

# 論文 AI を活用した表層品質目視評価システムの構築と現場運用に関する一考察

水野浩平\*1・水野健\*2・渡邊賢三\*3・巖侑真\*4

**要旨：**目視調査に基づくコンクリート構造物の表層品質評価手法について、コンクリート表層の画像から評価点を出力する、畳み込みニューラルネットワークを活用した AI 評価ツールを開発し、現場で活用・運用を行った。その結果、技術者による評価は、中程度のコンクリートの変状が認められる場合に偏差が大きくなることが確認された。また、AI による評価は、現場で取得したデータに対しても精度良く評価できることが確認された。ただし、評価点が低い範囲においては精度が低下することが課題として見出された。

**キーワード：**表層品質、目視評価、AI、機械学習、畳み込みニューラルネットワーク

## 1. はじめに

トンネルや橋梁をはじめ、社会インフラとして重要な役割を果たすコンクリート構造物には、高い耐久性が要求される。コンクリート構造物の耐久性を確保するためには、コンクリートの表層品質が重要であることが明らかになっている<sup>1)・2)</sup>。表層品質を確保・向上させる手法は多岐に亘り、近年、コンクリートの打込みの際に、施工方法の勘所を確認することで、コンクリート構造物の品質を高める手法（品質確保システム）<sup>3)</sup>や目視調査結果に基づいてコンクリート施工のPDCAサイクルを回して品質を確保する手法（目視評価法）<sup>4)</sup>などが運用されている。これらは、特別な材料や工法などを用いないため、多大なコストをかけることなく不具合の発生を防止し、構造物の耐久性などの品質を確実に確保することで社会資本整備費の抑制に貢献できるものと考えられる。

筆者らは、目視評価法について、評価者の所属、経験や技量などの属性が評価結果に及ぼす影響について検討し、評価者の経験が乏しい場合に、評価のばらつきが大きくなることを明らかにしている<sup>4)</sup>。この課題に対して、AIにより評価を補助することで、経験の乏しい評価者であっても熟練技術者と同等の評価ができる目視評価システムの構築を検討している<sup>5)・6)</sup>。

本稿では、コンクリートの画像と技術者による評価からデータセットを作成し、畳み込みニューラルネットワーク（Convolutional Neural Network：以下、CNN表記）を用いたモデルの検討を行った。さらに、技術者とAIモデルによる評価が容易となるよう、WEBアプリ化した目視評価システムを構築し、広く現場に展開することで集積されたデータに基づき、技術者による評価のばらつきやAIによる評価精度の分析を行った。

## 2. 目視評価法による表層品質評価

目視評価法<sup>2)・4)</sup>は、同じコンクリートを用い施工の良否によって外観に相違が生じた場合に、「美しいコンクリートは品質と耐久性の高いコンクリートである」<sup>2)</sup>という概念のもと、全6つの評価項目に対して技術者が目視によりコンクリートの表層品質を評価する手法である。本手法では、コンクリートの耐久性および見栄えに影響を及ぼすと考えられる、打重ね線、表面の色つや（以下、色つやと表記）、表面気泡、沈みひび割れ、型枠継ぎ目ののろ漏れ（以下、ノロ漏れと表記）および砂すじの6項目に対し点数化して評価する。評価の点数は、表-1に示すように、実際の構造物の表層品質をサンプル写真と照らし合わせて、一般的に「良」とされる範囲を評価点4,3,2および1に分類し、さらに補修を必要とする「否」の「0」を加えて評価を行う。また、2以上3未満と判断された場合は2.5と付けるなど、0.5刻みで評価点を付ける。本稿においては、AIモデルの構築時には特定の技術者3名により評価を行い、目視評価システムの運用時には、現場で担当となった不特定の技術者75名により評価を実施した。

## 3. AIモデルの構築

### 3.1 データセットの構築

本検討で使用した画像は、脱型後の実際のコンクリート構造物より、打重ね線、色つやおよび砂すじの3項目は5m程度、表面気泡、沈みひび割れおよびノロ漏れの3項目については2m程度離れた位置より撮影したものを640×640ピクセルに圧縮し、6項目毎にAIモデル用の学習データを準備した。学習データ数を表-2に示す。学習データは、コンクリート主任技士1名、同技士2名



\*1 鹿島建設(株)技術研究所 土木材料グループ 副主任研究員 修士(工学)

\*2 鹿島建設(株)土木管理本部 再生エネルギー部 学士(工学)

\*3 鹿島建設(株)技術研究所 土木材料グループ グループ長 博士(工学)

\*4 (株)UNAIIT 開発部 学士(工学)

表-1 目視評価におけるグレードの一例（打重ね線，型枠継ぎ目ののろ漏れ，砂すじを抜粋）

基準 項目	一般的に「良」とされる範囲				不適合 —
	4点	3点	2点	1点	
打重ね線	 ・近接では打重ね線が認められるものの、約 10m 離れた遠方からは認められない	 ・約 10m 離れた遠方から、打重ね線が認められる	 ・約 10m 離れた遠方から、打重ね線がはっきりと認められる	 ・2 点の状態よりも劣る	構造物のオーナーから不具合と判定される状況で補修を要するもの
型枠継ぎ目ののろ漏れ	 ・調査対象範囲にのろ漏れがほとんど認められない	 ・調査対象範囲の概ね 1/10 以上にのろ漏れが認められる	 ・調査対象範囲の概ね 1/3 以上にのろ漏れが認められる	 ・2 点の状態よりも劣る	
砂すじ	 ・調査対象範囲に砂すじがほとんど認められない	 ・調査対象範囲の概ね 1/10 以上に砂すじが認められる	 ・調査対象範囲の概ね 1/3 以上に砂すじが認められる	 ・2 点の状態よりも劣る	

豆板については、「なし」を 4 点、「ある」を不適合と評価する。

が全ての画像に対して評価点を付け、3 名の平均値を評価点とした。なお、学習データにおける 3 名の評価点の偏差は最大でも 0.47 点であり、学習データは誤差が極めて小さいものであることを確認した。

### 3.2 モデルの構築

色つや、表面気泡および沈みひび割れの 3 項目については、既報<sup>5), 6)</sup>にて AI モデルおよび学習方法を報告している。そこで本稿においては、打重ね線、ノロ漏れおよび砂すじの 3 項目に対する AI モデルおよび学習方法について検討した結果を示す。

打重ね線の判定モデル（以下、打重ねモデルと表記）の構造を図-1 に示す。打重ね線モデルは、元画像から打重ね線部を検出し、検出した打重ね線を線の濃淡や連続性を判定基準として 4 段階に分類した後に、写真から背景などを除いたコンクリート部分の検出ピクセル数および分類した打重ね線のピクセル数を入力データとして式(1)により評価点を算出する構造である。

$$E = \frac{A - X_1L_1 - X_2L_2 - X_3L_3 - X_4L_4}{A} \times 4 \quad (1)$$

ここで、E：評価点、A：コンクリート部分の検出ピクセル数、Ln：分類 n の検出ピクセル数（n=1, 2, 3, 4）、Xn：分類 n における係数（n=1, 2, 3, 4）とした。なお、4 段階への分類および係数を後述の学習時に最適化した。

表-2 学習データ数

項目	打重ね線	色つや	表面気泡	沈みひび割れ	ノロ漏れ	砂すじ
学習枚数	1500	891	891	1097	900	580

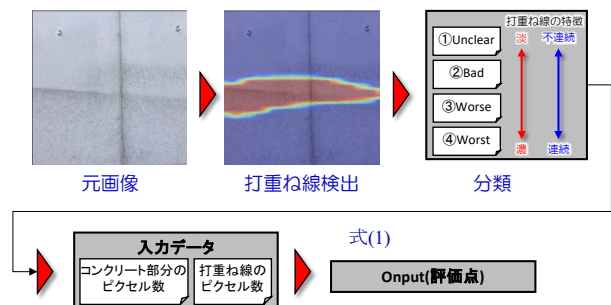


図-1 打重ね線の判定モデルの構造

ノロ漏れの判定モデル（以下、ノロ漏れモデルと表記）の構造を図-2 に示す。ノロ漏れモデルは、元画像から型枠の継ぎ目の範囲を検出し、ノロ漏れの範囲を検出した後に、型枠継ぎ目の本数と長さおよびノロ漏れ検出のピクセル数を入力データとして、Microsoft 社公開の決定木に基づく分析アルゴリズムである Light Gradient Boosting Machine（以下、Light GBM と表記）により評価点を算出する構造である。

砂すじの判定モデル（以下、砂すじモデルと表記）の

構造を図-3 に示す。砂すじモデルは、元画像から砂すじの範囲を検出し、砂すじ部の隣り合うピクセルの輝度差を示すエッジを検出した後に、砂すじ範囲の検出ピクセル数および砂すじ部の検出エッジのピクセル数を入力データとして、ノロ漏れと同様に Light GBM により評価点を算出する構造である。

### 3.3 学習方法および学習結果

打重ねモデル、ノロ漏れモデルおよび砂すじモデルにおいて、画像認識の機械学習として CNN によるセマンティック・セグメンテーションを用いた。CNN のネットワーク構造を示すアーキテクチャには U-Net を、CNN の特徴抽出構造を示すエンコーダには EfficientNet-b3 をそれぞれ用いた。なお、使用する画像は 512×512 ピクセルに圧縮してサイズを統一し、学習回数は 10~50 回程度とした。また、いずれのモデルについても、学習による評価結果の精度判定には、正解領域と予測領域の一致率を表す IoU を用いて評価した。なお、IoU は数値が高いほど一致率も高いことを示している。

打重ねモデル、ノロ漏れモデルおよび砂すじモデルにおける学習結果を図-4 に示す。なお、図には、既往の研究<sup>6)</sup>で報告されている、沈みひび割れのモデルにおける学習結果もあわせて示す。3 種類のモデルに共通して、学習回数の増加とともに、IoU が大きくなる傾向が認められた。既往の検討で実施した色つやおよび表面気泡の AI モデル学習において、100 回を超える学習回数ときに特定のデータに対してのみ推定精度が高くなる過学習の状態となったことから<sup>5)</sup>、本稿の 3 種類のモデルにおいては、IoU が概ね収束した時点で学習を終了した。最終学習回数における IoU は打重ねモデル、ノロ漏れモデルおよび砂すじモデルにおいてそれぞれ 0.72、0.90 および 0.85 ととなり、比較的高い精度で評価点を予測できることが示された。なお、色つやおよび表面気泡の判定モデルの学習については、判定の正解値と予測値との差の大きさを示す Loss を用いて精度を評価しており、2 つのモデルともに学習回数 87 回で Loss=0.01 と、差が少なく高い精度で予測できることを確認している。

## 4. AI モデルを搭載した目視評価システムの運用

### 4.1 目視評価システムと現場運用の概要

目視評価法について、3. に示した AI モデルを WEB アプリとして構築した。本システムは、技術者による判定入力画面に、表-1 に示したサンプル画像が表示されるため、実際のコンクリート表面の状態と見比べながら評価を実施することができる。また、技術者の評価後に、AI による評価結果が表示され、技術者自身の評価の振返りが即座に実施できる。

本システムを、橋梁、臨海、都市、トンネルなど様々な工種の 49 現場におけるコンクリート躯体構築のうち、

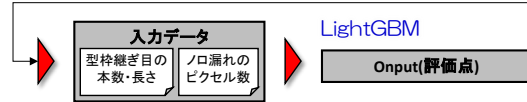
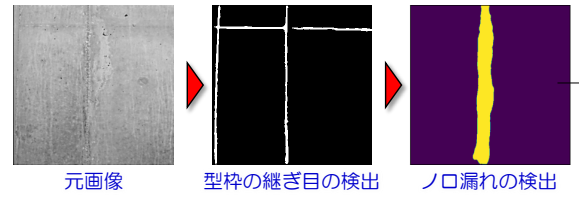


図-2 ノロ漏れの判定モデルの構造

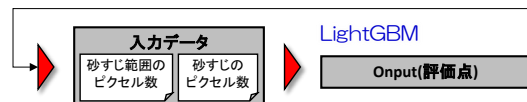
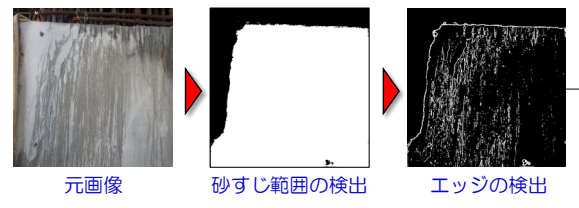


図-3 砂すじの判定モデルの構造

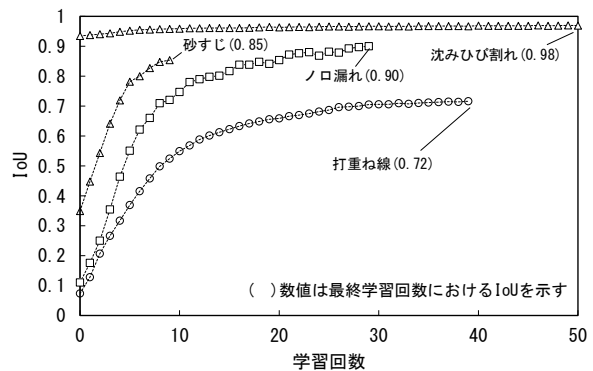


図-4 学習回数と IoU の関係

壁や柱など 849 ブロックを対象に適用した。目視評価は、同一ブロック毎に技術者 1~3 名の評価に加え、前述の AI モデルによる評価もあわせて実施した。なお、技術者による評価は現場や対象ブロックに応じて、376 ブロックを 1 名、406 ブロックを 2 名、67 ブロックを 3 名により評価を行い、AI モデルによる評価は、全 849 ブロックにて実施した。

### 4.2 現場運用の結果と考察

#### (1) 技術者による評価結果

技術者による目視評価の点数の度数分布を図-5 に示す。なお、目視評価の点数は、技術者 1~3 名による評価結果の平均値を示し、評価点の範囲は、0 以上 1 以下、1 大なり 1.5 以下、1.5 大なり 2.0 以下、2.0 大なり 2.5 以下、2.5 大なり 3.0 以下、3.0 大なり 3.5 以下および 3.5 大なり 4.0 以下の 7 種に区分した。評価点の度数分布は、6 項目いずれにおいても、評価点の高い 3.0~4.0 を中心に

分布し、美観に優れたコンクリート構造物が大半を占めることが確認された。打重ね線、表面気泡、沈みひび割れ、ノロ漏れおよび砂すじの5項目については、点数が高いほど頻度も高い傾向が認められる。また、沈みひび割れにおいては他の項目と比較して、3.5~4の範囲の頻度が顕著に高い結果であった。一方、色つやに着目すると、3.0~3.5に頻度のピークが認められた。色つやについては、実際のコンクリート構造物における発生のメカニズムが解明されておらず<sup>7)</sup>、評価点を改善する具体的な対策が確立されていないことが影響したものと推察される。

技術者による目視評価の傾向を分析するため、技術者2名および3名により目視評価を実施した473ブロックを対象に、技術者による評価点の偏差を算出した。偏差の範囲を0以上0.3以下、0.3大なり0.6以下、0.6大なり0.9以下および0.9以上の4種に区分した技術者の評価点の偏差の度数分布を図-6に示す。技術者の評価点の偏差は、6項目いずれにおいても、偏差の小さいものほど多く、技術者によるばらつきが小さく評価できているものと考えられた。沈みひび割れについては、他の5項目と比較して特に0~0.3の範囲の頻度が高い結果となった。

技術者による目視評価のばらつきが生じやすい条件が特定できれば、特定範囲のデータに対し重点的に学習するなど、今後のAIモデルの改良などに活用できる。そこで、図-5に示した全データのうち複数の評価者による473データを抽出し、このデータを得た際のばらつきに着目した。評価点の範囲を図-5と同様の7種に区分し、6項目それぞれに対して算出した技術者による評価点の偏差の平均値を図-7に示す。6項目全体としては、評価点が1.5~2.0および2.0~2.5の範囲において偏差の平均値のピークが認められ、評価点が低く、または高くなるほどに平均値も小さくなる傾向が確認された。コンクリートの変状が大きく明確に認められる場合には、技術者が判断の迷いなく低い点数を付けることが考えられる。同様に、コンクリートの変状が明確に認められず美観に優れている場合においても、技術者が判断の迷いなく高い点数を付けるものと考えられる一方、コンクリートの変状が中程度の場合に、技術者による評価のばらつきが大きくなることが示唆された。

## (2) 技術者とAIによる評価結果

目視評価の項目ごとにAIモデルによる評価精度を検証するため、施工ブロック毎に技術者とAIの評価点の差分(以下、評価点の差分と表記)を算出した。なお、評価点の差分は、-2.0から+2.0まで0.1刻みとした。全849ブロックに対する技術者とAIの評価点の差分の度数分布を図-8に示す。なお、評価点の差分の平均値

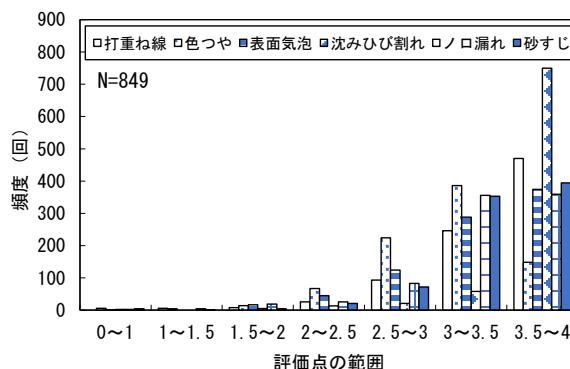


図-5 技術者による目視評価点の度数分布

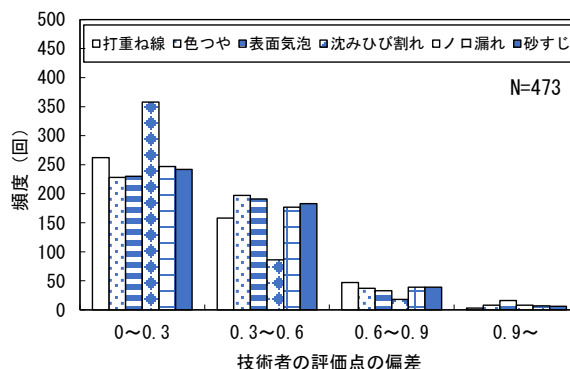


図-6 技術者による目視評価点の偏差の度数分布

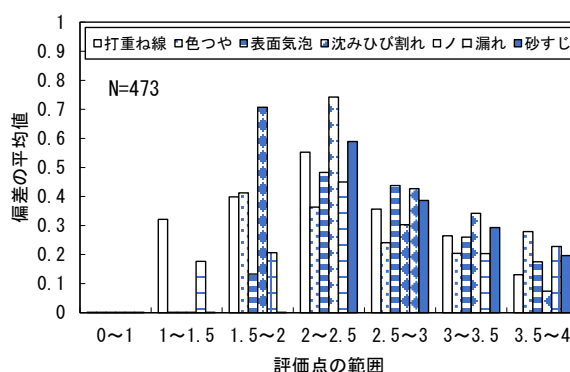


図-7 技術者による目視評価点の偏差

および標準偏差を正規分布曲線とあわせて図中に示す。

打重ね線については、評価点の差分0に頻度のピークが認められるものの、分布としては差分が負となる、すなわち技術者に比べAIの評価点が高くなる傾向となった。また、評価点の差分が-2.0以下となる施工ブロックは11か所認められ、AIが過大評価する場合もあることが確認された。

色つやについては、評価点の差分の頻度のピークが0.2付近に認められ、ピークを軸に正規分布に近い傾向の度数分布が得られた。また、技術者に比べAIの評価点が高くなる傾向が確認された。

表面気泡については、評価点の差分0に頻度のピークが認められ、正規分布に近い傾向の度数分布が得られ

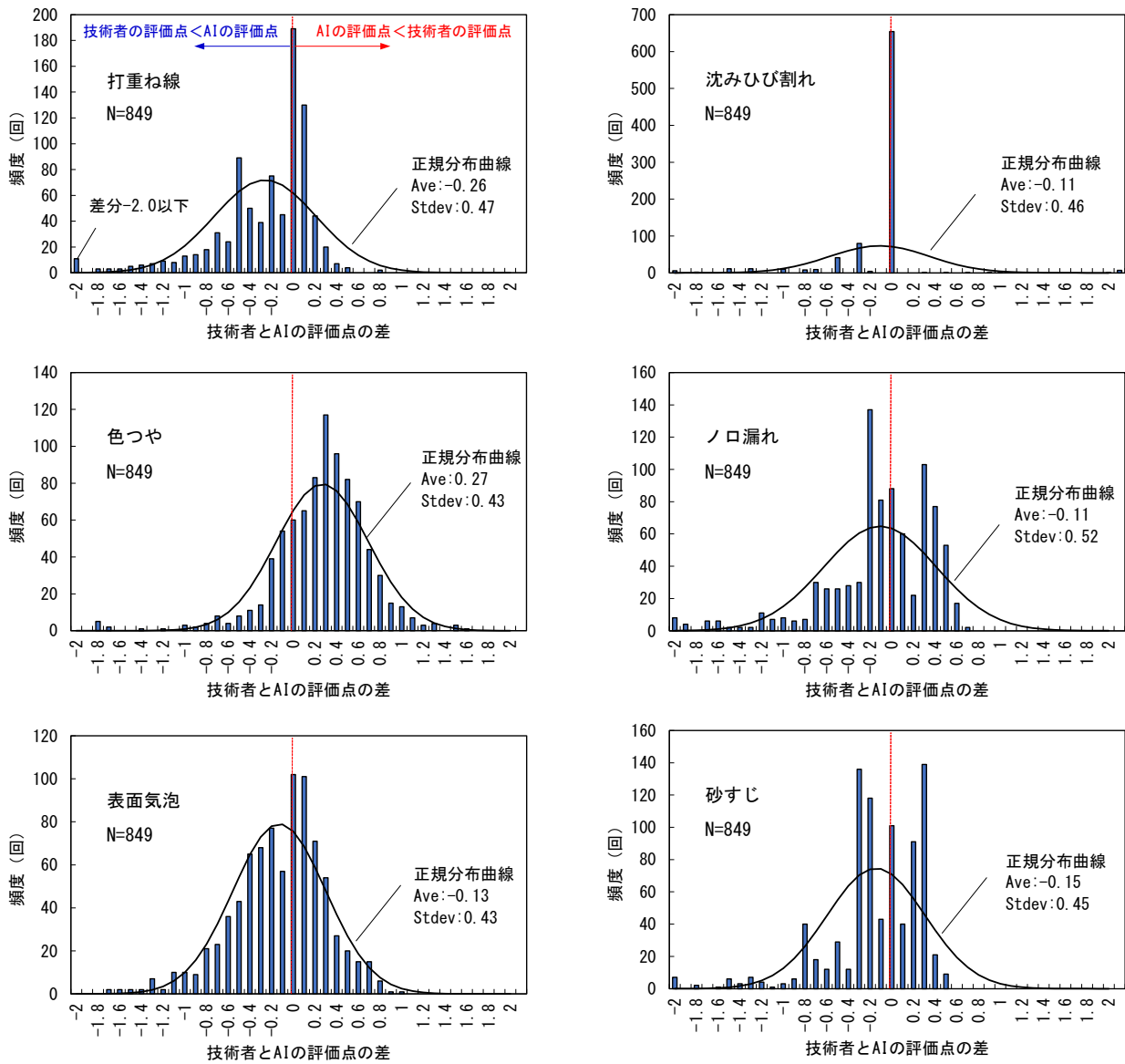


図-8 技術者と AI の評価点の差の度数分布

たものの、分布としては差分が負となる、すなわち技術者に比べ AI の評価点が高くなる傾向となった。

沈みひび割れについては、大半のデータが評価点の差分 0 であり、明瞭なピークが確認された。

ノロ漏れについては、評価点の差分の頻度のピークが -0.2 付近に認められ、差分が負となる、すなわち技術者に比べ AI の評価点が高くなる傾向が確認された。また、評価点の差分が -2.0 以下となる施工ブロックは 8 か所認められ、AI により過大評価する場合もあることが確認された。

砂すじについては、評価点の差分の頻度のピークが -0.3 および +0.3 付近の両者に認められ、差分が負となる、すなわち技術者に比べ AI の評価点が高くなる傾向が確認された。また、評価点の差分が -2.0 以下となる施工ブロックは 7 か所認められ、AI により過大評価する場合もあることが確認された。

以上より、技術者と AI の評価点の差の度数分布は、0 付近にピークが認められ、AI により概ね技術者と近い評価ができるものと考えられた。ただし、AI により過大・過小評価する施工ブロックもあり、AI による評価精度の向上のため、モデルの修正や学習の工夫などが必要となる課題も見出された。

学習データは、前述の表-2 に示すよう 1000 枚程度を用い AI モデルを構築したが、一般に、機械学習としては少ない。経験の乏しい評価者であっても AI の補助により熟練技術者と同等の評価ができることを目的とすると、前述の図-7 に示すように技術者の評価のばらつきが生じやすい中程度の評価点の範囲において、AI による評価精度を優先的に向上させることが、学習データの有効活用の観点からも効率的であるものと考えられる。そこで、技術者による評価点と AI モデルによる予測評価点の絶対誤差の平均値を示す MAE (mean absolute error) に基づ

表-3 学習時と運用時の MAE の比較

分類	打重ね線	色つや	表面気泡	沈みひび割れ	ノロ漏れ	砂すじ
学習時	0.32	0.27	0.23	0.08	0.23	0.24
運用時	0.34	0.40	0.34	0.17	0.37	0.33

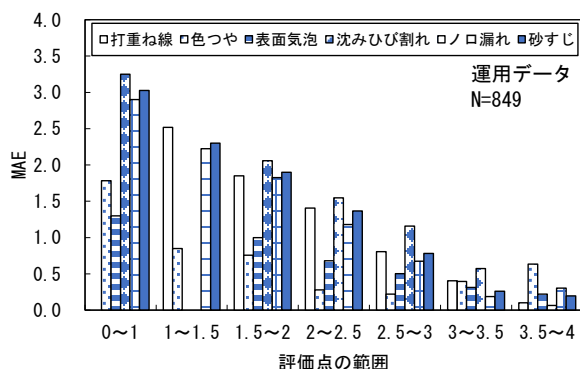


図-9 AI モデルによる MAE

き、AI による精度向上の方向性を考察することとした。例えば、MAE=0.5 のとき、AI による評価が概ね技術者の評価点±0.5 の範囲に収まることを表し、MAE が小さいほど AI による評価精度が高いことを示しており、本検討においては、MAE が 0.5 以内となることを目標に、AI モデルの構築を行った。6 項目に対して算出した、学習時と運用時の MAE の比較を表-3 に示す。6 項目いずれにおいても、学習時と比較して運用時の MAE は大きくなっているものの、学習時および運用時ともに MAE が 0.4 以内に収まり、AI により精度良く評価できているものと考えられた。ただし、前述の図-5 に示すように、3.0~4.0 の評価点の施工ブロックが大半を占めており、運用時の全 849 データを対象として算出した MAE は、評価点が高い施工ブロックの重みが大きい点に留意が必要である。次に、評価点の範囲を図-5 と同様に技術者による評価点に基づき 7 種に区分し、6 項目それぞれに対して AI モデルによる評価の MAE を算出した。運用時における AI モデルによる MAE を図-9 に示す。色つやについて、評価点が 2~2.5 および 2.5~3 の範囲で MAE は小さい結果が得られた。技術者による評価のばらつきの生じやすい中程度の評価点に対して MAE が小さく、色つやの AI モデルは、技術者の評価を補助するのに特に有効であるものと考えられた。打重ね線、表面気泡、沈みひび割れ、ノロ漏れおよび砂すじの 5 項目については、評価点が高いほど、MAE も小さくなる傾向が認められ、評価点が高い範囲においては、AI モデルによる評価精度が高い結果が得られた。また、AI モデルでは、評価点が高い範囲において、精度良く評価できないことが課題として見出された。この課題に対しては、現場運用で

得られたデータを活用し、評価点が低い範囲においても AI による評価精度を向上できるようモデルの修正や学習方法の工夫を検討する。

## 5. まとめ

コンクリート構造物の品質評価手法の一つである「目視調査法」について、撮影された画像から評価点を出力する、CNN を活用した AI モデルを構築し、広く現場運用した。得られた知見を以下に示す。

- (1) 技術者による評価の偏差は、大半の施工ブロックで小さいものの、2~2.5 点の中程度の評価点において偏差が高くなる結果となった。
- (2) 技術者と AI の評価点の差分は、ほとんどの施工ブロックで小さいものであったが、一部のブロックで差分-2.0 以下となり、AI により精度良く評価できない場合もあることが確認された。
- (3) 運用により取得したデータより、AI による評価精度は、AI モデル学習時と同程度となる結果が得られた。
- (4) 評価点が低い範囲においては、AI による評価精度が低下する傾向にあることが課題として見出された。

## 参考文献

- 1) 土木学会：構造物表面のコンクリート品質と耐久性検証システム研究小委員会 (335 委員会) 成果報告書、コンクリート技術シリーズ 97, 2012
- 2) 坂田昇, 渡邊賢三, 細田暁：コンクリート構造物の品質向上と表層品質評価手法、コンクリート工学, Vol.50, No.7, pp.601-606, 2012
- 3) 細田暁, 二宮純, 田村隆弘, 林和彦：ひび割れ抑制システムによるコンクリート構造物のひび割れ低減と表層品質の向上、土木学会論文集 E2, Vol.70, No.4, pp.336-355, 2014
- 4) 渡邊賢三, 小林聖, 坂田昇, 細田暁：目視調査に基づくコンクリートの表層品質評価手法の特徴・傾向に関する分析、コンクリート工学年次論文集, Vol.37, No.1, pp.1813-1818, 2015
- 5) 濱田那津子, 中村真人, 有坂壮平, 渡邊賢三：機械学習を活用した目視評価による表層品質評価システムに関する位置検討、コンクリート工学年次論文集, Vol.42, No.1, pp.1186-1191, 2020
- 6) 水野健ら：AI を用いた沈みひび割れ目視評価点の判定に関する位置検討、土木学会第 77 回年次学術講演会, CS14-24, 2022
- 7) 温品達也, 渡邊賢三, 坂田昇, 柳井修司：コンクリート表面の色むら発生に及ぼす要因に関する実験的検討、コンクリート工学年次論文集, Vol.35, No.1, pp.679-984, 2013