

ディープラーニングを用いた 開削トンネルの健全度自動判定

大原 勇¹・仲山 貴司²・三輪 陽彦³・清水 達貴⁴

¹正会員 鉄道総合技術研究所 構造物技術研究部 (〒185-8540 東京都国分寺市光町 2-8-38)
E-mail: ohara.yu.98@rtri.or.jp

²正会員 鉄道総合技術研究所 構造物技術研究部 (〒185-8540 東京都国分寺市光町 2-8-38)
E-mail: nakayama.takashi.61@rtri.or.jp

³正会員 鉄道総合技術研究所 構造物技術研究部 (〒185-8540 東京都国分寺市光町 2-8-38)
E-mail: miwa.akhiko.53@rtri.or.jp

⁴正会員 鉄道総合技術研究所 構造物技術研究部 (〒185-8540 東京都国分寺市光町 2-8-38)
E-mail: shimizu.tatsuki.78@rtri.or.jp

日本の鉄道トンネルは70%以上が経年60年を超えており、安全性を確保するため、事業者により2年に1回の定期検査が行われ、熟練した検査員による健全度の判定等が行われている。ここで、健全度判定は経験的な判断基準に基づく部分が多く、検査結果にばらつきが生じる恐れがある。また、今後は少子化に伴い熟練検査員の確保が困難になることも予想される。筆者らはこうした状況に対応していくため、ディープラーニングを用いて変状を自動抽出したうえで健全度を自動判定する手法を検討している。本研究では、既存の変状写真を用いて学習済みモデルを作成するとともに、検査における実際の健全度判定を正答として精度検証を行ったところ、安全側の判定を許容すれば一定の精度が得られていることを確認した。

Key Words: railway tunnel, inspection, water leakage, soundness judgement, deep learning

1. はじめに

日本には約4,800本、延長3,700kmの鉄道トンネルがある。これらの70%以上が経年60年を超えており、延命化に対する対応が必要な時代に入っている。経年が進むなか、管理者により適切な検査が行われている。一般的には、日々の巡回のほか、定期的な検査（通常全般検査や特別全般検査）が義務付けられ、熟練した検査員による軌道面からの目視調査や高所作業車を用いた打音調査（図-1）を通じて、健全度の判定および要注意箇所の入念な確認が行われている。このように、トンネルの検査は目視や打音調査が主体であるため、検査員にかかる負担が大きい。また、健全度判定は経験的な判断基準に基づく部分が多く、検査結果にばらつきが生じる恐れがある。今後労働人口が減少する中で、検査業務をいかに効率化・高精度化するかが課題となっている。

こうした状況を受け、近年トンネル検査の効率化に向けたデジタル技術の開発が進められている。具体的には、図-2に示すような覆工内面の連続画像の撮影技術^(例えば1)や、機械学習を用いたひび割れ抽出技術^(例えば2)などがあ

る。このように、検査のうち調査を補助する技術については開発され、実現場への導入が進められている。しかしながら、調査結果を踏まえた性能の確認（≒健全度の判定）については、既存技術で実用化に到達していない。

そこで本研究では、ディープラーニングを用いて覆工の連続画像から変状を自動抽出するとともに、抽出した変状の種類や範囲などに基づき健全度を自動判定する手法を検討している³⁾。本紙では、既存の変状写真から学習モデルを作成し、実際の健全度と比較検証を試行した結果について報告する。



(目視調査) (打音調査)

図-1 鉄道トンネルの検査の様子

2. 検討の流れ

検討の流れを図-3に示す。本研究では、健全度自動判定を2ステップに分けて実施した。まずステップ1として、ディープラーニングによる学習を行い、画像から変状を自動抽出する学習済みモデルを構築した。この際、教師データの数量やハイパーパラメータのチューニングを行い、変状抽出の精度向上を図った。次にステップ2として、ステップ1で抽出した変状の種類や範囲などと健全度の関係をまとめたマトリクス（以下、判定マトリクス）を作成し、これを用いて健全度を自動判定し、実際の健全度と比較することで精度を検証した。

(1) ディープラーニングによる変状抽出

ディープラーニングを用い、物体検出により画像から変状を抽出する学習済みモデルを構築した。物体検出とは、ある画像に含まれる対象物体の位置を矩形のボックス（バウンディングボックス）で特定するとともに、その物体のクラス（種類）を識別する手法である（図-4）。着目する変状のクラスについては、十分な教師データが確保できること、健全度に与える影響が大きいと考えられることを踏まえ、エフロレンス（以下、エフロ）、漏水、錆汁に着目した。

ディープラーニングのフレームワークには Tensorflow Object Detection API を使用した。初期のネットワークモデルには CoCo データセットを事前学習した EfficientDet(D0) を使用し、後述する教師データを転移学習させた。物体検出では検出ボックスごとに判定の信頼性を示すスコアが出力され、事前に設定した閾値（パススコア）を超える検出ボックスを有効とするが（図-4）、パススコアについては良好な結果が得られるよう目視で調整し、本研究では20%を採用した。なおハイパーパラメータについては、チューニングを行ったうえで最良と考えられるものを選定した（3章で後述）。

教師データについては、実際の鉄道トンネルの変状のスナップ写真から適当な画像（明確な変状を捉えたもの）を選定し、各画像に変状の位置とクラスの情報を付与することで作成した。教師データの数量は表-1に示す通りである。また、学習時にランダムで画像処理（水平反転、鉛直反転、サイズ変更）を施しデータ数量を増強した。

テストデータには、ある鉄道事業者2社（事業者X、Y）で実際に撮影された覆工連続画像を使用した。なお、ここで用いた覆工連続画像は、トンネル軸方向に10mスパンで区切られたものであり、開削トンネル区間を対象に、事業者Xから215枚、事業者Yから102枚（計317枚）の画像を収集した。

(2) 健全度自動判定

上述の変状抽出結果を判定マトリクス（表-2）に当てはめることで健全度を自動判定した。ここで、判定マトリクスは鉄道構造物等維持管理標準（トンネル）⁴を参考に、比較的高い専門性が求められるトンネル構造の安定性に着目して、やや安全側の判定となるよう作成した。また、外力パターン（トンネルの変形・沈下・移動時に

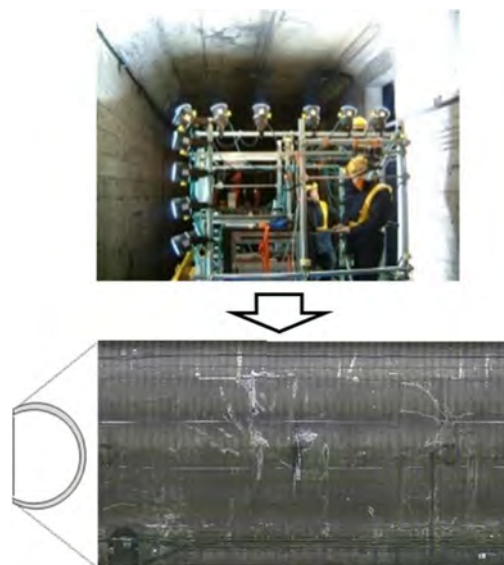


図-2 覆工内面の連続撮影画像

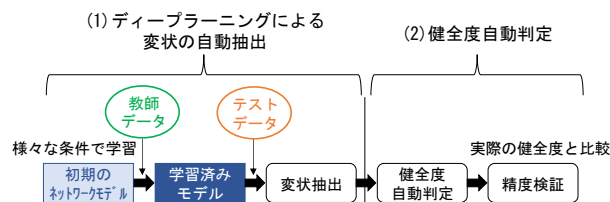


図-3 検討の流れ

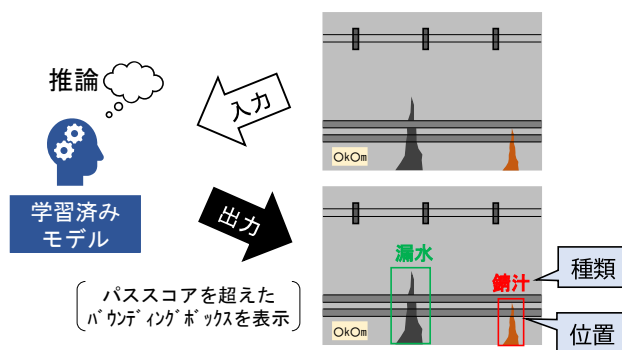


図-4 物体検出

表-1 教師データ数量（増強前）

| 変状種別 | 数量（枚） |
|------|-------|
| 錆汁 | 225 |
| 漏水 | 505 |
| エフロ | 230 |

表-2 判定マトリクス

| 性能項目 | 判定指標・方法 | | | 部材の劣化 | | | | |
|------------|---------|---------------------|----------------------|--------------------|-------|-------|--------|--------|
| | | | | 鉄筋露出 ^{※2} | 錆汁が併発 | 漏水が併発 | エフ・漏水跡 | ひび割れのみ |
| トンネル構造の安定性 | ひび割れ | 下記以外 | 下記以外 | A | B | C | C | S |
| | | | 外力パターン ^{※1} | A | A | B | C | C |
| | | 新規・進展 ^{※3} | 下記以外 | AA or A | | A | | |
| | | | 外力パターン ^{※1} | AA or A | | | | |

※1：トンネルの変形・沈下・移動の際に見られる特徴的なひび割れパターン。今回は検出ボックスの比率から該当の有無を推定

※2：300mm×300mm以上，教師データに含まれない変状のため今回は対象外とした

※3：今回は一時期の画像を対象に試行したため，全て上段の「下記以外」として取り扱った

見られる特徴的なひび割れパターン) については，健全度が悪い場合に変状が密集する傾向にあることに着目し，検出ボックスが画像の延長方向に占める比率が閾値(25%)を超えた場合に外力パターンに該当させることとした。

3. 学習済みモデルの作成

本研究では，既存の変状写真のアノテーションとともに，学習時にハイパーパラメータをチューニングしながら学習済みモデルを作成した。なお，Tensorflow Object Detection API は自動チューニング機能を有していないため，ハイパーパラメータについては，表-3の範囲で手動で変化させた。

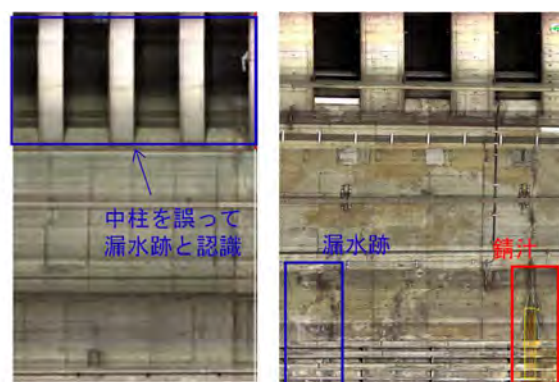
図-5に検出結果の一例を示す。図中の赤枠，青枠は学習済みモデルの検出結果を示しており，アノテーションの精度やハイパーパラメータの設定によって，得られた学習済みモデルの検出精度が異なる。図-5(a)は，比較的精度の悪いモデルの検出例であり，中柱を誤って漏水跡と認識していることがわかる。一方，図-5(b)は，比較的精度の良いモデルを用いたの検出例であり，錆汁や漏水跡を正しく検出できていることがわかる。学習済みモデルの良し悪しは mAP などの指標で評価されるのが一般的であるが，本研究ではこれに加えて実際に覆工連続画像から変状を検出させた結果を確認しながら，良好な検出が得られるアノテーション位置やハイパーパラメータの調整を行った。

この結果得られた最良と考えられる学習済みモデルの学習曲線を図-6に示す。学習曲線は学習済みモデルの精度のほか，過学習の発生を評価するために用いられるグラフであり，一般的に学習ステップが進むにつれて validation loss が小さくなっていく傾向を示すと良い学習とされている。今回の学習では，validation loss が下限値を取ったあとの上昇傾向をある程度抑制できたものの，下限値の低下には至らなかった。そのため，学習済みモ

表-3 学習条件

| | チューニング範囲 | 採用値 |
|------------------------------------|--------------------------|-------------|
| 最適化アルゴリズム | Momentum | Momentum |
| Batch size [※] | 2~8 | 8 |
| Learning rate | Base | 0.05 |
| | Warmup rate | 0.0009 |
| | Warmup step [※] | 3,000~5,000 |
| L2 regularizer weight [※] | 0.000001~0.5 | 0.001 |
| Step [※] | 5,000~500,000 | 100,000 |

※チューニングしたパラメータ



(a) 精度の悪い学習済みモデルの場合 (誤った検出例) (b) 精度の良い学習済みモデルの場合 (正しい検出例)

図-5 検出結果の例

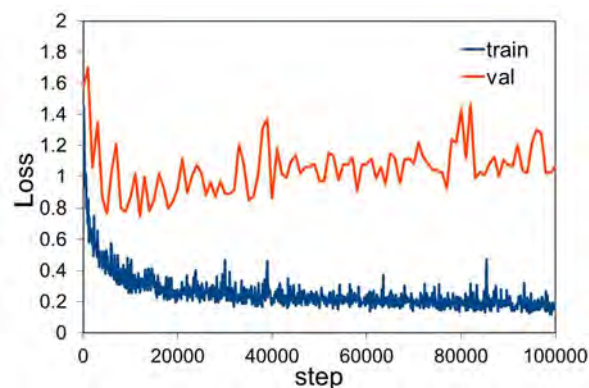


図-6 学習曲線 (最適ケース)

デルの改良の余地は残るものの、まずこの学習済みモデルで得られる健全度判定の精度を検証することとした。

4. 健全度自動判定結果

学習済みモデルを用いて変状抽出および健全度自動判定を実施し、実際の健全度と比較した。なお推論精度の評価指標としては、一般的に用いられるPrecision, Recall, Accuracyを採用した。Precisionは「ある推論がどれだけ当たっているか」、Recallは「ある正答に対してどれだけ正しく推論できたか」、Accuracyは「推論結果全体がどの程度当たっているか」を評価する指標であり、計算式は表-4および式(1)~(3)に示す通りである。

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

表-5に既往の検討³⁾における精度検証結果を示す。表-5より、健全度自動判定において一定の精度が得られていることが確認できる。ただし、ここでいう「正答」とは人の目で抽出した変状に判定マトリクスを適用して作成したものであり、実際の健全度ではない。

そこで本研究では、実際の検査における健全度判定を正答として精度検証を行った。精度検証結果を表-6に示す。全体としてはAccuracyが30%程度、安全側の判定を許容した場合のAccuracyが60%程度となっており、一定の精度で健全度を自動判定できていると考えられる。

また、事業者Xでは健全度Aの画像2枚がそれぞれB, Sと危険側の判定となった。これらの変状は新規・進展の変状であったため、検査において健全度Aと判定されていた。一方で、現状の自動判定ではある一時期の画像を対象に健全度判定が行われており、変状の進展性については対象としていないため、危険側の判定になったと考えられる。

事業者間の比較では、YはXよりも全体的に精度が低くなっているが、この原因としては、発生している変状の傾向・数量や補修・補強状況といった覆工の状況が事業者ごとに異なることが考えられる。このことから、今後健全度自動判定を実際の検査に活用していく上では、適宜事業者ごとに学習済みモデルや判定マトリクスを調整する必要性が示唆された。

5. まとめ

本研究では、開削トンネルを対象としてディープラーニングを用いた健全度の自動判定を実施した。その際、ハイパーパラメータのチューニングを行うとともに、検査における実際の健全度判定を正答として精度検証を行い、適用性を検討した。以下に得られた知見を示す。

[1] 今回の検討範囲内ではチューニングの効果は限定的であった。今後、教師データの追加等、別の観点からの改善が必要と思われる。

表-4 混同行列 (健全度Aのみの場合)

| | | 正答 | |
|------|-----|----|-----|
| | | A | A以外 |
| 自動判定 | A | TP | FP |
| | A以外 | FN | TN |

表-5 健全度自動判定結果の精度検証 (既往の検討³⁾)

| | | 正答 ^{※1} | | |
|------|-------|------------------|----|----|
| | | AAorA | B | C |
| 自動判定 | AAorA | 5 | 2 | 0 |
| | B | 0 | 32 | 2 |
| | C | 0 | 5 | 36 |
| | S | 0 | 0 | 3 |

※1：人の目で抽出した変状に健全度自動判定を適用した結果

※2：青字：安全側の判定，赤字：危険側の判定

表-6 健全度自動判定結果の精度検証

| 事業者X | | 正答 ^{※1} | | | | Precision |
|--------|---|------------------|------|------|------|--------------------|
| | | A | B | C | S | |
| 自動判定 | A | 0 | 2 | 5 | 0 | 0.00 |
| | B | 1 | 51 | 56 | 4 | 0.46 |
| | C | 0 | 7 | 25 | 3 | 0.71 |
| | S | 1 | 11 | 47 | 2 | 0.03 |
| Recall | | 0.00 | 0.72 | 0.19 | 0.22 | 0.36 ^{※3} |
| | | | | | | 0.69 ^{※4} |

| 事業者Y | | 正答 ^{※1} | | | | Precision |
|--------|---|------------------|------|------|------|--------------------|
| | | A | B | C | S | |
| 自動判定 | A | 0 | 0 | 1 | 4 | 0.00 |
| | B | 0 | 8 | 21 | 9 | 0.21 |
| | C | 0 | 1 | 13 | 3 | 0.76 |
| | D | 0 | 8 | 34 | 0 | 0.00 |
| Recall | | 0.00 | 0.47 | 0.19 | 0.00 | 0.21 ^{※3} |
| | | | | | | 0.58 ^{※4} |

※1：検査における実際の健全度判定

※2：青字：安全側の判定，赤字：危険側の判定

※3：Accuracy

※4：安全側の判定を許容した場合のAccuracy

- [2] 健全度自動判定の結果を検査における実際の健全度判定と比較したところ、安全側の判定を許容すれば一定の精度が得られていることを確認した。
- [3] 事業者によって自動判定の精度に差があることが確認された。この原因としては、発生している変状の傾向・数量や補修・補強状況等が事業者ごとに異なることが考えられる。このことから、今後健全度自動判定を活用していく上では、適宜事業者ごとに学習済みモデルや判定マトリクスを調整する必要性が示唆された。

謝辞：本研究は国土交通省交通運輸技術開発推進制度 (JP002223) により実施しました。

参考文献

- 1) 全邦釘, 井後敦史: Random Forest によるコンクリート表面ひび割れの検出, 土木学会論文集 F3(土木情報学), Vol. 71, No. 2, pp. I_1-I_8, 2015.
- 2) 海瀬忍, 伊藤哲男, 前田佳克, 八木弘, 前田洗樹, 進士正人: トンネル覆工表面画像撮影技術による近接目視点検の代替の可能性, 土木学会論文集 F1(トンネル工学), Vol. 76, No. 1, pp. 62-78, 2020.
- 3) 大原勇, 仲山貴司: ディープラーニングを用いた開削トンネルの健全度自動判定の試行, 第 77 回土木学会年次学術講演会, 2022.9
- 4) (公財) 鉄道総合技術研究所: 鉄道構造物等維持管理標準・同解説 (構造物編) トンネル, 丸善出版, 2019.

AUTOMATIC SOUNDNESS JUDGEMENT ON OPEN CUT TUNNELS BASED ON DEEP LEARNING

Yu OHARA, Takashi NAKAYAMA, Akihiko MIWA and Tatsuki SHIMIZU

More than 70% of Japan's railroad tunnels are over 60 years old. To ensure safety, regular inspections once every two years are obligated by law and soundness judgement is conducted by skilled inspectors. However, the soundness judgment is largely based on qualitative criteria, and this could increase variability of judgements. In addition, it is expected to be more difficult to secure skilled inspectors in the future due to the declining birthrate. Therefore, we used deep learning to automatically judge the soundness of open-cut tunnels for the purpose of improving efficiency of the inspections. In this study, the accuracy of automatic judgement verified by using the actual soundness judgement in the inspection and the accuracy was confirmed to be approximately 60% if the safe side judgement was treated as correct.