(劣化度)の推定については検討していない。列車前方映像からの設備の劣化度を推定する際の課題として、見た目から設備の劣化を定量的な数値で表現することが難しいということが挙げられる。画像から劣化していないのか、大きく劣化しているのかなど大まかな判断は可能だが、少しでも劣化しているような中間の状態を数値として判断できないため、回帰モデルの教師データの作成が難しい。本稿では、この問題に対して Deep Metric Learning (DML)を用いて、劣化度の定量的な値を教師データとして付与することなく設備の劣化度を推定する手法を提案する。また、器具箱を対象に適用した結果についても報告する。

2. ハンディカメラのみを用いた沿線設備管理システム

提案システムにはハンディカメラで撮影した列車の前方映像を入力する。列車前方映像を撮影する際のカメラ位置

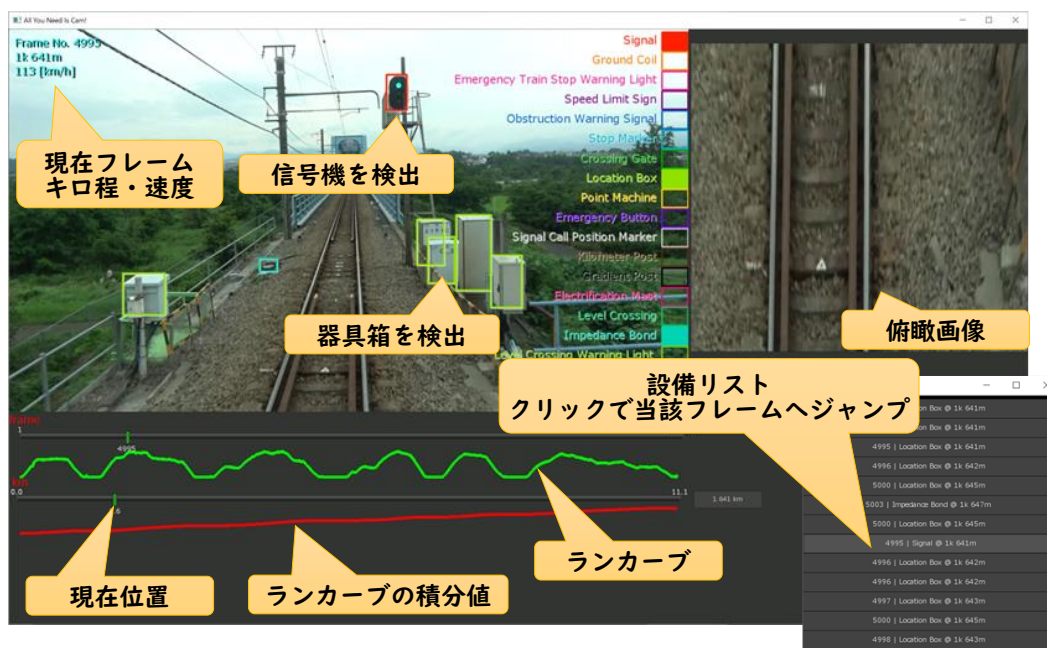


図1 システムの動作例

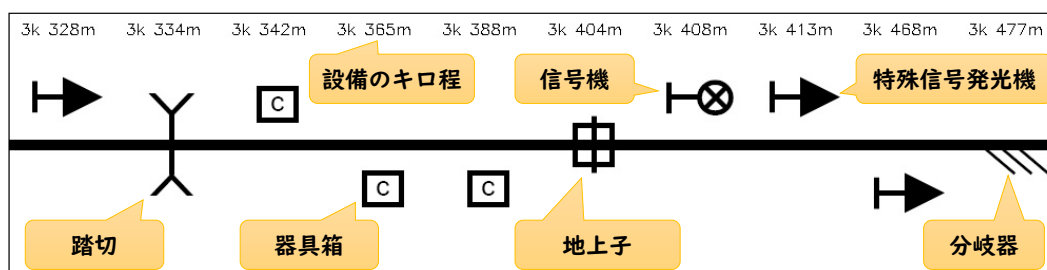


図2 システムで生成する簡易図面の例

は、運転台の助手席側や貫通扉の中央部とし、運転士から見えている風景を撮影する。

入力された列車前方映像に対して画像処理を施し、得られた情報をユーザに提示することで、メンテナンス業務の支援を行う。システムの動作例を図1に示す。システムは次の5つの機能を持つ。

(1) 設備の認識

深層学習によって各フレームから沿線設備を認識する。図1の動作例では、YOLOv3²⁾を用いて設備を認識しており、認識対象としている設備は信号設備を中心とした20設備である。

(2) 軌道面の俯瞰画像の生成

各フレームの軌道面を射影変換することで、軌道面を真上から見たような俯瞰画像に変換する。軌間 1,067mm とした既知のスケールを基準に俯瞰画像に対して1画素あたりの実寸法を計算することで、画像内での設備の距離を計測することも可能である。

(3) キロ程の推定

(2) 項で生成した俯瞰画像のフレーム間のオプティカルフローから各フレームのキロ程を推定する。具体的にはフレーム間のオプティカルフローを列車の走行速度とみな

し、映像全体で累積することで、走行距離を求める。この時、走行距離は画素を基準とした値であるため、撮影時の全体の走行距離や設備や駅の情報を与えることで画素に対する距離を割り当て、各フレームでキロ程を推定する。

(1) 項で述べた設備認識の結果と合わせることで、各設備の設置キロ程も推定できる。

(4) 簡易図面の作成

(3) 項で述べた各設備位置のキロ程から、図2に示すような簡易図面を生成する。画像中での設備の認識位置から、進行方向に対して左右どちらにあるか判定している。

(5) 設備の劣化度の推定

(1) 項で認識した設備について、深層学習を用いて設備の劣化度を推定する。推定手法の詳細については3章以降で述べる。

3. 深層学習による劣化度推定手法

前述した通り、列車前方映像に映る設備に対して、定量的な劣化度の数値を付与することは難しい。そこで、設備の劣化度を回帰問題として解くのではなく、距離を学習させるDMLを用いて推定する。

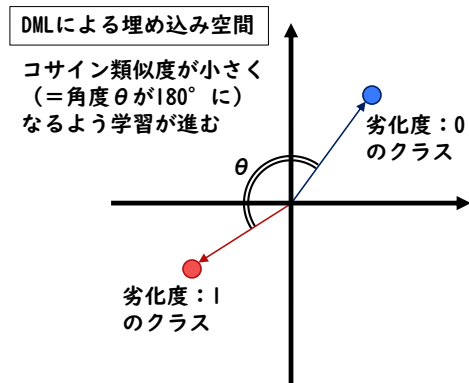


図3 DMLにより得られる埋め込み空間のイメージ図

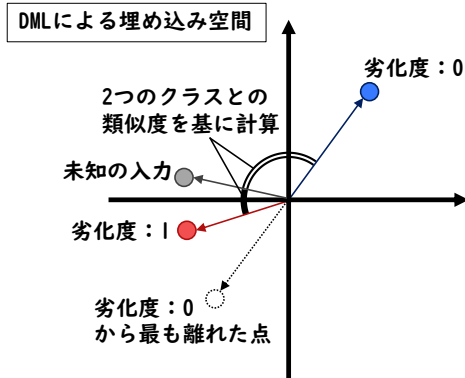


図4 劣化度の推定方法のイメージ図

3.1 Deep Metric Learning (DML)

Metric Learning とは入力空間におけるサンプル同士の類似度がコサイン類似度などの尺度と対応するような別空間に埋め込む変換を学習する方法である。変換後の別空間を埋め込み空間 (Embedding 空間) と呼び、埋め込み空間への変換を深層学習によって学習させるのが Deep Metric Learning である³⁻⁵⁾。DML は画像検索やクラスターリング、異常検知等さまざまなタスクに応用可能であり近年注目を集めている。

提案する劣化度推定手法では、設備が全く劣化していない状態 (劣化度: 0) と設備が非常に劣化している状態 (劣化度: 1) は教師データとして与えることが可能であることを前提とする。この 2 つのクラス対して、DML により埋め込み空間への変換を学習し、劣化度がわからない未知の入力の各状態への距離を求めることで劣化度を推定する。

3.2 劣化度推定手法

DML の手法として提案手法では L2 Constrain Softmax Loss⁵⁾を用いる。劣化度 0 と劣化度 1 の 2 つのクラスで学習した場合、2 つのクラスは埋め込み空間で式(1)に示すコサイン類似度が小さくなるように学習される (図 3)。

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \dots \dots \dots (1)$$

ここで、 $\mathbf{a}, \mathbf{b}$ は入力画像を埋め込み空間へ変換した際の特徴ベクトルであり、 $\cdot$ はベクトルの内積を、 $\|\mathbf{x}\|$ は L2 ノルムを表す。この時、学習対象のクラスが 2 つであるため、この 2 クラスを埋め込み空間で離そうとすると、2 つのクラス間のコサイン類似度が -1 (2 つのクラスの特徴ベクトル



図5 データセットの例

ルのなす角度が 180°) に近づくと考えられる。

未知の入力画像 $\mathbf{x}$ について、学習された埋め込み空間での劣化度 0 と劣化度 1 との距離を基に劣化度を推定する。入力画像 $\mathbf{x}$ を埋め込み空間に投影した特徴ベクトルを $\mathbf{x}'$ としたとき、埋め込み空間内で最も近い劣化度 0 と劣化度 1 の代表点 $\mathbf{d}_0, \mathbf{d}_1$ から劣化度 D を次のように定義する。

$$D(\mathbf{x}') = \frac{1}{4} \cos(\mathbf{x}', \mathbf{d}_1) - \frac{1}{4} \cos(\mathbf{x}', \mathbf{d}_0) + \frac{1}{2} \dots \dots \dots (2)$$

(2) 式中の 1/4 や 1/2 は劣化度を [0,1] に収めるための定数である。図 4 のように、DML による埋め込み空間において、劣化度 0 と劣化度 1 のクラスの代表点間のコサイン類似度は必ずしも -1 にはならないため、どちらかのクラスのみでの比較では劣化度を正しく表現できない可能性がある。そこで、2 つのクラス間の類似度を平均するようにして劣化度を推定している。

4. 器具箱を対象とした劣化度推定実験

前章で述べた劣化度推定手法について、器具箱を対象として DML の学習と劣化度の推定を実施した。

4.1 ネットワークの学習

本実験で使用したデータセットは器具箱を対象とした劣化度 0 : 518 枚、劣化度 1 : 160 枚の合計 678 枚である。このうち 8 割を学習データ、残り 2 割をテストデータとして用いた。データセットの例を図 5 に示す。劣化度の基準としては錆びている度合いを主観的に判断し、正解付けを実施した。

ネットワークアーキテクチャには WideResNet50_2⁶⁾ を、損失関数に L2 Constrain Softmax Loss⁵⁾ を用いた。L2 Constrain Softmax Loss の正則化パラメータは文献 5 を参考に $\alpha = 16$ としている。学習時には、ImageNet⁷⁾ によって事前学習されたモデルに対して器具箱を劣化度 0 と 1 で 2 クラス分類するようファインチューニングを実施した。

テストデータを学習後の埋め込み空間に変換し t-SNE⁸⁾ により可視化したものを図 6 に示す。劣化度 0 と 1 の 2 つのクラスタに分かれていることが確認できる。テストデータに対する劣化度 0 と 1 の分類精度は 98.5% であった。デ

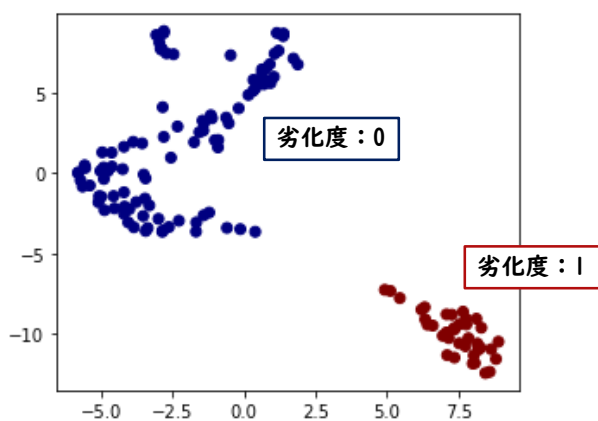


図6 埋め込み空間でのテストデータの可視化

ータセットが小さく、分類問題としては容易な問題であったため、高精度な結果になったと考えられる。

4.2 劣化度の推定結果および考察

学習用データセットとは別に、ある程度劣化しているものの、非常に劣化している（劣化度 1）とは言えないいくつかのデータについて劣化度の推定実験を行った。なお、劣化度の推定に必要な劣化度 0 および 1 の代表点は各クラスの学習データの埋め込み空間における平均値とした。

劣化度の推定結果を図 7 に示す。図 7 の結果から、劣化度 0 や 1 に分類したクラスでも劣化度が 0 や 1 にはならないことが分かる。これは今回のクラスの代表値が学習データの平均値となっているためである。代表値の決定方法については、学習データの埋め込み空間での分布などに基づいて複数持たせる改良が考えられる。

劣化度の判断が難しいデータについては、①のように塗料が剥がれていたり、②のように塗料が変色していたりするような画像は劣化度が 0.11、0.18 と低い結果となった。これは DML の学習の際に錆びの度合いを基準として劣化度 1 の画像を学習したためであると考えられる。また、③、④、⑤を比べると錆びている大きさはそれほど違いないように見えるが、錆びの色が⑤が最も濃く撮像されており、その結果⑤の方が劣化度 0.82 と高くなったと考えられる。

本実験では DML の学習に Softmax 関数を用いた損失関数を用いているため、2 つのクラスのどちらからも離れたデータは劣化度が 0.5 に近くなる。逆に、どちらかにある程度類似していれば、劣化度は 0 か 1 に近づくため、器具箱の劣化の度合いによらず、提案手法では劣化度 0.5 周辺の値が取りづらくなっている可能性がある。また、学習時

には劣化度 0 と 1 という極端に離れた 2 つのクラスしか用いていない。これらのことから、DML による埋め込み空間における距離に対して、劣化度が均一に割り当てられていない可能性がある。劣化度を均一に対応付ける方法としては、0 と 1 の間を示すクラスを追加した 4 クラス程度で、中間の状態を含めて距離を学習させることが考えられる。

5. まとめ

設備のメンテナンス業務を省力化させるための安価なシステムとしてハンディカメラと画像処理を用いたシステムを提案した。設備の劣化度を推定するために、DML を用いた手法を提案し、器具箱に対して適用した。

現状では、劣化度と DML の埋め込み空間での距離の対応が均一ではない可能性があり、今後はクラス数や学習データ数を増やして検証を進めていく。

参 考 文 献

- 1) Hiroki, Mukojima and Nozomi, Nagamine, "Recognition the installation place of signal equipment using one handy camera", the International Journal of Transport Development and Integration, Vol.4, No. 3, pp. 264–273, 2020.
- 2) Redmon, Joseph and Farhadi, Ali, "YOLOv3: An Incremental Improvement", arXiv:1804.02767, 2018.
- 3) Wang, Jiang, et al. "Learning fine-grained image similarity with deep ranking", Proc. of CVPR2014, pp.1386–1393, 2014.
- 4) Sohn, Kihyuk, "Improved Deep Metric Learning with Multi-class N-pair Loss Objective", Advances in NIPS2016, pp.1857–1865, 2016.
- 5) Ranjan, Rajeev, Carlos D. Castillo and Rama Chellappa, "L2-constrained softmax loss for discriminative face verification", arXiv:1703.09507, 2017.
- 6) Zagoruyko, Sergey, and Nikos Komodakis, "Wide residual networks", arXiv:1605.07146, 2016.
- 7) Deng, Jia, et al., "Imagenet: A large-scale hierarchical image database", Proc. of CVPR2009, pp.248–255, 2009.
- 8) Maaten, Laurens van der and Geoffrey Hinton. "Visualizing data using t-SNE", Journal of machine learning research, Vol.9, No. 86, pp. 2579–2605, 2008.

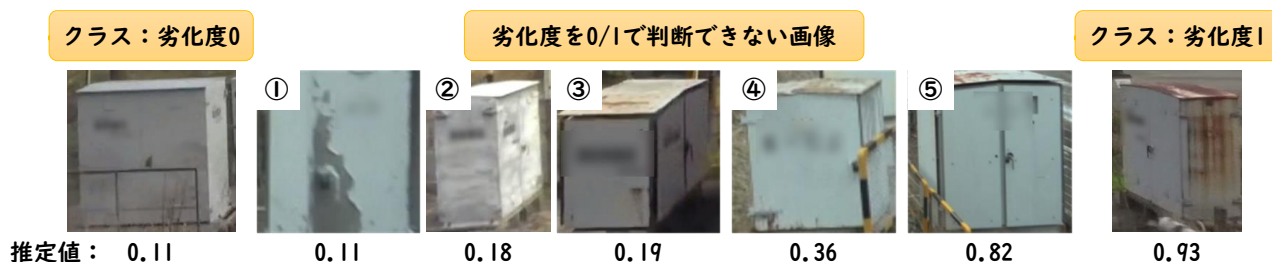


図7 劣化度の推定結果