構造物のひび割れ検出における物体検出技術に関する研究

中央大学大学院　学生員　○ 八木　笙太
中央大学　正会員　樫山　和男

1. はじめに

近年，国土交通省の推進する BIM/CIM，i-Construction といった取り組みにより，建設業界における ICT の活用は一層活発化している．これらの取り組みの一つに，IoT 機器などによって収集されるビッグデータを有効的に活用する AI(Artificial Intelligence) 技術がある．近年の AI 技術は深層学習 (Deep Learning) の登場により，入力されたデータからその特徴を自動で学習することが可能となった．中でも画像認識技術は様々な分野において研究・活用がされており，土木・建設分野でも維持管理等に用いられている．維持管理において画像認識技術は高所や他の危険な箇所など立ち入り困難な場所において安全確保・業務効率化等を目的に発展が進んでいる．先行研究では実時間での構造表面のひび割れ検出に関して，物体検出手法 YOLOv2(You Only Look Once v2) に加えてひび割れ検出を二段階認識で行う手法[1)]，また生成モデルを用いて行う手法[2)] がそれぞれ提案されているが，どちらも既存の物体検出手法 YOLOv2 に加えて別の提案を行っているものである．

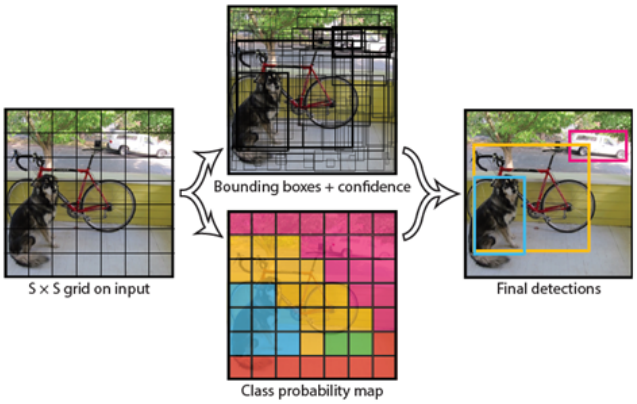
本論文では物体検出手法 YOLO を用いるひび割れ検出において，YOLO 内部のデータ拡張やネットワーク構造を変化させることによる検出精度への影響について確認を行う．

2. 物体検出手法 YOLO

YOLO は領域検出とクラス認識を単一のネットワークで同時に行うことにより，非常に高速な物体検出を可能としている．また，YOLO は同作者からその後継となる YOLOv2[4)] や YOLOv3[5)] といった手法が提案されているが，以下では YOLOv3 における推論の流れと誤差関数，および各モデル構造について述べる．

(1) 推論手順

図-1 に YOLO の推論手順を示す．手順として，まず入力画像を $S \times S$ のグリッドセルに分割する．領域検出においては Anchor Box と呼ばれる一定のアスペクト比を持つ矩形を各グリッドにおいて B 個オブジェクトを予測するように配置し，その中心座標 $(\hat{t_x}, \hat{t_y})$ と矩形の大きさを示す幅と高さ $(\hat{t_w}, \hat{t_h})$ を取得する．同時に各矩形においてその領域内に物体が存在する可能性を示す信頼度 t_o が与えられる．物体が存在する信頼度が高い場合には，CNN によるクラス予測が行われ，クラス予測値 $\hat{z_k}$ が取得される．以上の手順から得られる各予測値とそれに対応する教師データとの誤差を計算し，その最適化を図ることによって画像内の対象クラスの位置と大きさおよびクラス名を決定する学習が行われる．



図－1　YOLO 推論図[3)]

(2) 誤差関数

学習にあたって，教師データと予測データの誤差を取得する誤差関数 L を以下に示す．

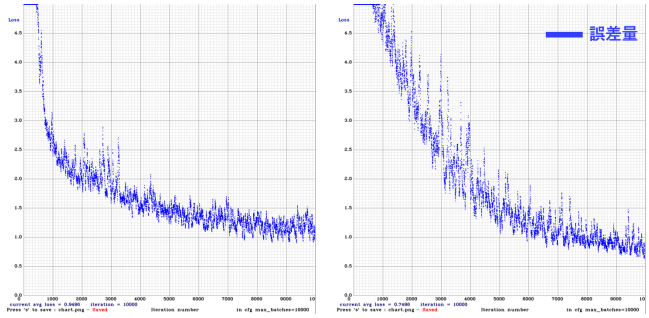
$$
\begin{aligned}
L = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{i,j}^{obj} [(t_x - \hat{t_x})^2 + (t_y - \hat{t_y})^2 \\
+ (t_w - \hat{t_w})^2 + (t_h - \hat{t_h})^2] \\
+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{i,j}^{obj} [-log(\sigma(t_o)) + \sum_{k=1}^{C} BCE(\hat{z_k}, (\sigma(s_k)))] \\
+ \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{i,j}^{noobj} [-log(1 - \sigma(t_o))] \qquad (1)
\end{aligned}
$$

式 (1) 中の BCE は BinaryCrossEntropy の略であり，交差エントロピーによるクラス分類誤差を示す．ハットがつく変数は，YOLOv3 モデルによる推測値であり，ハットがつかない変数は教師データから得られる正解値のデータである．1_{ij}^{obj} は，j 番目の Anchor Box の中心座標が i 番目のグリッドに存在するときに 1，それ以外の場合に 0 を示す関数であり，1_{ij}^{noobj} は，その逆を示す関数である．なお，λ_{coord}，λ_{noobj} はプログラム上で任意に設定できる値であり，本論文では初期設定値である λ_{coord}= 5，λ_{noobj}=0.5 を採用している．学習はこれら全ての項がそれぞれ 0 に収束するように行われる．

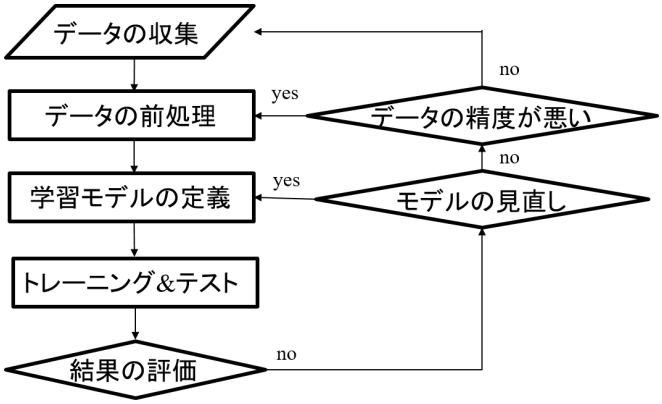
(3) モデル構造

YOLOv2 では特徴抽出機として DarkNet19 という独自の畳み込みニューラルネットワーク (CNN) モデルを採用している．DarkNet19 には，FCN(Fully Convolutional Network) 構造の採用や Batch Normalization の採用，活性化関数を ReLU 関数から LeakyReLU 関数への変更などが行われており，前モデルよりも精度が向上していること

KeyWords：　深層学習，物体検出，ひび割れ検出
連絡先：　〒112-8551 東京都文京区春日 1-13-27　TEL：03-3817-1815　E-mail :a16.365j@g.chuo-u.ac.jp



図－2　仮データにおける YOLOv2(左) と YOLOv3(右) の誤差量推移の比較



図－3　開発のフローチャート

が確認されている．また，YOLOv2 ではデータ拡張 (Data Augmentation) として内部でランダムな教師画像のリサイズ，鏡像反転，回転処理などが施されている．

YOLOv3 では DarkNet53 という CNN モデルを採用している．DarkNet53 は，FPN(Feature Pyramid Networks) 構造の採用によって小さいオブジェクトの検出の精度が向上している．また，SoftMax 関数の廃止による人間と男性・女性といったラベルが重なる多重分類への対応などが行われている．YOLOv3 のデータ拡張としては YOLOv3 の内部処理において彩度，露光，色相，変動などによる処理が施されている．

今回のようなひび割れ検出問題などといった検出対象クラス数が少ない場合，YOLOv3 を採用するメリットは少なく，むしろ学習と検出速度が遅くなるデメリットが存在する．そのため，先行研究[1) 2)] においても物体検出手法には YOLO 系の中でも検出速度が最も早い YOLOv2 が採用されている．しかし，図-2 に示すように，そのモデル構造やデータ拡張方法の差異などによる YOLOv2 と YOLOv3 の学習過程の違いもあり，最適なモデルの構築には両手法の比較も必要であると考える．

3.　適用例

本システムのフローチャートを図-3 に示す．

(1)　データの収集と前処理

今回対象とするひび割れ画像は Python ライブラリの icrawler を採用して収集を行い，収集した画像に対して手作業にて選別を行った上で，200 枚程度を対象とした．また，教師データ作成のためのアノテーションツールとして labelImg を採用し，これを用いて検出領域の中心座標の情報，大きさを示す幅と高さ情報，およびクラス情報を取得している．

(2)　学習モデルの定義

ネットワークの学習にはファインチューニングを採用し，初期の重みには YOLO の開発者があげている学習済みの値を採用している．学習回数は 10000epoch とし，検出には学習時における最も mAP 値が高いものを採用する．

今回精度検証において考慮する要素としては主に 2 つである．1 つ目はデータ拡張に関する検討であり，YOLOv2 や YOLOv3 の内部で行われている既存のデータ拡張ではなく，AutoAugument や RandAugment 手法などの適用によってデータ拡張を行ったものとの精度の検討を行う．2 つ目はモデル構造に関する検討であり，YOLOv2，YOLOv3 それぞれにおいて活性化関数に LeakyReLU ではなく，新たに Mish 関数を採用することや，DarkNet 以外の CNN モデルの採用による差異を検討する．

(3)　結果の評価

結果の評価は混同行列 (Confusion Matrix) に従って行い，モデルの汎化性能の比較には mAP 値による比較を行う．また，物体検出手法として YOLO を用いる上で他の検出手法との違いは主にその検出速度にあるため，各条件下における精度に加えてその検出速度も比較を行う．

(4)　結果

今回考慮する各要素において，その精度および検出速度の差異を示す結果は講演時に示す．

4.　おわりに

本研究では，物体検出手法 YOLOv2，YOLOv3 を用いたひび割れ検出において，その内部構造である活性化関数やネットワーク構造，またデータ拡張方法による精度と検出速度に及ぼす影響の確認を行った．今後の予定としてはさらなる精度向上を目的としたネットワークの構築を行っていく予定である．

参考文献

1) 野村 泰稔, 重村 知輝, 深層学習に基づく物体検出と生成モデルを用いた構造表面損傷の実時間検出技術の開発, 「材料」(Journal of the Society of Materials Science, Japan), Vol.68, No.3, pp.250-257, Mar.2019
2) 重村 知輝, 野村 泰稔, 深層学習に基づく物体検出･認識技術を用いた二段階構造表面ひび割れスクリーニング,「材料」(Journal of the Society of Materials Science, Japan), Vol.69, No.3, pp.218-225, Mar.2020
3) J.Redmon, S.Divvala, R.Girshick, A.Farhadi, You Only Look Once: Unified, Real-Time Object Detection, arXiv:1506.02640, (2020 年 3 月 30 日閲覧)
4) J.Redmon, A.Farhadi, YOLO9000: Better, Faster, Stronger, arXiv:1612.08242, (2020 年 3 月 30 日閲覧)
5) J.Redmon, A.Farhadi, YOLOv3: An Incremental Improvement, arXiv:1804.02767, (2020 年 3 月 30 日閲覧)