# 金属の火花試験の自動評価のための3D CNNの適用

松山 弘樹 † 永田 寅臣 † 加藤 博久 ‡ 渡辺 桂吾 §

†山口東京理科大学大学院工学研究科 ‡山口東京理科大学工学部機械工学科 § 岡山大学 E-mail: nagata@rs.socu.ac.jp

# 1 はじめに

火花試験とは回転する砥石 (グラインダ) と鋼材を接 触させ,研削時に飛び散る火花を観察し,その形状や色 の特徴から鋼材に含まれる成分及び含有量を識別する 試験のことであり, JIS において規格化されている [1]. これまでに画像解析を応用して火花試験の自動評価を 目指した研究がある.例えば,吉岡らは高速度カメラ で撮像した火花の画像解析から鋼材の C%識別技術を 開発した [2].また,山下らは一定条件で火花を発生さ せ,それを撮影するための火花試験装置および撮影さ れた画像から鋼材成分を判別する画像解析プログラム の開発を行った[3].

筆者らは15層という,AlexNetなどの転移学習で得 られた2クラス分類用 CNN と比べると非常に浅い構造 でありながら,ラップフィルムの欠陥検出に有効な2D CNN(sssNet)を提案している[4]. 欠陥検出用の CNN を構築していく上では良品と不良品の分類性能が同じ であれば , 訓練時間 , メンテナンス性 , 拡張性を考慮す るとより浅く,よりフィルタ数が少ない構造であるこ とが望ましい [5].本論文では,火花試験の映像データ から鋼材の炭素含有量を自動評価することができる 3D CNN モデルの提案と,動画から抽出した 3D ボリュー ムのフレーム数および 3D フィルタの深さが分類性能に 与える効果について報告する.



Figure 1 Movie samples used for steel spark test.



Figure 2 Conceptual diagrams of 2D CNN and 3D CNN

#### 火花試験のための 3D CNN の設計 $\mathbf{2}$

#### 2.1 3D CNN について

本研究では,鋼材の炭素含有量(C%)によって変化 する図1のような3種類の火花の映像データを用いた. C7, C19, C26 は分類ラベルとしての表記であり,実際 の炭素含有量はそれぞれ 0.07 %, 0.19 %, 0.26 %であ る.それぞれの映像データを10フレームあるいは20 フレームごとに分割,抽出し,さらに解像度をダウンサ イジングした 3D ボリュームを訓練用データセットとし て用いて 3D CNN の学習を行う.学習時には動きの特 徴抽出に重要な3Dフィルタの深さを変化させて訓練を 行い,未学習のテスト用データセットをどの程度正しく 分類できるかを評価する. 3D CNN は従来の CNN に 時間情報を加え,畳み込み層を3次元に拡張すること で,2次元空間上の特徴に加えて時系列に沿った特徴も 抽出することができる.図2に2DCNNと比較した場 合の 3D CNN の概念図を示す. 2D CNN では,入力画 像に対して畳み込み処理,プーリング層による次元圧縮 を行う. 一方, 3D CNN の場合は,時間軸の方向にフ レーム数による深さ (Depth) の情報が追加され, 動画 から抽出後にリサイズした 3D ボリューム  $(H \times W \times D)$ を入力とすることができる.この 3D ボリュームに 3D フィルタ  $(h \times w \times d)$  を適用し畳み込みを行い,その後 のプーリング層でボリュームサイズの圧縮を行ってい る[6].

Off On OMRON CAM Categories and nu	mber of	f images fo	or train	ning set Number of categories 2 Close
Auto Classification Option Help OK 100	N5	0	~	Help
Cropping from center NG 100	N6	0	(I)	Folders for training(OK, NG,), for classified images(OK3, NG3,)
O CNN SVM 0.3 Period [s] N1 0	N7	0	_	
Ethernet(TCP/IP) or USB	N8	0	2	Folders for test images(OK2,NG2)
Help Help Help N3 0	N9	0	Hole	
CheckneintBath BatebNermalizationStatistics N4 0	NA	0	11. SVI	M for One Class Learning Nu 0.2 OutlierFraction Test (③Folder)
population moving	1 50 1			
Over fitting check & its Auto stop		JD C	:NN (s	s V Crocker Threshold linear V Order Test (an image)
MaxEpochs 80 Accuracy 0.999 GAN one label	Help	Auto Opt	imizatio	ion of TCSVM Standardize $\gamma$ 0 Bias 0
MiniBatchSize 30 Loss 0.001 GAN multi labe	Help Is	13. SVM f	for Two	Class Learning C 1 Test (③Folder) Check of input image
Allowed times of not Improving 20	Catego	ies and nu	umber	of images for test set
LearnRateSchedule GPU	OK2	1	N52	
InitialLearnRate 0.0001 ONone Single Multi	NG2	1	N62	0 BMP 14. Transfer Learning ⇒ sssNet Not transfer
LearnRateDropPeriod 3 GPU None	N12	0	N72	0 PNG LearnRateFactor of fully connect layer 10
LearnRateDropFactor 0.1 Help multi-gpu	N22	0	N82	0 TIF Freeze weight for 1~ 10 layers Help
Montage 5 Test of CNN with a test image or volume	- N37		NIQ2	O MP4 Augmentation Help Test (@Folder)
Eile name	P N/32		NA2	Figures check!
	1942			Validation Frequency 3 Copy
Width 0 Height 0 Channel 0 Frame 0	Convo	lution layer	r 2	Help analyzeNetwork
Plot training-progress	Num	Channels	1	Filter's feature 1 Filter's feat
Plot scores 0 0 minutes	Nu	umFilters	1 📤	Filter's feature 2 Close Otherwise 0.5
Finish Training ! Pinpoint training for	Lave	r's structure	•	all figures Scores Not trained samples
7 Additional Training of CNN scores under 0.5 (fi	IterSize	x filterSize	x Num	nChannels x NumFilters)
MaxEpochs 10	A			16. Convolutional Test (3)
Note Accuracy 0.7	Auto Tra	ining of Mu		Auto Encoder (CAE) Folder)
ONNX_ex 8. Save AI (*.mat)				BlockSize 224
ONNX_im 9. Load AI (*.mat)				Help
Help				Result 123
10. Classification ⇒ ③ Folder				Help
Classified images are copied into the folders (OK3, NG3, N13, N23,	)	Check of "	'OK" or	r "Otherwise" using 2 CNNs for binary classification Yolo

Figure 3 A main dialog developed on MATLAB system to user-friendly design 3D CNNs.



Figure 4 Our original 3D CNN model named sssNet\_2A designed using application shown in Fig.3.

# 2.2 3D CNN の設計

図 3 には MATLAB 上で開発した 2D と 3D の CNN の設計,訓練,転移学習などを行うことができるソフト ウェアを示す.筆者らはこのソフトウェアを用いて 15 層という非常に浅い構造の 2D CNN である sssNet を設 計し,欠陥検出への有用性を確かめている [4].動画か ら抽出した 3D ボリュームを扱うことができる 3D CNN は,この sssNet をベースとして 10 フレームの訓練デー タを用いた 3D CNN(sssNet\_2A, sssNet\_3A)を 2 種類, 20 フレームの訓練データを用いた 3D CNN(sssNet\_2B, sssNet\_3B, sssNet\_4B) を 3 種類設計する.それぞれの 3D フィルタの深さは,2,3,2,3 および 4 である.例 として図 4 に sssNet\_2A の構造を示す.これら 5 つの 3D CNN はともに 3 つの 3 次元の畳み込み層と 2 つの 全結合層を含む全 15 層から構成されており,入力層には 108 × 192 × 10 × 3 のサイズのボリュームが入力され,32 個の 5 × 5 × 2 × 3 の 3D フィルタにより 32 個 の特徴ボリュームを生成する.第2,第3 の畳み込み層



Figure 5 A part of main dialog developed on MATLAB system to extract 3D volumes from videos.

も同様に 3D フィルタと同数の特徴ボリュームを生成する. なお, 全結合層は 32 × 3 個の重みを有する.

3 3D ボリュームの抽出と3D CNN の訓練

### 3.1 訓練のための 10 フレームから構成される 3D ボ リュームの生成

図5にはMATLAB上で開発した画像オーギュメン テーションソフトウェアのメインダイアログの一部を 示す.主に画像処理を加えて効率的な訓練データセット の水増しを行ったり,テンプレートマッチングを適用 するために開発されたアプリケーションであるが, 3D CNN の学習にあたり映像データから 3D ボリュームを 生成する新たな機能を追加した.まず,(5) Folder to save converted images のボタンをクリックし, 生成さ れる 3D ボリュームの保存先フォルダを選択する.3D のチェックボックスにチェックすることにより,その隣 にある Depth の設定ボックスがアクティブになり,1 つの 3D ボリュームに含ませたいフレーム数を設定す ることができる.また,Width,Heightでは,フレー ムの解像度を自由に変更することができる.C7,C19, C26, 3 つの mp4 データの再生時間はそれぞれ, 49 秒, 84秒,44秒で解像度は1920×1080,フレームレート は 30 fps である.今回は 3D ボリュームの時間軸方向 にあたるフレーム数を10に設定した場合,生成された 3D ボリュームの数はそれぞれ 147, 253, 136 であった. また,20 フレームに設定した場合に抽出できた 3D ボ リュームの数はそれぞれ, 73, 126, 67 であった. なお, 学習にかかる時間的な負荷とメモリ消費の負荷を抑え るために訓練に使用するフレームの解像度は192×108 にダウンサイジングした.これは同じ分類性能が得られ るのであれば , より規模の小さいデータセットで CNN を構築したいというニーズに応えるものである.

#### 3.2 3D CNN の訓練

まず, C7, C19, C26 の映像データから生成した 3D ボリュームのうち, それぞれ 122 個, 210 個, 111 個, 計 443 個を訓練用データセット A として用意する.学 習条件は最大エポックを 80, ミニバッチサイズを 30, 学習率を 0.0001 に設定して訓練を行い, sssNet\_2A と sssNet\_3A を得る.例として図 6 に sssNet\_2A におけ るミニバッチ学習の反復数に対する精度と損失の関係 を示す.なお, 37 エポックの時点で精度が 100 %に達



Figure 6 Accuracy and loss versus iterations during mini-batch learning of sssNet\_2A.

# したので学習を終了させている. 3.3 3D CNN による分類結果

学習を終えて得られた sssNet\_2A と sssNet\_3A の汎 化性能を評価するために C7, C19, C26 の 3 クラスで それぞれ 25 個,43 個,23 個,計 91 個の 3D ボリュー ムから成るテスト用データセット A を用いた.表1に は sssNet\_2A の全体的な認識結果を示す混同行列を示 す.また,表2 にも同様に sssNet\_3A の混同行列を示 す.sssNet\_2A, sssNet\_3A はともに同じ結果が得られ た.C7, C19 のクラスにおいては100 %の精度を示し ている.しかし,C26 クラスにおいては2 個のデータ を C7 クラスと誤判定している.

Table 1 Confusion matrix classified by sss-Net\_2A.

Predicted True	C7	C19	C26
C7	25	0	0
C19	0	43	0
C26	2	0	21

Table 2Confusion matrix classified by sss-Net\_3A.

Predicted True	C7	C19	C26
C7	25	0	0
C19	0	43	0
C26	2	0	21

# 4 20 フレームで構成された 3D ボリューム の場合

さらに深い 3D フィルタを適用するためには,入力される 3D ボリュームのフレーム数を増やす必要がある、図5 で示したソフトウェアを用いて,C7,C19,C26のmp4 データからフレーム数を20 に設定した 3D ボリュームを生成する.このとき,生成された 3D ボリュームの数はそれぞれ73,126,67 であった.訓練用とテスト用

に9:1の割合で抽出した3Dボリュームを分割し,それ ぞれを訓練用データセット B, テスト用データセット B とする.訓練用データセットBは3クラスでそれぞれ66 個,114個,60個,計240個の3Dボリュームから成る. 深さが 2, 3, 4 の異なる 3D フィルタを適用した sssNet において ,それぞれ訓練用データセット B で訓練するこ とで,sssNet\_2B,sssNet\_3B,sssNet\_4Bを得る.学習条 件は訓練用データセット A を用いた場合と同一とした. その後,テスト用データセットBを用いて sssNet\_2B, sssNet\_3B、sssNet\_4B による分類実験を行い, 汎化性 能を評価・比較する.表3、表4および表5にはそれぞ れ,sssNet\_2B,sssNet\_3B,およびsssNet\_4Bの分類結 果を表す混同行列を示す.sssNet\_2BとsssNet\_3Bでは 同様の結果が得られた. C7, C19 のクラスにおいては 100%の精度を示しているが, C26クラスにおいては1 個のデータを C7 クラスと誤判定している.さらに深い 3D フィルタを適用した sssNet\_4B ではその誤判定も改 善され , 全てのクラスにおいて 100 %の分類精度を示 している.

Table 3 Confusion matrix classified by sss-Net\_2B.

Predicted True	C7	C19	C26
C7	7	0	0
C19	0	12	0
C26	1	0	6

Table 4 Confusion matrix classified by sss-Net\_3B.

Predicted True	С7	C19	C26
C7	7	0	0
C19	0	12	0
C26	1	0	6

Table 5 Confusion matrix classified by sss-Net\_4B.

Predicted True	C7	C19	C26
C7	7	0	0
C19	0	12	0
C26	0	0	7

# 5 結言

本研究では非常に浅い構造をもつオリジナルの 3D CNN を設計し,火花試験における C %の自動分類を行 うことができるシステムを提案した.C7,C19,C26の 3 クラスを対象として火花の mp4 データから 10 フレー ムと 20 フレームを含む 3D ボリュームを抽出すると同 時に,ダウンサイジングすることで訓練で扱えるフォー マットに変換した.オリジナルの 3D CNN(sssNet) は 畳み込み層における 3D フィルタの深さが 2,3 および 4 のように異なる 3 種類のものを設計し,3D ボリュー ムで構成したデータセット A とデータセット B を用い て訓練し,最終的に sssNet\_2A, sssNet\_3A, sssNet\_2B, sssNet\_3B および sssNet\_4B を得た.その後,訓練用 データセットに含まれていないテスト用データセット を用いて 5 種類の sssNet を評価,比較した.その結果, 20 フレームで構成した 3D ボリュームに深さ4の3D フィルタを適用した場合に高精度な分類性能が示され た.このように,3D フィルタを深くすることでテスト 用の3D ボリュームの分類に対する汎化性能の向上がみ られたため,今後は更に深い3D ボリュームと3D フィ ルタを適用し,汎化性能にどのような効果を及ぼすの かを検討していく.

### References

- [1] 日本工業規格,社団法人日本鉄鋼連盟,"JIS G0566-1980 鋼の火花試験方法".
- [2] 吉岡孝宜,大場康英,小林宏,"画像解析を用いた火花試験による鋼種識別システムの開発",山陽特殊製鋼技報 (Sanyo Technical Report), Vol. 26, No. 1, pp. 23–27, 2019.
- [3] 山下 裕貴,小林 宏,"火花に基づく鋼材の自動識別 火花試験の自動化",ロボティクス・メカトロニクス講演会講演概要集,2A1-V03,2014.
- [4] K. Nakashima, F. Nagata, A. Otsuka, K. Watanabe, M. K. Habib, "Defect Detection in Wrap Film Product Using Compact Convolutional Neural Network," *Artificial Life and Robotics*, Vol. 26, No. 3, pp. 360–366, Springer, 2021.
- [5] 松山 弘樹,永田 寅臣,三木 康平,加藤 博久,渡辺 桂吾,"3D CNN による金属の火花試験の評価シス テムの提案",第22回計測自動制御学会システムイ ンテグレーション部門講演会論文集,pp. 328-331, 2021.
- [6] Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, Manohar Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks," *Proceedings of the IEEE International Conference on Computer Vision (ICCV2015)*, pp. 4489– 4497, 2015.