

インフラマネジメントシステムへの AI 実装に向けた実証的研究

宮崎 文平¹・前田 紘弥²

¹ 法人正会員 三菱総合研究所 次世代インフラ事業本部（〒141-0031 東京都千代田区永田町 2-10-3）

E-mail: bumiyaza@mri.co.jp

² 非会員 東京大学 生産技術研究所（〒153-8505 東京都目黒区駒場 4-6-1）

E-mail: maedahi@iis.u-tokyo.ac.jp

社会基盤施設が団塊的な老朽化を迎える中で、点検や診断を補助する AI 技術に注目が集まっている。近年、国内外において、産官学様々な分野で AI 技術の開発が進められているが、インフラ管理の実態が AI 技術を適用する上で必要となる条件と整合していなかったり、AI 技術を調達する仕組みの整備が不十分であったりして、本格的な社会実装に至っていない。

本研究においては、具体的橋梁点検データに対して、AI 技術を適用し、その効果検証を図るとともに、インフラマネジメントシステムにおいて、AI 技術の活用を促進する上での課題を整理し、新たな調達の仕組みについて提案を行うものである。

キーワード：アセットマネジメントシステム、AI、調達

1. はじめに

高度経済成長期に大量に建設された社会基盤施設が団塊的な老朽化を迎えている。老朽化したインフラの増加に伴い、点検に対する需要は年々増大する一方で、専門知識を持った熟練技術者の減少などの問題が顕在化しており、より効率的な維持管理・更新を実施するための技術開発が進んでいる。

上述した課題の解決策の一つとして、点検や診断を補助する AI 技術に注目が集まっている。近年、国内外において、産官学様々な分野で AI 技術の開発が進められているが、インフラ管理の実態が AI 技術を適用する上で必要となる条件と整合していなかったり、AI 技術を調達する仕組みの整備が不十分であったりして、本格的な社会実装に至っていない。

本研究においては、具体的橋梁点検データに対して、AI 技術を適用し、その効果検証を図るとともに、インフラマネジメントシステムにおいて、AI 技術の活用を促進する上での課題を整理し、新たな調達の仕組みについて提案を行うものである。なお、ここで述べている AI は特定の用途に特化した特化型 AI であり、橋梁点検等のインフラマネジメントシステムにおいて、画像処理や機械学習手法を用いることにより人間の作業を補助・自動化するものを指す。

2. 国内における AI の活用状況

平成 29 年度に策定された「第 4 期国土交通省技術基本計画」では、計画の 3 つの柱として、「人を主役とする IoT, AI, ビッグデータ等の活用」が位置づけられており、IoT, AI, ビッグデータ等の導入により、人を主役とすることで「人の力」を高め、「新たな価値」を創造するとしている。

平成 30 年度より、国土交通省では、AI を活用した構造物点検などを目指し、AI の研究開発を促進するために、土木技術者の正しい判断を基にした良質な「教師データ」を国が整備・提供し、開発された AI の性能評価を行う取り組みとして「AI 開発支援プラットフォーム」の設立を検討している。この AI 開発支援プラットフォームにおいては、維持管理の生産性向上を実現する上で、「人の作業」の支援のみならず、「人の判断」の支援が重要と位置付けられている。

しかしながら、現状においては、インフラ管理者側で十分な量のデータ整備がなされておらず AI 開発者がアクセス、利用できるデータは限られている。一部の先進的な組織では、データの蓄積が進んでいるが、カメラの画角や解像度など、データの質の面で、AI の利用に適していないデータセットであるケースも多い。

また、AI 開発者側で、AI による評価・診断を行えた

としても、その結果を正しく評価・判断できる技術者がインフラ管理者側に存在しないため、導入が進まない現状も見受けられる。とりわけ、こうした AI 活用による評価・診断の効率化のニーズは、経営資源が限られている自治体でこそ高いと考えられるが、実態としては、高速道路会社などにおける取組が中心になっており、自治体レベルでの取組は進んでいない。

自治体のインフラマネジメントへ AI 技術の導入・実装を進めていくためには、AI 活用による効果を定量的に評価するとともに、その適用限界を示し、既存のインフラマネジメントシステムを AI 活用を前提として、見直していかなければならない。そこで、本研究においては、具体的橋梁点検データに対して、AI 技術を適用しその精度評価を行うとともに、用意したデータやモデル、構築環境等に関する課題を抽出し、今後の活用に向けた提案を行うものとする。

3. 実証分析

(1) データセットの概要

本研究で用いるデータセットは、米国において主に橋梁床板の点検を実施している NEXCO-West USA, Inc 社（以下「NWU 社」）から提供頂いた。米国の橋梁上面は、日本と異なりアスファルト舗装が施されていないケースがほとんどであり、コンクリート床板が剥き出しになっている。NWU 社では DTSS (Deck Top Scanning System) と呼ばれる独自のシステムを採用している。同システムは、赤外線カメラとラインカメラと呼ばれる高精度カメラを車両に搭載し、最大 70mile/h の速さで橋梁上を走行しながら、橋梁床板の画像を取得する。赤外線カメラでは、床板の温度差を取得し、Delamination と呼ばれる床板の浮きを検出する。一方のラインカメラでは、0.3mm 以上のひび割れの検出が可能である。NWU 社は、撮影した画像に対して、損傷箇所をマーキングし、州 DOT (Department of Transportation) やエンジニアリング会社に提出している。

本研究においては、このラインカメラにより撮影した高解像度画像を利用する。具体的に、NWU 社が過去に撮影した 1 橋梁（片側 2 車線、橋長約 4,000m）分の床板画像から、深層学習モデルの入力として、1 フィートに相当する 256×256 ピクセルの画像（以下「チャンク」）を切り出した。チャンク数は合計で 3,984 枚となった。

米国においては、部材別の橋梁点検が義務付けられているが、具体的な基準は、米国全州道路交通運輸行政官協会 (American Association of State Highway and Transportation Officials 「以下 AASHTO」) が発刊する Manuals for Bridge Element Inspection¹⁾に定められている。同マニュアルの中では、橋梁点検において、検出すべき損傷が、その損傷

表-1 対象データ

損傷名	チャンク枚数
ひび割れ (Crack)	493
剥離 (Spall)	29
パッチ (Patch)	4
鉄筋露出 (Exposed Rebar)	11
損傷なし	3,447

コードとあわせて体系的に整備されている。各損傷は、その程度に応じて、1 から 4 の 4 段階で健全度 (Condition State) が判断される。例えば RC 床板のひび割れに関しては、0.012inch から 0.05inch が健全度 2、0.05inch 以上が健全度 3 と定められている（健全度 1 は損傷無し、健全度 4 は直ちに橋梁を閉鎖する余裕がある緊急性の高いものであり具体的な数値基準は存在しない）。NWU 社の通常業務では、この AASHTO のマニュアルに従って、損傷を検出しており、本研究においても、マニュアルに定められているカテゴリごとにデータを整理した。カテゴリごとチャンク数の内訳は、表-1 の通りである。表-1 より、剥離やパッチ、鉄筋露出のチャンク数が少なく、ひび割れのチャンクがほとんどであることが読み取れる。一般に、深層学習モデルの学習においては、十分な判定精度を担保するために、各カテゴリごとに 1,000 から 5,000 枚程度の画像を教師データとして用意することが望ましい²⁾と言われており、ひび割れを除く画像の蓄積は十分ではなかった。そこで、本研究においては、チャンクの中にひび割れがあるか、ないかの 2 択を判別するに留めることとした。ひび割れ (Crack) と損傷なし画像をそれぞれ 6:2:2 の割合でランダムに分割し、学習・モデルの評価・テストに使用した。本論文で言及する結果は K=3 のクロスバリデーションを実施した結果である。

(2) 適用モデル

(1)で説明したデータに対して、深層学習モデルの学習を行った。本研究では、ResNet-50³⁾を学習するニューラルネットワークとして使用している。事前学習済みのモデルは使用せず、本研究で用意したチャンクのみを用いてモデルを学習させた。

ここでは、ピクセルごとにラベルをつけるアプローチ (Segmentation) ではなく、一定の大きさのチャンクにラベルをつけるアプローチ (Classification) を採用していることに注意されたい。実際、AASHTO のマニュアルにおいて、ひび割れを 1 本 1 本描画することは求められていない。納品物となる損傷図に求めるアウトプットは州によって考え方が異なり、プロジェクトや対象としている橋梁によってもカスタマイズが必要である。厳格な州では、ひび割れ展開図を求めるが、多くの州では、すぐに補修のアクションが必要な、損傷の進展が著しい範囲

のみを大きく枠囲みして特定して欲しい、というニーズがほとんどである。Classificationモデルを導入する場合、教師データはひび割れをピクセルレベルでアノテーションする必要があり、膨大な作業が発生するのに加え、計算機のコストも大きくなる一方で、顧客の要求水準をはるかに上回る。このような理由で、チャンクごとに損傷の有無を判定する Classification モデルを構築した。

学習に要する時間は、Amazon EC2 の p2.xlarge と呼ばれる GPU を搭載したインスタンスを用いた場合、5時間程度であった。学習には膨大なパラメータを更新するために膨大な計算機コストがかかるが、予測（新規画像を入力して、予測ラベルを得る）には CPU インスタンスで十分であることが多い。

(3) 適用結果

チャンク画像の分類モデルの評価結果を表-2 に示す。ひび割れ検出の再現率は $0.9(=90/(90+10))$ 、適合率は $0.3(=90/(90+222))$ となった。

図-1 は、ひび割れ分類モデルによる判定の結果、False Negative（見逃し）となった画像である。左側の画像は、ひび割れが明瞭でなく、AASHTO の基準に照らしても判断が難しい損傷であった。また、右側の画像には、タールと呼ばれる汚れが写り込んでおり、誤った特徴量を認識し、誤判定につながったと考えられる。この場合には、同じような汚れが写り込んでいる画像を複数枚収集し、分類モデルのカテゴリに追加することで、見逃しを減らすことが可能である。

また、深層学習による画像分類では、分類するカテゴリごとの確信度を出力することが一般的であり、本研究でも確信度を閾値として、モデルの出力結果を変更しながら検討を行った。例えば、閾値を 0.1 とした場合、False Negative（見逃し）が 10 枚から 18 枚に増えるが、False Positive（誤検出）は 224 枚から 124 枚となり、両者の関係はトレード・オフにある。後述するが、False Negative と False Positive、それぞれのエラーへの対応プロセスは大きく異なり、実際の閾値の設定に関しては、分類モデルを活用する事業者がそれぞれのケースに応じて、設定を変更し、業務改善につなげていくことが重要である。

4. 実装に向けた課題と対応策

前章で、具体の点検データに対して、AI を適用することにより明らかとなった課題は主に以下の通りである。

- ・画像中に汚れや落下物、白線、陰などが存在した場合の誤検知
- ・画質の一貫性のなさ（距離、画角、解像度など）
- ・False Positive と False Negative への対応

表-2 ひび割れ分類モデルの評価結果

		モデルの予測	
		ひび割れ	損傷なし
正解	ひび割れ	90	10
	損傷なし	222	461

※確信度 0.5 を閾値とした場合

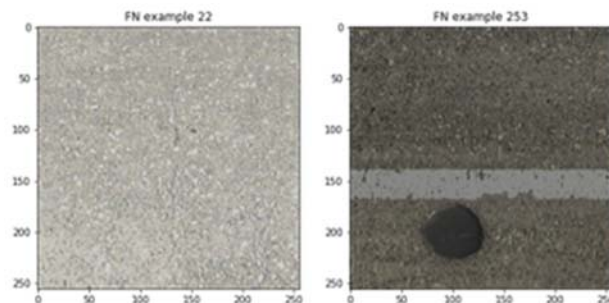
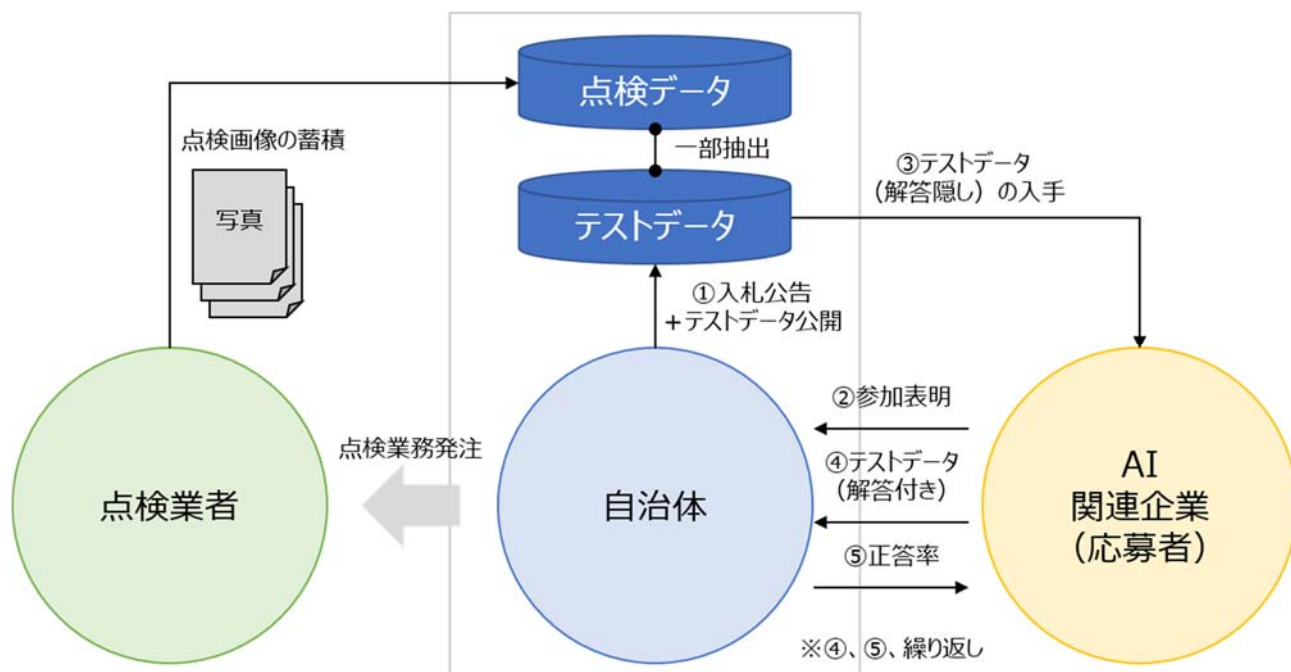


図-1 False Negative（見逃し）の例

一つ目に、画像中に検出対象以外の物体が写り込んでしまう場合に、モデルの精度が低下している。例えば、今回のモデルの場合、チャンクにタール（汚れ）が写り込んでいる場合に、ひび割れを検出することができず、False Negative が発生した。先述の通り、このような場合に対しては、新たなカテゴリをモデルの出力に追加し、教師データを収集して対応できることが望ましい。しかしながら、現状の維持管理においては、損傷箇所のみを撮影しているケースが多く、汚れや白線、日陰などのデータも個別に取得しているケースは少ないと考えられる。国や研究機関が、モデルの精度に影響を与える要因を整理して公表し、各自治体にデータの蓄積を促していく必要がある。また、2章で述べた「AI 開発支援プラットフォーム」でも取組みがはじまっているように、国として大量の写真データを教師データとして整備して公開したり、自治体間でもデータを共有する仕組みを構築したりするなどして、不足分を補っていく必要がある。

二つ目に、画質の一貫性のなさの問題がある。本研究で使用した画像データは、橋梁床版を車両上に固定搭載したラインカメラで撮影したものであるため、画質や画角は一貫していた。一方で、日本国内の橋梁は、床版上面にアスファルト舗装が敷かれていることもあり、橋梁上部からではなく、下部から目視点検を実施し、損傷を発見した場合に、当該損傷箇所を撮影して、管理者に提出することが一般的である。この場合、損傷箇所に対する撮影距離や画角が一定ではなく、自動化アルゴリズムの構築が難しい。また、損傷箇所のみを撮影しているため、損傷がない場合の教師データを作ることが難しいという問題もある。

モデルの精度は、画像の質に大きく依存するため、点



検査業務を発注する自治体は、発注仕様書の中で、写真の撮影方法や納品時の分類方法に対して、より具体的かつ詳細な要求をしていくことが望ましいと考えられる。一方で、過度な要求をし過ぎた場合、点検業者側の負担が大きくなり、効率化につながらないという問題もある。点検業務と分析業務のトータルで最適化が図れるよう、自治体、点検業者、分析業者（AI 関連業者）を交えて議論を深めていく必要がある。

三つ目に、分析実施後の False Negative と False Positive への対応がある。False Positive については、自動検出された範囲の中に誤った検出結果が含まれていることを示す。これについては、自動検出が実施された後、点検に関する専門知識を持ったエンジニアがその結果を目視チェックするプロセスを組み込むことで対応可能である。目視チェックして判定を見直したデータは、また新たな教師データとして活用していくことが出来る。

一方で、False Negative については対応が難しい。False Negative は、そもそも損傷として検出されていない画像であるため、False Negative をなくす作業を行うことは橋梁全画像をチェックすることに等しく、目視による分析との重複作業に他ならない。先述した通り、閾値を調整することにより、False Positive の数を増やして、False Negative を強制的に減らすことは可能であるが、ゼロにすることは不可能である。インフラの点検、評価、診断のプロセスにおいて、AI 導入が進まない最たる要因は、この False Negative の可能性を排除できないためであり、分析の中身はブラックボックスであるから、結果として、従来通りの目視による作業が対外的に明確な説明も行い

やすく、採用され続けているのが現状である。

しかしながら、実際には、目視によるチェックの結果にも誤りが含まれている可能性は否定できない。目視によるチェックの結果と AI によるチェックの結果の時間・精度の違いについては今後も研究の蓄積が不可欠であるが、ステークホルダーである市民へのアカウントビリティを果たしていくためには、目視による分析か AI による分析かに関わらず、成果物に誤りや見通しがないよう、品質を担保するための仕組みを既存のマネジメントプロセスの中に組み込み、分かりやすい形で説明できる必要がある。

例えば、AI を適用し、損傷として抽出されなかった範囲 (False Negative が含まれている可能性のある範囲) に対しても、ランダムにその数%を抽出し、目視による確認を行うなどの仕組みを入れることで、False Positive を減らすことは可能である。また、橋梁でいえば桁中央部など損傷が発生しやすい箇所、また損傷が発生した場合クリティカルな場所は、構造的に明らかになっているため、そういった箇所に部分的な目視チェックを導入することで、品質を高めることが可能である。すべてを AI で置き換えるのではなく、人を介在させながら機能する仕組みを構築することが重要である。

5. 新たな調達制度に関する提案

平成26年、トンネルや橋梁に対しては5年に1回の近接目視点検が義務付けられた。このことは、各道路管理者にとっては、大きな負担となっている。今後、2巡目、

3 巡目の点検を行っていく中で、点検の重点化や効率化を図っていく必要がある。

平成 30 年の定期点検要領の改定では、「自らが近接目視による時と同等の健全性の診断を行うことが出来ると定期点検を行う者が判断した場合には、その他の方法についても近接目視を基本とする範囲と考えてよい。」とされた。近接目視が絶対ではなくなり、今後ドローンや AI などの新しい技術の導入を検討しやすい環境が整ってきたといえる。しかしながら、自治体側には、新しい技術に精通している職員がおらず、導入を躊躇してしまふケースは少なくない。AI 関連業者などから個別に技術提案をもらっても、その技術の妥当性を判断することが難しい。

そこで、図-2 に新しい調達フロー（案）を示す。点検業務の発注を検討する自治体は、入札公告とあわせて、過去の点検業務で蓄積した今回の点検業務の対象構造物に関するテストデータを公開する。点検データは、すべてアノテーションされたものであるが、テストデータとして利用する際には、解答（カテゴリごとの損傷の有無）は伏せた状態で公開する。応札を希望する業者は、参加表明とあわせて、同テストデータ入手し、自社のモデルを利用して分析を行う。分析が完了した場合は、自治体に対して、解答付きのテストデータを送付し、自治体から正答率のみを通知される。業者側では、その結果を受けて、アルゴリズムの改善を図る、というものである。

このプロセスを繰り返し、最終的にある一定水準の精度を満たした AI 関連業者にのみ、入札参加資格が付与される。こうすることで、自治体側は、自組織が管理する構造物の画像に対して最も適したモデルを選択することができ、また、国内外に現存する AI による分析の精度の限界も把握することが出来る。そのため、成果物に対してどのような検収を行う必要があるのか、現状の点検データの蓄積方法に関し、今後どのような修正が必要であるかなどの改善に向けたアクションを検討することが可能となる。

なお、図-2 のフローでは、業務発注ごとにデータコ

ンペティションを実施し、その都度、基準を満たす業者を抽出する流れを想定しているが、将来的には、段階的な認証制度を設けていくことが望ましいと考えられる。具体的には、各自治体で定期的に同様のデータコンペティションを実施し、業者の精度に応じて、認証を実施、今後の業務発注時には、応募要件として、一定のレベルの認証取得を要求するような形で、手続きを簡略化することが可能である（現状では、点検業務の中に分析業務も含まれているケースがほとんどであるが、AI 活用を前提とする場合には、より参入の障壁が下がるように、分析部分のみを切り出して業務発注することも検討していく必要があると考えられる）。

6. おわりに

本稿では、橋梁床板の点検データに AI 技術を適用することで、その効果を検証するとともに、今後 AI 技術を活用する上での課題を抽出し、対応策の検討を行った。AI による分析の精度を向上させるためには、データの量と質を高めていくことが不可欠であるが、自治体が単独で取り組むには非常に負荷が大きいため、国や研究機関、近隣の自治体と連携して進めていく必要がある。また、新たに提案した調達制度・認証制度に関しても、国レベルで、制度のあり方について検証するとともに、特定の自治体やプロジェクトなどで、試行的な取組をはじめていくことが重要と考えられる。

参考文献

- 1) Manual for Bridge Element Inspection, 2nd Edition, AASHTO, 2019
- 2) Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron. Deep learning. MIT press, 2016.
- 3) He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.