[共通セッション] 土木分野におけるAIの活用

2024年9月6日(金) 10:40 ~ 12:00 第2講義室(経済学部)(川内南キャンパス文科系総合講義棟)

土木分野におけるAIの活用（その7）

座長：植田 知孝（オリエンタルコンサルタンツ）

10:50 ~ 11:00

[CS14-45] 道路施設点検における基盤モデルを用いたZero-shot類似画像検索

*都築 幸乃[1]、全 邦釘[2] (1. 八千代エンジニヤリング株式会社、2. 東京大学)

キーワード：人工知能、深層学習、基盤モデル、CLIP、類似画像検索

国土交通省では，道路構造物の諸元や点検結果に関するデータを全国道路施設点検データベースとして公開している．本研究では，点検データベースを活用した維持管理作業の効率化を目標とし，基礎研究として，類似損傷画像検索における基盤モデルの有用性について調査する．画像認識において代表的な基盤モデルの一つであるCLIPを用いて特徴量を抽出し，事前に作成したデータベースの中から入力した特徴量に対する近傍特徴量を探索し，それに対応する画像を類似画像として出力する．実験の結果，CLIPで抽出した特徴量は同じ部材および損傷で類似度が高い傾向にあり，損傷画像に適用可能な特徴量抽出器として基盤モデルの有用性が示された．

道路施設点検における基盤モデルを用いたZero-shot類似画像検索

八千代エンジニヤリング（株）　正会員　○都築　幸乃
東京大学　正会員　全　邦釘

１．はじめに

国土交通省では，デジタル道路地図等を基盤として，構造物データやリアルタイムデータを紐付けた 3 次元道路データプラットフォーム「xROAD」を構築している．さらに xROAD の一環として，道路構造物の諸元や点検結果に関するデータを「全国道路施設点検データベース」として公開している．

膨大な点検・診断データを活用した維持管理の効率化の取組として，既往研究では AI を用いた類似損傷画像検索システムが提案されている[1, 2]．これらは学習済みの CNN モデルを特徴量抽出器として活用している．

一方，発展の著しい AI 分野では，大規模なデータセットで事前学習された基盤モデルが注目されている．膨大なパラメータを持つ基盤モデルは，新しいラベルに対する再学習をすることなく様々なタスクに適用することができる（Zero-shot transfer）．中でも CLIP（Contrastive Language-Image Pre-training）[3)]は画像認識における代表的な基盤モデルの一つであり，多様なデータ表現を保持していることから，損傷画像にも適用できる特徴量抽出器であることが期待できる．

本研究では，点検データベースを活用した維持管理作業の効率化を目標とし，基礎研究として類似損傷画像検索における基盤モデルの有用性について調査する．

２．CLIP と Faiss による類似損傷画像検索

CLIP はテキストと画像のペアからなる大規模データを学習したマルチモーダルな基盤モデルである．テキストと画像それぞれにエンコーダがあり，本研究では画像エンコーダのみを特徴量抽出器として使用する．

まず，CLIP で抽出した特徴量をデータベース（DB）として保存する．次に，DB の中から入力画像の特徴量に対する任意の数の近傍特徴量を抽出し，それらに対応する画像を類似画像として出力する．本研究では，近似最近傍探索ライブラリの Faiss[4)]を使用して L2 距離に基づく近傍探索を行う．

図-1　データ記録様式（その 10）例（一部抜粋）

３．実験・結果

診断する損傷画像に対して，過去の記録から類似度が高い画像を順に複数枚提示することで点検技術者の判断を補助することを想定し，ここでは入力画像と同じ部材または損傷を持つ画像が類似度が高い上位 n 枚（top-n）に含まれているかどうかを評価する．

３．１．データセット

実験で使用する画像と部材・損傷クラス名は全国道路施設点検データベースのデータ記録様式（その 10）から取得した（図-1）．データ数の都合により，2014～2018 年度の点検画像 2,034 枚を推論，2019～2023 年度の点検画像 10,694 枚を DB に使用する．

３．２．実験１：部材検索

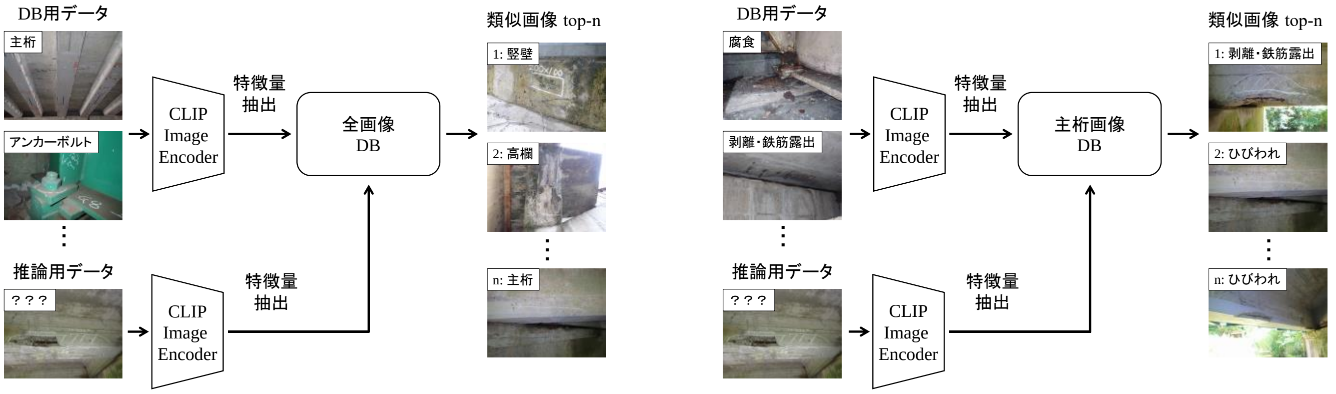
実験 1 では，入力した損傷画像に対して同じ部材クラスの画像が類似画像として出力されるかを検証する．図-2 に示すように，すべての部材の画像特徴量を保存した全画像 DB から top-n を出力する．

top-n の中に入力画像と同じ部材クラスの画像が含まれていた割合を表-1 に示す（$n = 1, 10$）．これより，top-10 では 90%以上の画像で正しいクラスの画像が出力されていることが確認できる．

３．３．実験２：損傷検索

実験 2 では部材を 1 クラスに限定し，入力した損傷画像に対して同じ損傷クラスの画像が類似画像として出力されるかを検証する．図-3 に示すように，すべての損傷の画像特徴量を保存した主桁画像 DB から top-n

キーワード　人工知能，深層学習，基盤モデル，CLIP，類似画像検索
連絡先　〒111-8648 東京都台東区浅草橋 5-20-8 CS タワー
八千代エンジニヤリング（株）　TEL03-5822-7032



(a)　実験 1: 部材検索　　(b)　実験 2: 損傷検索

図-2　類似損傷画像検索概要

を出力する．

top-n の中に入力画像と同じ損傷クラスの画像が含まれていた割合を**表-2** に示す（n = 1, 10）．これより，損傷の種類によって類似画像の抽出しやすさが異なることが確認できる．特に，腐食や亀裂は精度が高く，漏水・滞水は精度が低い．

３．４．実験３：同一箇所の類似度

実験 3 では，二時期に撮影された同一箇所の損傷が類似度が高い画像として出力されるかを検証する．

実験 2 と同様に主桁 DB で類似画像検索を行った結果，ペアが存在する損傷画像 287 枚のうち 95 枚に対して同一箇所の画像が top-1 類似画像として出力された．目視で確認した結果，**図-3** のように，画角等の撮影条件が似ている場合に類似度が高くなる傾向があった．

４．おわりに

本研究では，CLIP と Faiss を用いて道路施設点検データに対する Zero-shot 類似画像検索を行った．実験により，CLIP で抽出した特徴量は同じ部材および損傷で類似度が高い傾向にあり，類似損傷画像検索における特徴量抽出器として基盤モデルが有用であることが示唆された．

表-1　全データ内訳と部材検索結果

クラス	DB 用データ	推論用データ	top 1	top 10
主桁	1,881	301	55.8	93.7
床版	1,210	300	53.7	95.7
支承本体	1,051	211	66.8	95.3
竪壁	766	334	48.8	93.4
その他 (51 クラス)	5,786	888	-	-
合計	10,694	2,034	-	-

表-2　主桁データ内訳と損傷検索結果

クラス	DB 用データ	推論用データ	top 1	top 10
腐食	450	91	79.1	100
亀裂	327	44	88.6	95.5
剥離・鉄筋露出	49	17	47.1	88.2
漏水・滞水	29	11	27.3	54.6
その他 (12 クラス)	1,026	138	-	-
	1,881	301	-	-

(a)　入力画像　　(b)　top-1 類似画像

図-3　同一箇所を撮影した損傷画像が top-1 類似画像として出力された例

参考文献

1)　熊谷宗一郎，片山直道，全邦釘：橋梁維持管理のための意味的特徴の抽出による類似損傷画像検索，AI・データサイエンス論文集, Vol. 4, Issue 3, pp.142-148, 2023.

2)　柏貴裕，宮﨑文平，竹末直樹：橋梁類似画像検索 AI の活用に関する研究, JAAM, pp. 91-95, 2021.

3)　A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever: Learning Transgerable Visual Models From Natural Language Supervision, ICML, Vol. 139, pp. 8748-8763, 2021.

4)　M. Douze, A. Guzhva, C. Deng, J. Johnson, G. Szilvasy, P. E. Mazaré, M. Lomeli, L. Hosseini, and H. Jégou: The Faiss library, arXiv preprint arXiv: 2401.08281, 2024.