

列車前方画像を用いた木まくらぎ劣化度判定 AI システムの開発

前田 梨帆* 長峯 望 合田 航 坪川 洋友 加藤 爽 (鉄道総合技術研究所)

Development of AI System for Inspection of Wooden Sleepers Deterioration Using Train Forward View Images from Train Cab

Riho Maeda*, Nozomi Nagamine, Wataru Goda, Yosuke Tsubokawa, So Kato (Railway Technical Research Institute)

In order to save labor work in the inspection of the track facility by a low-cost and simple inspection method, we have developed a wooden sleeper inspection method analyzing forward view images from a train cab. This paper presents an overview of the system and the results of comparison and analysis of the system's output results with the results of those made by experts.

キーワード: 木まくらぎ検査, 列車前方画像, 画像処理, Deep Learning

Key Words: Wooden sleeper inspection, Forward view image from train cab, Image processing, Deep learning

1. はじめに

線路に敷設された木まくらぎは、保線技術者が 1 本ごとにその劣化度を検査することによって、管理されている。木まくらぎの不良箇所が連続すると、道床横抵抗力が弱まり、車両走行にかかる荷重による軌間拡大やレール温度上昇による座屈等によって、脱線を引き起こす可能性がある。したがって、木まくらぎの検査は走行の安全を考える上で非常に重要な項目となっている。

しかしながら、木まくらぎの検査は作業員の徒歩での目視検査を基本とし、必要に応じて打音等を使用するものであるため、木まくらぎの膨大な敷設数を考慮すると労力が大きい。加えて、将来的な労働人口の減少に伴う保線技術者不足を踏まえると、木まくらぎの検査を今後も維持していくためには、検査の効率化が喫緊の課題である。

そこで、我々は市販のハンディカメラを用いて撮影した列車前方画像による木まくらぎの検査システムの開発を行っている⁽¹⁾⁽²⁾。市販のハンディカメラを使用するため低コストでシステム構築が可能であり、木まくらぎ区間の多い閑散線区や地方鉄道など、高価なシステムの導入が難しい事業者に対しての適用が可能である。

本論文では、開発したシステムの概要について述べたのち、保線技術者による判定とシステムの判定との比較検証を行った結果について示す。

2. 木まくらぎ検査

〈2・1〉 概要 不良まくらぎの一例を図 1 に示す。このように、腐食等によりレールとまくらぎの締結力が機能し

ない不良まくらぎが連続すると、走行する車両の自重によりレールが外側に押され、軌間が拡大する。最悪の場合、レールから車輪が外れ、脱線事故を引き起こすため、木まくらぎ検査は軌道の重要な検査の 1 つである。

〈2・2〉 現行の検査と課題 木まくらぎ検査は、1 年の周期を超えない範囲でまくらぎ 1 本ごとに割れや折れ、腐食等の状態を目視や打音検査により劣化状態を判定している。また、木まくらぎの敷設数は膨大であることや線路内の点検には安全管理上、1 パーティー4 人程度で行うことから検査に多くの時間的・人的な労力を必要としている。

得られた検査結果は管理台帳に記録し、連続不良箇所の抽出や各線区における不良率を算定している。

したがって、木まくらぎ検査を自動化するためには、まくらぎの劣化状態の判定だけでなく、1 本単位で木まくらぎを判別し、劣化度の判定結果を出力する機能が求められる。

システムによる検査の自動化が可能となれば、従来担当者や線区ごとに異なる等曖昧性を含んでいた判断基準が一律化でき、劣化度を定量的に管理できるようになることが期待される。



図 1 不良まくらぎの例

〈2・3〉 研究の目的 以上前節で述べたことから木まくらぎの検査に対しては大きく次の3つが望まれている。

1. 線路内を歩かずに検査ができること
2. 時間的・人的コストが少なく検査ができること
3. 定量的な検査結果が得られること

そこで、ハンディカメラを用いて列車添乗した前方映像から AI を用いて木まくらぎの劣化度を判定するシステムを開発することとした。

これにより、1回の列車添乗で検査ができるため、線路内に立ち入らずとも、低コストで定量的な検査が実現できると考える。

3. 列車前方画像を用いた木まくらぎ検査システム

〈3・1〉 木まくらぎ検査システムの概要 木まくらぎ検査システムの概要を図2に示す。本システムでは低コストかつ容易な撮影のために、市販のビデオカメラ (4K 解像度) と吸盤式の固定マウントを用いて列車前方の窓 (車内) から軌道を撮影する (以下、列車前方画像という)。

車上で撮影ののち、地上処理では木まくらぎ劣化度判定アルゴリズムを用いてまくらぎの劣化度・直角変位・間隔の診断や測定を行う。得られた結果は、「判定画像」と「台帳」形式で出力される。

開発したビューアや出力される台帳を活用することで、まくらぎの連続不良箇所の抽出や不良箇所の状態確認を効率的に行える。さらに、判定画像については、レール締結装置やバラストの状態確認にも応用が可能である。

〈3・2〉 アルゴリズム概要 本アルゴリズムの主な特徴は、①列車前方画像の床下画像化、②列車の位置推定、③ Deep Learning による木まくらぎ劣化度判定である。

①列車前方画像の床下画像化では、図3に示すように列車前方画像を床下画像の様に俯瞰した画像 (以下、疑似床下画像という。) に変換する。列車前方画像は画像手前から奥にかけて物体が細く映るため、射影変換処理を通じてま

くらぎの形状を統一し、現場の状況に近い状態で診断・把握することを可能とする。

②列車の位置推定では、オプティカルフローを用いて疑似床下画像から走行距離や列車速度を求める。これにより動画の各フレームに対して距離情報を付与することができ、画像中のまくらぎのキロ程としての位置を把握することができる。

③Deep Learning を用いて疑似床下画像から木まくらぎの検出と劣化度判定を行う学習済みモデルを構築する。このモデルを劣化度判定モデルと呼ぶ。この劣化度判定モデルを用いて疑似床下画像から木まくらぎの検出と劣化度判定を行う。なお、まくらぎの表面がバラスト等により鮮明に確認できない場合は、誤判定しないようマスクをするために未判定のカテゴリを設けた。

また、Deep Learning には、ネットワークアーキテクチャとして YOLOX⁽³⁾を用いた。YOLOX は、Deep Learning における物体の検出と識別を同時に実施する手法である。

〈3・3〉 学習データセットおよびデータ拡張 〈3・2〉の③で述べた劣化度判定モデルを構築するためのデータセットには、現状の検査分類と同等にするため図4に示す判定の分類方法を用いた。建築限界支障の恐れのある状態を A1 とし、軌間保持機能低下を A2、まくらぎの機能低下を B、軽微な損傷を C、良好なものを D とした。また、まくらぎの表面に多くのバラストが被っていたり、草等で隠れていたり、埋没して判定が困難なものを「未判定」、PC まくらぎを「PC」とし、合計7つのクラスとした。これらをもとに、まくらぎの外形と劣化度情報を専門の保線技術者により与えたデータ約 10 万枚 (劣化度情報の内訳は A1 : 12 本, A2 : 3,339 本, B : 12,165 本, C : 30,343 本, D : 93,323 本, 未判定 : 2,138 本) を学習データとして用意した。

また、実際の撮影時に起こり得る環境変化として、天候

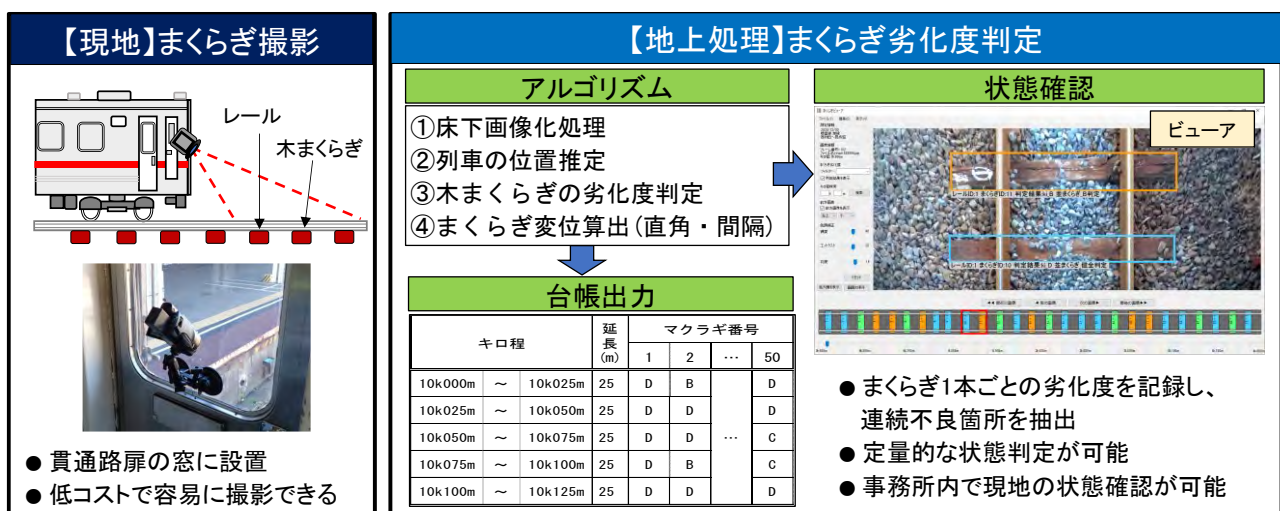
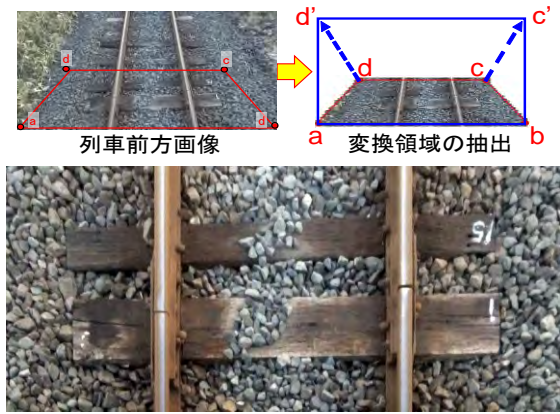


図2 木まくらぎ検査システムの概要



射影変換画像(床下画像化)
図 3 射影変換による床下画像化処理

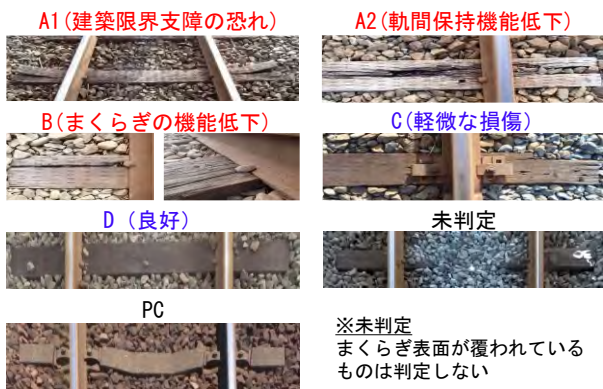


図 4 木まくらぎの劣化度判定標準



図 5 木まくらぎ検出と劣化度判定結果

や太陽の動きによる色調の影響を想定した明度・明暗の変更や、撮影時のホワイトノイズやブレを想定したノイズ付加等の画像処理を実施しデータを拡張した。これにより用意した約 10 万枚のオリジナル画像を 2 倍の約 20 万枚に拡張し、モデルの構築を行った。まくらぎの数としては、約 20 万枚の画像に対し、約 37 万個のまくらぎにタグ付けを行っている。

〈3・4〉モデルによる劣化度推定の例 前節の学習データセットを用いて学習した YOLOX モデルによる木まくらぎ劣化度の判定出力結果の一例を図 5 に示す。図 5 では劣化度が A2 および C である判定されていることを表しており、実際の保線技術者の判定と一致していることを確認した。

4. 精度評価

〈4・1〉 保線技術者による評価用まくらぎの劣化度分類

218 本分のまくらぎ画像を評価用データとして用意し、これらのまくらぎ画像 1 枚 1 枚に対し、まくらぎ検査に従事する保線技術者 7 人による正解付けを行った。保線技術者による判定結果を図 6 の上部に示す。あるまくらぎに対して判定を下した人数が最も多いランクを正解として、正解のランクごとに保線技術者全員による判定結果の割合グラフを表している。すなわち、判定結果の最頻値ごとに図 6 より、およそその傾向として劣化度の判定は一つのランクに偏る一方で、全ての結果が同一のランクに一致するわけではなく、他ランクにも判定がばらついていることが見てとれる。このばらつきについて 4 章 3 節で評価を行う。

〈4・2〉 Deep Learning による劣化度判定精度の検証 同様に評価用のまくらぎ画像 218 枚に対して劣化度判定システムで判定した結果を図 6 下部に示す。この結果から判定精度の検討を行う。保線技術者 7 人のいずれかの判定結果とシステムの出力結果が一致した場合を正解とし、その割合から正解率を算出した。各劣化度ランクにおいて正解率を算出した結果を表 1 に示す。A2 を除くすべての劣化度ランクにおいて 90%以上の判定精度であった。A2 では正解率 81%と他ランクよりも低くなっているが、不正解となった A2 まくらぎ 4 本については B ランク判定となっており、A2~B ランクを不良まくらぎとして管理していることを鑑みると、実用上大きな影響はきたさないと考えられる。また、全てのまくらぎ判定結果を総合して精度を算出した結果は 90.2%となり、高い精度で判定が可能であることが明らかとなった。

上記結果を踏まえ、この評価用まくらぎを標本集団として考えたときに、他まくらぎ画像での劣化度判定精度について推定による検討を行う。いま、標本集団となる評価用画像のまくらぎサンプル本数が 218 であり、このときの判定精度が 90.2%である。したがって信頼度 95%で z 推定を実施した結果、このシステムの判定正解率は信頼区間 86.3%~94.1%の間にあると推定できる。以上より、他のまくらぎ画像に対しても十分な精度で劣化度の判定が可能であるといえる。

〈4・3〉判定結果のばらつきの分析 次に、劣化度のシステム判定結果のばらつきについて、保線技術者の判定結果のばらつきと比較しながら評価を行う。作業員の目視による検査を置き換えるシステムを最終的な目標としているため、開発したシステムのばらつきは、人間の目視による判断結果のばらつきと同程度、あるいはそれよりも小さいことが求められる。前節で図 6 下部に示した劣化度ランクごとの判定結果のヒストグラムからわかるとおり、システムの判定結果にもばらつきが生じている。したがって、シ

表 1 木まくらぎ劣化度の判定割合

ランク	A2	B	C	D	未判定
システム判定結果 (本)	21	49	29	99	6
正解率	81%	90%	97%	90%	100%

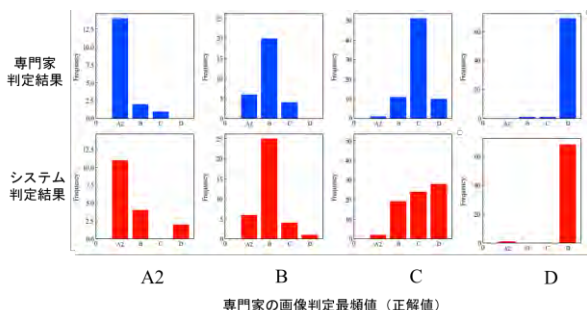


図 6 専門家およびシステムによる判定結果分布

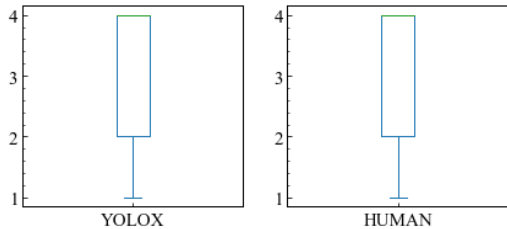


図 7 専門家およびシステムによる判定結果の箱ひげ図

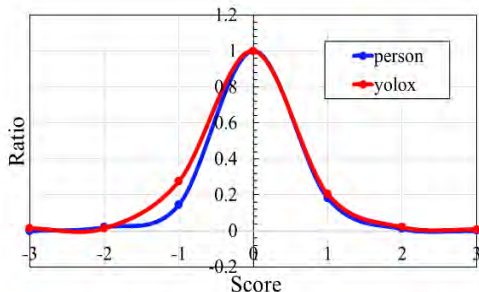


図 8 判定結果のスコアの分布

システムによる判定結果を保線技術者による判定結果と比較し、その分布について統計的解析を実施した。

図 7 にシステムと保線技術者の判定結果の箱ひげ図を示す。以下、システムの結果を YOLOX、保線技術者の結果を HUMAN とする。まず YOLOX の箱ひげ図を見ると、中央値がほぼ HUMAN の中央値と同値であることがわかる。また四分位数についても一致しており、このことからシステムと保線技術者の判定結果は高い類似度を有しているといえる。このことについて定量的に検証するため、効果量を用いて二つの群の差の大きさを評価した。両者の差の効果量（対応あり）を以下の式により算出した。ここで M はサンプルの平均であり、 SD は標準偏差を表す。

$$d = \frac{M_{\text{YOLOX}} - M_{\text{HUMAN}}}{\sqrt{\frac{SD_{\text{YOLOX}}^2 + SD_{\text{HUMAN}}^2}{2}}}$$

計算の結果 $d=0.019$ となり、一般に効果量は 0.2 以下のときに効果量が小さいとされることから、両群の分布の差は小さいと考えられる。

次に両者の正答率のばらつきについて分析する。例えば B ランクのまくらぎについて、システムあるいは保線技術者が B ランクと正答した場合のスコアを 0 とし、A2 ランクと判定をした場合+1、C ランクと判定した場合-1、D ランクとしたとき-2 のスコアをつけるとする。同一の手順で

まくらぎの判定結果をスコア付けした結果のヒストグラムを図 8 に示す。両者のヒストグラムは形状が似ていることが視覚的に読み取れ、このことから全体の傾向として HUMAN の評価を全体の傾向として推定できている可能性がある。そこで、この結果について統計的仮説検定における F 検定により検証する。

F 検定を用いるためには、両群のヒストグラムが正規分布である必要がある。そこで、シャピロウィルク検定を用いて両群を分析した結果、それぞれの p 値が YOLOX で 0.072、HUMAN で 0.064 であった。一般的な有意水準 5% を採用すると、YOLOX、HUMAN とともに p 値が 0.05 より大きいため、二つの分布はともに正規分布であると仮定できる。上記を踏まえ、それぞれの分布について分散の F 検定を実施した。それぞれの標本分散の値は YOLOX が 0.130、HUMAN が 0.132 であり、また標本数は YOLOX が 218、HUMAN が 1,485 であった。帰無仮説として $H_0: \sigma_{\text{YOLOX}}^2 = \sigma_{\text{HUMAN}}^2$ を立て、対立仮説に $H_1: \sigma_{\text{YOLOX}}^2 \neq \sigma_{\text{HUMAN}}^2$ を立てる。有意水準 0.05 で検定すると、 $F_0 = 1.29$ が導かれる。有意水準 0.05 のとき $R = 1.17$ であることから、 $F_0 \geq R$ であり、すなわち帰無仮説 H_0 が棄却され、対立仮説 H_1 が採択される。すなわち、システムの判定結果である YOLOX の判定結果のヒストグラムは、保線技術者による判定結果である HUMAN の判定結果ヒストグラムに類似していることが肯定できるといえる。

〈4・4〉 総合評価 システムと保線技術者との判定結果を比較したところ 〈4・2〉 示した通り 86.3%~94.1% の精度で一致することが示された。また、システムと保線技術者との判定結果のデータ分布の差は小さいこと、両者の正答率のばらつきが類似していることを、それぞれ効果量及びヒストグラムを用いた分析により示した。すなわち、専門の保線技術者の作成した学習データを用いて YOLOX による開発した劣化度モデルが、保線技術者の判定と同様の精度とばらつきを有することが示された。

5. おわりに

まくらぎ検査効率化のため、列車前方面像からまくらぎ検査を行うシステムを提案し、専門家による判定結果との比較評価を行った。このシステムによって従来保線技術者が徒歩巡視によって行っていた作業を省力化し、まくらぎ検査にかかるコストの削減や頻度向上などを旨とする。

参考文献

- 1) 糸井 謙介, 坪川 洋友, 長峯 望, 合田 航, 大場 久良, 加藤 爽: 列車前方面像を用いたまくらぎ劣化度判定アルゴリズムの開発, 土木学会第 76 回年次講演会, 2021
- 2) 合田 航, 長峯 望, 糸井 謙介, 坪川 洋友, 大場 久良, 加藤 爽: 列車前方面像を用いたまくらぎ識別手法, 電気学会資料, TER-20-085 2020
- 3) Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, Jian Sun. "YOLOX: Exceeding YOLO Series in 2021" arXiv preprint arXiv:2107.08430 2021