

単眼カメラと Encoder-Decoder モデルを用いた損傷サイズ同定法に関する基礎研究

JIP テクノサイエンス(株)

正会員○林 煒松 非会員 星野隼人 正会員 戸田圭彦

正会員 巻幡憲俊 正会員 今野将顕 正会員 田中克弘 非会員 阿野 翼

1. はじめに

維持管理分野において、画像あるいは映像から構造物の損傷を検知する試みは多々存在する¹⁾。しかしながら損傷の大きさを自動判定することは難しいといった課題がある。損傷の大きさを写真から求める方法として、(1)長さが既知のスケールを貼り付け、損傷と一緒に撮影する方法、(2)写真測量技術²⁾を利用する方法などが考えられる。しかしながら、(1)は橋梁下面などの手が届かない損傷には使えない、(2)はカメラあるいはレンズが2台以上必要で単眼カメラ間の距離が既知である必要があるなど、現場での運用は簡易とは言えない。写真1枚の情報から簡易に損傷サイズを同定する方法の確立が課題である。

深層学習の発展により、深層学習を単一写真から被写体までの距離を推定する深度推定技術へ応用した事例が現れており、様々なモデルが提案されている³⁾。写真の各ピクセルの深度値をRGB値から推定できれば、損傷サイズを求めることができる。そこで本研究では図-1に示すような損傷検知システムにおいて損傷サイズを同定することを目的とした深度推定モデルについて基礎的な検討を行った。

2. 深度推定モデルの概要

既往の研究において、深度推定のための深層学習モデルとして画像セグメンテーションに基づくニューラルネットワークが有効であることが示されている³⁾。本論文ではこのような背景を踏まえ、U-Net⁴⁾を採用し、深度推定モデルを構築した。図-2に示すように、深度カメラによりRGB-D画像を取得し、そのRGB画像と深度の関係をモデルに学習させることで、RGB画像から深度を推定できるAIモデルを開発した。以下にモデルの詳細を示す。

2.1 ネットワーク構造

図-3に示されるように、EncoderとDecoderにはそれぞれ4つのブロックを使用し、各ブロックには2つの畳み込み層が含まれる。また、各ブロック間ではMax-poolingとUp-samplingを使用して特徴量を処理する。

2.2 損失関数

損失関数を次式に示す。

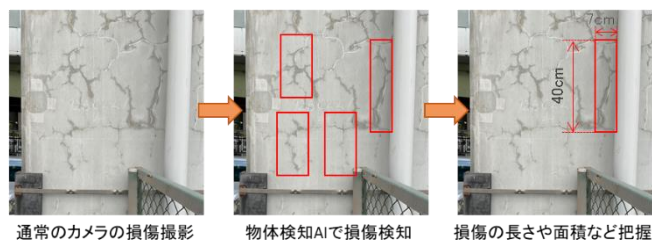


図-1 システム概念図

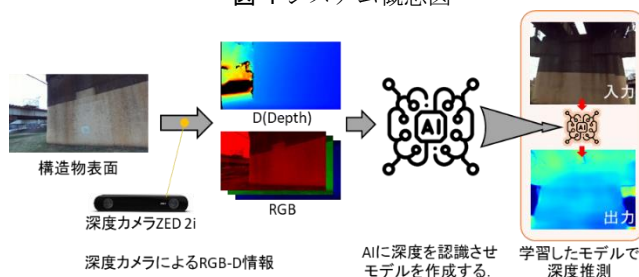


図-2 深度推定 AI モデルの作成および運用

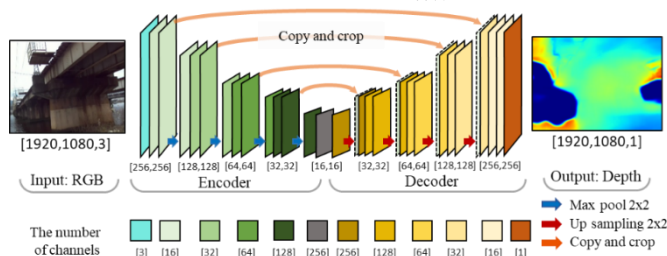


図-3 ネットワーク構造図

$$Loss = W_{SSIM}L_{SSIM} + W_{L2}L_{L2} + W_{Smth}L_{Smth} \quad (1)$$

ここに、 L_{SSIM} , L_{L2} , L_{Smth} はそれぞれ構造類似性指標(SSIM), L2 損失, 深度の平滑性損失であり、次式で表される。

$$L_{SSIM} = \frac{1}{n} \sum_i^n (1 - SSIM) \quad (2)$$

$$L_{L2} = \frac{1}{n} \sum_i^n (d_i - d_i^{gt})^2 \quad (3)$$

$$L_{Smth} = \frac{1}{n} \sum_i^n S_x(d_i, d_i^{gt}) + \frac{1}{n} \sum_i^n S_y(d_i, d_i^{gt}) \quad (4)$$

$$SSIM = \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)} \quad (5)$$

$$S_x(d_i, d_i^{gt}) = \left| \partial_x d_i \cdot \exp \left(\frac{1}{n} \sum_i^n |\partial_x d_i^{gt}| \right) \right| \quad (6)$$

$$S_y(d_i, d_i^{gt}) = \left| \partial_y d_i \cdot \exp \left(\frac{1}{n} \sum_i^n |\partial_y d_i^{gt}| \right) \right| \quad (7)$$

また、 W_{SSIM} , W_{L2} , W_{Smth} は各項の重みであり本研究では0.1, 0.5, 0とした。 d_i は推定深度、 d_i^{gt} は正解深度、 μ_a, μ_b は画

キーワード 深層学習, 深度推定, U-Net, 単眼カメラ

連絡先 〒532-0011 大阪市淀川区西中島 2-12-11 JIP テクノサイエンス(株) TEL : 06-6307-5404

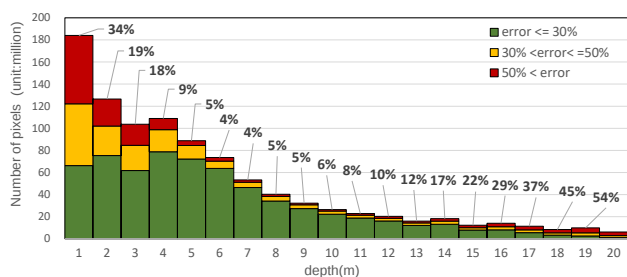


図-4 深度ごとの推測誤差率分布
(図中の数値は誤差率 50%以上が占める割合)

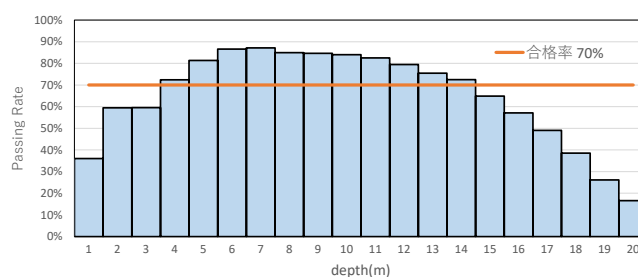


図-5 誤差率 30%以下の割合分布

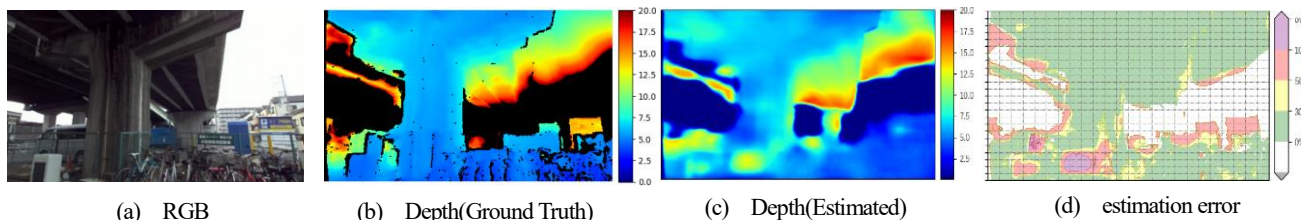


図-6 深度推定 AI で推測した結果の一例 (b), (c)の単位 : m)

像 a, b の小領域内の深度の平均, σ_a, σ_b は画像 a, b の小領域内の深度の標準偏差, σ_{ab} は画像 a, b の小領域内の共分散, そして C_1, C_2 は定数である. なお, SSIM は画像の深度推定値につき, 推定値と教師データの間の明るさ, コントラスト, および構造の類似度を比較して定量化したものである. 一方, 深度の平滑性損失は, 隣接したピクセルの深度の勾配にペナルティを課すことにより, 推定深度が滑らかになるようにするものである. ただし本研究では深度変化の大きい写真を対象としたため, より構造物の境界部が明瞭となるように, 平滑性損失の重みは 0 とした.

3. 教師データおよび学習

ZED 2i 深度カメラ⁵⁾ (35mm 換算焦点距離 : 15mm) を使用して, コンクリート構造物の表面を録画した映像から, 合計 2,886 枚の撮影された写真(RGB)と深度ファイル(D)を学習データ(教師ありデータ)として取得した. モデルに与えた教師データのうち, 学習用のデータと検証用のデータ (以下, テストデータ) 数は 8:2 の割合であり, テストデータを使用して, その学習に基づく推定値に対する損失を計算した. なお, 本研究では 90 エポック程度で学習データとテストデータに対する損失曲線に変化が見られなくなったとして学習を打ち切った.

4. 検証

テストデータを用いてモデルが推測する距離の誤差を観察した. 図-4 に各写真の正解深度ごとのピクセル数の分布, および正解値毎の誤差率の割合を示す. 誤差率は正解深度と推定深度から求め, さらに誤差率が 30%未満, 30%~50%, 50%以上の 3 つの範囲に分類した. 図より, 距離 4m 以上の遠距離において推定精度が高く, 距離 4m~11m において誤

差が 50%以上の比率が 10%未満に抑えられていた. しかし, 距離 4m 未満の近距離での推定精度は低く, 距離 1m 以内での推測において誤差が 50%以上の比率が 34%を占めていた. 次に, 実用上の観点から損傷サイズの許容誤差を 30%と仮定し, その許容誤差を下回った割合を合格率として算出した. 距離と合格率の分布を図-5 に示した. この図より, 距離 4m~14m では合格率が 70%を上回っており, 開発した深度推定モデルを用いた損傷サイズ同定がある程度の性能で可能であることが示された. また, 合格判定となった画像の 1 例を図-6 に示す. 深度推定の結果(c)は正解深度 (b)の傾向を概ね再現することができている.

5. まとめ

本研究では, U-Net を用いた深度推定モデルの実用性を検証し, 本研究で用いた条件であれば, 十分な精度で絶対深度を推定することができることを確認した. 今後の研究では, 異なる焦点距離, 異なる撮影デバイスへ適用範囲の拡大について検討予定である.

参考文献

- 1) Zhang, G., Wang, B., Li, J. et al.: The application of deep learning in bridge health monitoring: a literature review, *Advances in Bridge Engineering*, Vol.3, Article No.22, 2022.
- 2) 織田 和夫 : 解説 : Structure from Motion (SfM) 第一回 SfM の概要とバンドル調整, 写真測量とリモートセンシング, Vol.55, No.3, pp.206-209, 2016.
- 3) Masoumian, A., Rashwan, H. A., Cristiano, J, Asif, M. S., Puig, D.: Monocular Depth Estimation Using Deep Learning: A Review, *Sensors* Vol.22, Article No.5353, 2022.
- 4) Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol.9351, pp.234-241, 2015.
- 5) ZED 2i | Stereolabs: <https://www.stereolabs.com/zed-2i/> (参照 2023/03/27)