

ロータ表層の UT 欠陥自動検出モデルの開発

Development of an Automatic Defect Detection Model for Near Surface UT of Rotor Shafts

長谷川 諒*
Ryo Hasegawa博士 (工学) 柳沢 祐介*
Dr. Yusuke Yanagisawa星 延幸**
Nobuyuki Hoshi南谷 昌弘**
Masahiro Minamiya

要 旨

発電所で使用されるタービンロータ軸材（ロータ）は、製造中の内部品質を評価するために、非破壊検査の一つである超音波探傷検査（UT）が行われる。検査工程のリードタイム短縮や省人化を目的として、当社では自動 UT 装置を導入している。ロータ表層の UT 検査において、検査員は本装置から得た B-Scan 画像、および A-Scan のエコー高さから欠陥を目視で確認するが、本作業に時間を要することが課題である。そこで本稿では、検査員の作業負荷の軽減を目的に、B-Scan 画像中の欠陥を自動で検出する機械学習モデルを開発し、実機 20 部材を用いてその予測精度を検証した。その結果、エコー高さ 40 % 以上の欠陥に対する検出漏れ率は 0 %、過検出率は 2.4 % であり、提案手法の実機運用の目途が得られた。

— Synopsis —

Turbine rotor shafts used in power plants are subjected to ultrasonic testing (UT), a type of non-destructive testing, to evaluate their internal quality after manufacturing. We installed an automatic surface UT machine in order to shorten the lead time of the inspection process and to save manpower. When performing near surface UT of rotor shafts, the inspector visually checks for defects based on the B-scan image obtained from the machine and the echo height of the A-scan. However, this process is time-consuming. Therefore, in this paper, we developed a machine learning model to automatically detect defects in B-scan images to reduce the workload of the inspectors and verified its prediction accuracy using 20 actual rotor shafts. The detection failure rate for defects with an echo height of 40 % or greater was 0 %, and the over-detection rate was just 2.4 %. This suggests that the proposed model is operational in practice.

*: イノベーションマネジメント本部 マテリアル技術研究所 / Material Technology Research Laboratory, Innovation Management Headquarters

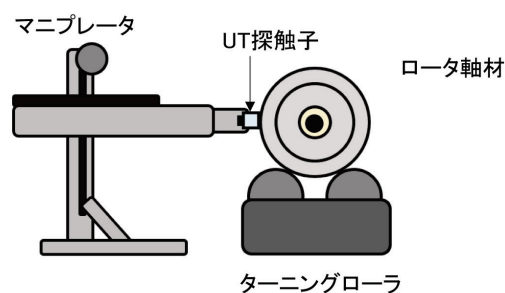
**: 日本製鋼所 M&E (株) 品質管理部 / Quality Control Department, Japan Steel Works M&E, Inc..

1. 緒 言

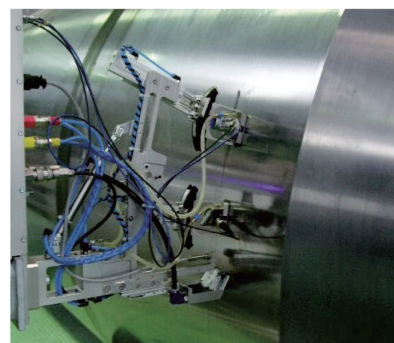
当社が製造する発電所向けタービンロータ軸材（ロータ）では、製造中の内部品質を評価するために、非破壊検査の一つである超音波探傷検査（UT）が行われる。検査員が探触子を手で動かしながら走査する手動探傷では、検査に時間を要することと、熟練した技術が必要となる。そこで当社では、検査工程のリードタイム短縮や省人化を目的に、図 1 に示す自動 UT 装置（TUROMAN5, Actemium Cegelec 製）を導入している^{(1)・(2)}。探触子をロータ表面と接触させた状態でロータを回転させ、探触子の位置をマニプレータで微調整しながら、連続的に検査データを取得する。

自動 UT 装置では最大 $\phi 3000$ mm までのロータの内部検査を実施できるが、本稿では表層約 120 mm 範囲を探傷する表層 UT を検討対象とした。自動 UT 装置から得たエコー波形の一例を図 2 に示す。図 2 (a) に示す A-Scan と呼ばれるエコー波形に着目すると、欠陥(ind.)エコーの近傍に表層特有の疑似エコーが認められ、欠陥の判定を難しくする一因となる。図 2 (b) は A-Scan を周方向に連結してエコー高さで表示した B-Scan 画像であり、図中の赤枠部が欠陥を表す。検査員は部材全体の B-Scan 画像、および A-Scan のエコー高さから、欠陥を判断する。ここで、B-Scan 画像はロータの軸方向に沿って複数枚収録され、検査条件やロータの長手寸法によるが、その枚数は 1000 枚以上になる。そのため、画像の確認作業に時間を要することが課題であった。

UT 検査に機械学習を適用した事例は、これまでにいくつか報告されている^{(3)・(4)}が、材料や欠陥種別が多様であること、画像の規格も現場によって異なること等が要因で、機械学習に必要な計測データを集めることは容易ではない⁽⁵⁾。この問題は、自動 UT 装置によって大量の探傷画像を収集することで解決が見込まれる。本稿では、検査員の作業負荷の軽減を目的に、B-Scan 画像中の欠陥を自動で検出する機械学習モデルを開発し、実機部材を用いてその予測精度を検証した。



(a) 概略図



(b) 探触子拡大図

図 1 タービンロータ軸材の自動 UT 装置

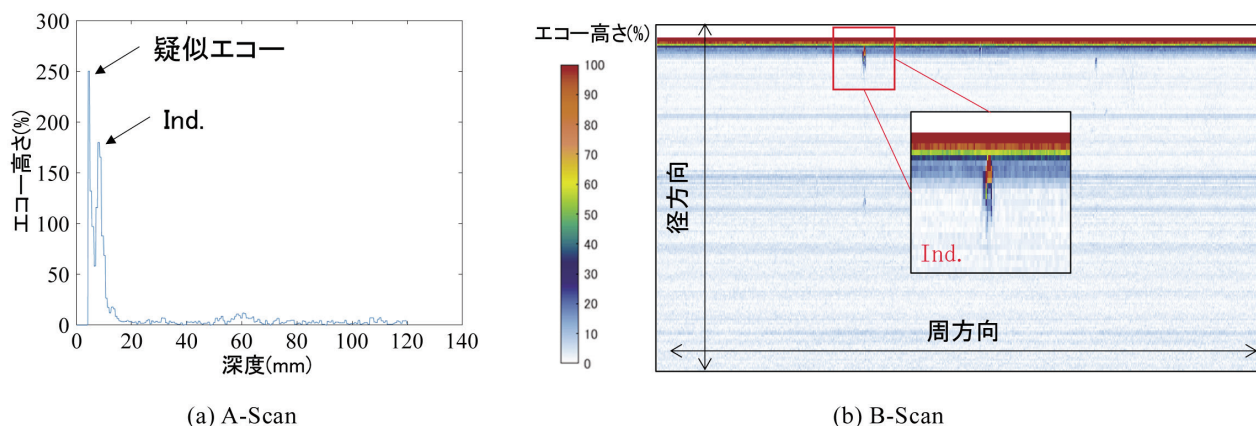


図 2 欠陥部におけるエコー波形の一例

2. 欠陥自動検出手法

2.1 処理フローの概要

機械学習による画像認識の手法は、画像分類、物体検出、セマンティックセグメンテーションに大別される。図 2 に示したように、B-Scan 画像には判別が難しい疑似エコーと欠陥エコーを含むため、物体検出で欠陥を識別するアプローチは困難と考えられる。また、欠陥エコーは B-Scan 画像サイズに対して非常に小さく、セマンティックセグメンテーションによるラベリングは容易ではない。以上を踏まえて、欠陥検出の機械学習モデルは、ルールベースによる判定と、画像分類手法の一つである畳み込みニューラルネットワーク(CNN)を組み合わせたものとした。

機械学習モデルの開発には数値解析ソフトウェア MATLAB R2022a を使用した。本モデルの欠陥評価フローを図 3 に示す。まず、自動 UT 装置から取得した検査データを入力データとして扱い、B-Scan 画像を読み取る。このとき、平均エコー高さに閾値を設けて、明らかなエラー画像は除外した。

B-Scan のデータサイズは、ロータの直径や自動 UT 装置の測定間隔に依存して変動する。そこで、B-Scan 画像を画像データとして読み取りやすくするために、データサイズの補正、RGB 変換等の前処理を行った。

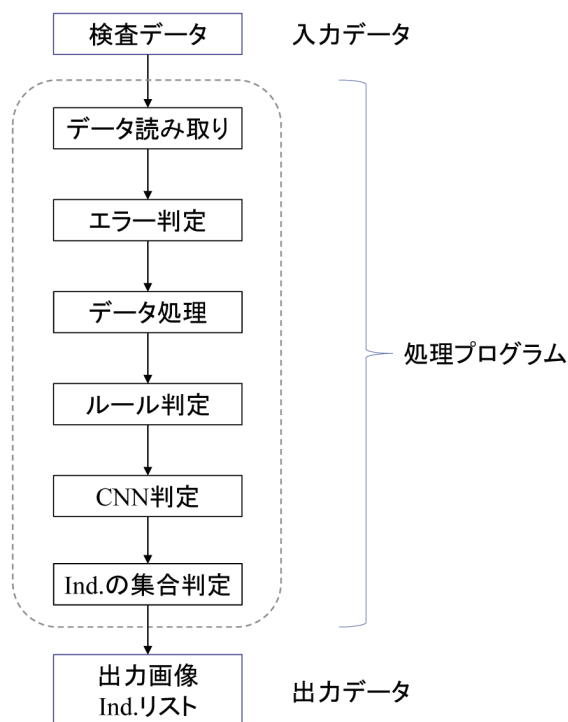
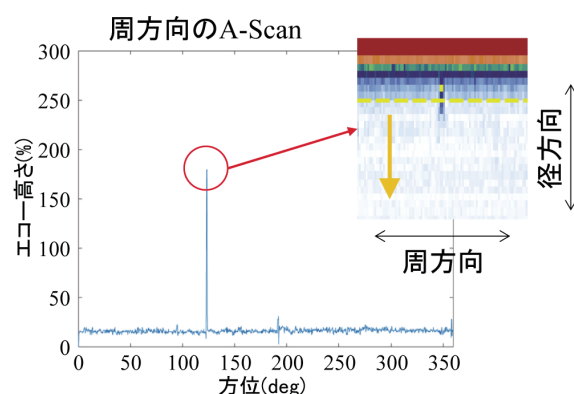


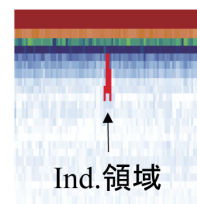
図3 欠陥評価フロー

続いて、データ処理後の画像に対してルール判定を適用し、欠陥の候補を抽出した。図 4 に示すように、ここでは周方向のラインごとにエコーが高い領域を判定し、径方向に連続的に走査することで欠陥領域を検出する。さらに、得られた欠陥領域ごとに最大エコー高さを算出し、検出目標である $\phi 0.8$ mm の欠陥サイズをエコー高さ 100 % として、エコー高さ 40 % 以上の欠陥のみを検出した。

最後に、ルール判定のみでは欠陥数を過大に見積もることから、CNN を用いて欠陥を絞り込んだ。以上の処理フローによって、検査データ中の欠陥座標およびエコー高さの一覧が出力される。



(a) 周方向のエコーピーク位置の検出方法



(b) 検出した欠陥

図4 ルール判定の概略図

2.2 CNN の学習条件

CNN の学習用データセットの作成にあたり、まず実機ロータ 4 部材の B-Scan 画像から、検査員の判定結果を基に正解(欠陥)のラベル付けを実施した。正解ラベルの一例を図 5 に示す。図中の赤枠部は、検査員が評価した欠陥領域を表しており、ボックスラベルのサイズを 128 pixel $\times$ 128 pixel とした。また学習データには、最大エコー高さが 20 % 以上の欠陥のみを扱った。

図 6 のように、学習用データセットには上記の正解ラベルに加え、ランダムに抽出した健全画像および疑似エコー画像を使用した。ここで、データセットに疑似エコーを与えた理由は、疑似エコーが欠陥として誤検出されることを防ぐためである。

CNN の学習条件を表 1 に示す。学習モデルには GoogLeNet を適用し、初期 10 層のパラメータは凍結して最

終層を重点的に学習するファインチューニングを行った。ソルバーはモーメント項付き確率的勾配降下法 (SGDM) とし、学習・検証データの割合は 50 : 50 とした。

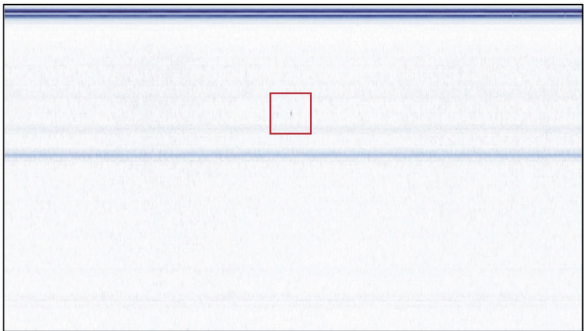
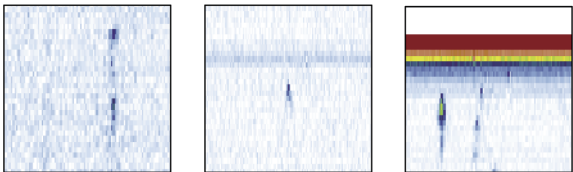
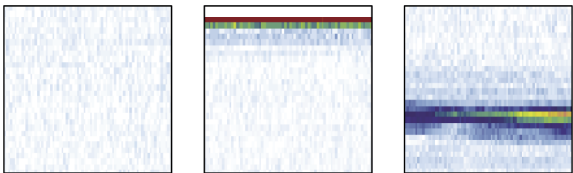


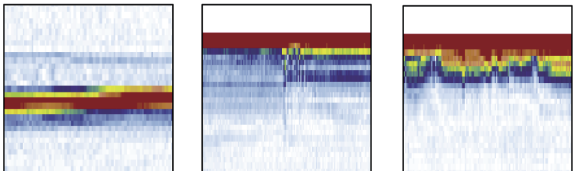
図5 検査員による正解ラベルの一例



(a) 欠陥画像（正解ラベル）



(b) 健全画像（ランダム抽出）



(c) 健全画像（疑似エコー抽出）

図6 CNN の学習に使用した画像例

画像サイズ		128 pixel × 128 pixel
カラー		RGB
欠陥画像枚数		934
健全画像 枚数	ランダム	5085
	疑似エコー	1914
学習モデル		GoogLeNet
ソルバー		SGDM
ミニバッチ数		64
エポック数		30
学習データ：検証データ		50 % ： 50 %

表1 CNN の学習条件

2.3 テストデータを用いた精度検証

CNN における健全画像と欠陥画像の分類精度を、混同行列で評価した結果を図 7 に示す。図 7 (a) から、学習データの再現率・適合率はともに 99 % 以上であり、十分な学習が行えていると判断した。また図 7 (b) より、検証データに対する再現率は 97 %、適合率は 96.2 % と十分に高い精度である。僅かに見逃しが生じているが、これらはエコー高さが 20 % ~ 40 % の範囲の比較的微小な欠陥であった。

推定結果	欠陥	健全	適合率
	11.7%	0.1%	99.6%
	0.1%	88.2%	99.9%
再現率		99.4%	99.9%

欠陥 健全
正解ラベル

(a) 学習データ

推定結果	欠陥	健全	適合率
	11.4%	0.5%	96.2%
	0.4%	87.8%	99.6%
再現率		97.0%	99.2%

欠陥 健全
正解ラベル

(b) 検証データ

図7 混同行列による CNN の精度評価結果

CNN の学習に用いた実機 4 部材の全検査データを対象に、本 CNN を組み込んだ検出手法(図 3)による欠陥の検出精度を評価した。ここではエコー高さ 40 % 以上を欠陥と見なしており、このときの評価結果を表 2 に示す。表中の全画像数は B-Scan 画像の枚数を、欠陥画像数は検査員が実際に評価した欠陥数のことを表している。過検出率および検出漏れ率は、次の(1)、(2)式よりそれぞれ算出した。

$$\text{過検出率} = (\text{欠陥画像を過検出した数}) / (\text{全画像数} - \text{欠陥画像数}) \times 100 \quad \dots (1)$$

$$\text{検出漏れ率} = (\text{欠陥画像を見逃した数}) / (\text{欠陥画像数}) \times 100 \quad \dots (2)$$

過検出率は部材 No. 2 が最大となり、その値は 4.0 % であった。いずれの部材も検出漏れは発生しておらず、本手法では欠陥有無を安全側に評価していることが判る。

部材 No.	全画像数	欠陥画像数	評価結果[%]	
			検出漏れ	過検出
1	1369	73	0.0	3.5
2	1020	19	0.0	4.0
3	1655	6	0.0	0.7
4	1732	82	0.0	3.7
合計	5776	180	0.0	2.8

表 2 テストデータの欠陥検出精度 (エコー高さ 40 % 以上)

3. 実機部材の欠陥分類精度の評価

3.1 評価方法

表 3 に示す実機ロータ 20 部材に提案手法を適用し、欠陥の分類精度を検査員の評価結果と比較した。ここで使用した 20 部材は、前章の学習・検証に用いた 4 部材とは異なるものである。評価対象の欠陥はエコー高さが 40 % 以上のものとした。表 3 における欠陥画像数は、40 % 未満の欠陥を省いた検査結果であり、その総数は 127 個であった。

3.2 検証結果

実機 20 部材の欠陥検出結果を図 8 に示す。推定した欠陥数は計 996 個あり、いずれの部材も検査員が検出した欠陥を全て検出できており、検出漏れは認められなかった。一方で、20 部材分の過検出率の平均値は 2.4 % であり、提案手法は僅かに過検出寄りだった。機械学習モデルの実機運用を実現するには、欠陥の見逃しを絶対に避けなければならない、そのためには欠陥を過検出側に評価する必要がある。これを踏まえると、提案手法の欠陥分類精度は良好だ

部材 No.	全画像数	欠陥画像数	部材 No.	全画像数	欠陥画像数
1	1168	1	11	1889	1
2	2120	10	12	1719	0
3	1751	16	13	2008	7
4	1405	34	14	2042	13
5	1652	5	15	1848	4
6	1531	3	16	1679	0
7	1508	12	17	2453	5
8	2478	2	18	2302	3
9	1123	2	19	972	7
10	1758	0	20	2545	2

表 3 精度検証に使用した実機 20 部材

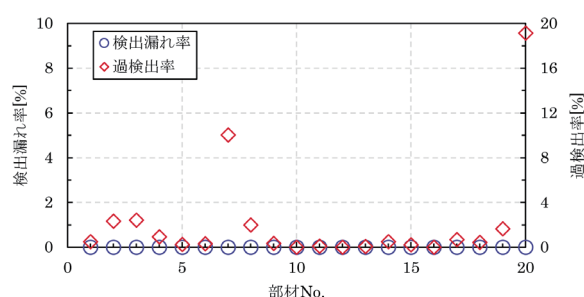
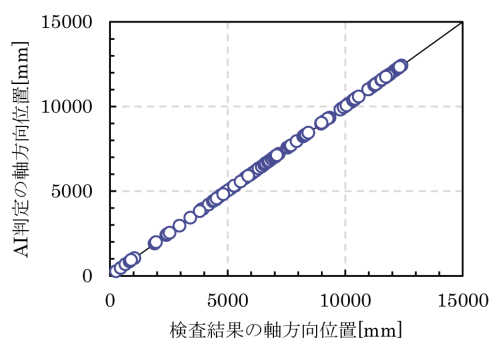


図 8 実機 20 部材の欠陥検出率

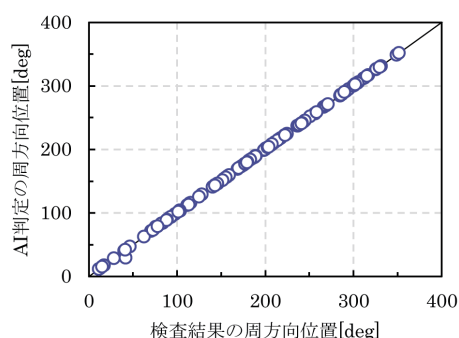
と考える。過検出率は部材ごとにバラつきが見られ、特に部材 No. 7, 20 では 10 % 以上となった。これらの部材は、学習データにはない特徴を有した疑似エコーを欠陥として検出しており、学習が不十分だったことで欠陥と誤認識したと考えられる。

検査員と提案手法が判定した欠陥位置の比較結果を図 9 に示す。なお、軸方向に連続的に検出された欠陥はまとめて一つと扱い、そのエコー高さが最大の位置を代表座標として示した。両者の欠陥位置は、軸・周・径方向ともに概ね一致しており、十分な精度を有することから、自動検出モデルの実機運用の目的が得られた。

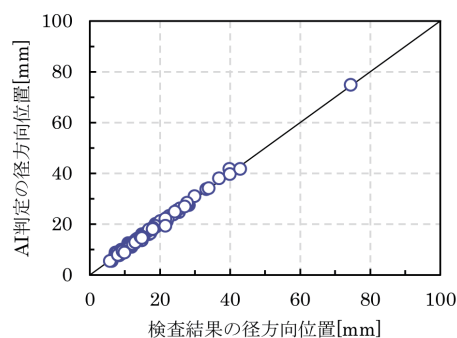
本検討では比較的少ない 4 部材の学習データから推定モデルを作成したが、さらに学習データを増やすことで推定精度の向上が期待される。ただし、疑似エコー画像の学習数を増加させた場合、かえって欠陥画像の分類精度が低下する可能性があるため、学習データに供する各画像数のバランスも重要である。



(a) 軸方向位置



(b) 周方向位置



(c) 径方向位置

図9 自動判定した欠陥位置と検査結果の比較

4. 結 言

本稿では検査員の作業負荷の軽減を目的に、B-Scan 画像中の欠陥を自動で検出する機械学習モデルを開発した。得られた結果を以下に示す。

- 1) 自動検出のアルゴリズムは、ルール判定と CNN を組み合わせたものであり、欠陥の絞り込みを行った後に検査データ中の欠陥座標およびエコー高さの一覧を出力する。
- 2) CNN における健全画像と欠陥画像の分類精度を混同行列で評価した結果、学習・検証データの再現率・適合率はともに 96 % 以上であり、十分な学習を行っていると判断した。
- 3) 実機 20 部材の自動検出結果と検査結果を比較したところ、エコー高さ 40 % 以上の欠陥に対する検出漏れ率は 0 %、過検出率は 2.4 % であった。
- 4) 上記で推定した欠陥位置を検査結果と比較すると、軸・周・径方向ともに概ね一致しており、本手法の実機運用の目的が得られた。

参 考 文 献

- (1) 平順一, 仁村弘樹, 幅寺亮二, 成ヶ澤秀明, 星延幸: “大型タービンロータシャフトへの自動 UT 装置の適用による信頼性の向上”, 日本製鋼所技報, No.60 (2009), pp. 48-53
- (2) 星延幸, 吉田一, 仁村弘樹, 成ヶ澤秀明: “フェーズドアレイ UT による大型ロータシャフトの内部品質評価方法の確立”, 日本製鋼所技報, No.65 (2014), pp. 89-94
- (3) 神田昂亮, 東海林一: “多項ロジスティック回帰を用いたきず判定機械学習プログラムの開発—ステンレス鋼溶接部の応力腐食割れ検出への適用—”, 電力中央研究所報告 (2023), pp. 1-9
- (4) 和田貴裕, 井岡良太, 篠田薫, 片山猛, 安部正光, 竹中俊哉, 服部洋: “管端溶接部の溶接欠陥有無自動判定システムの開発”, Hitz 技報, Vol.81 (2020) No.1, pp. 2-8
- (5) 斎藤隆泰: “NDE4.0 の実現に向けた高性能波動解析技術とデータサイエンスの融合”, 学際大規模情報基盤共同利用・共同研究拠点 2021 年度共同研究 最終報告書, pp. 1-8