

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.

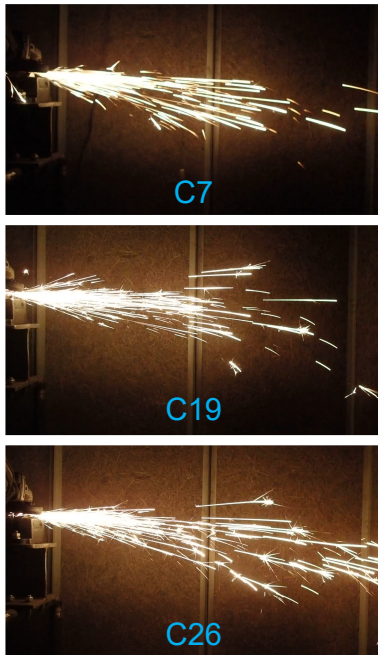


Figure 1 Movie samples used for steel spark test.

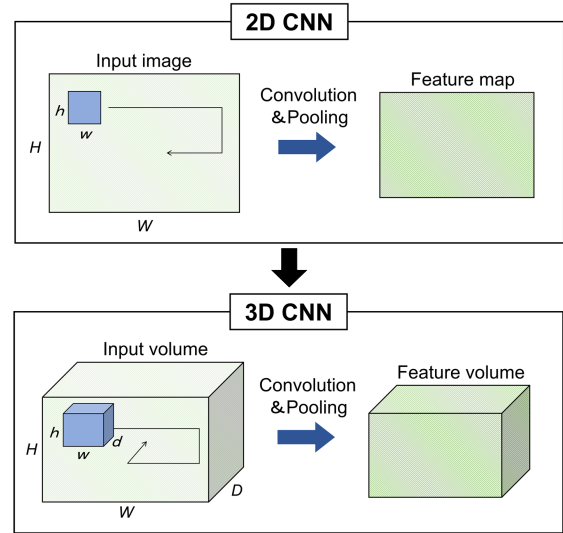


Figure 2 Conceptual diagrams of 2D CNN and 3D CNN

2 火花試験のための 3D CNN の設計

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 2D CNN 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].

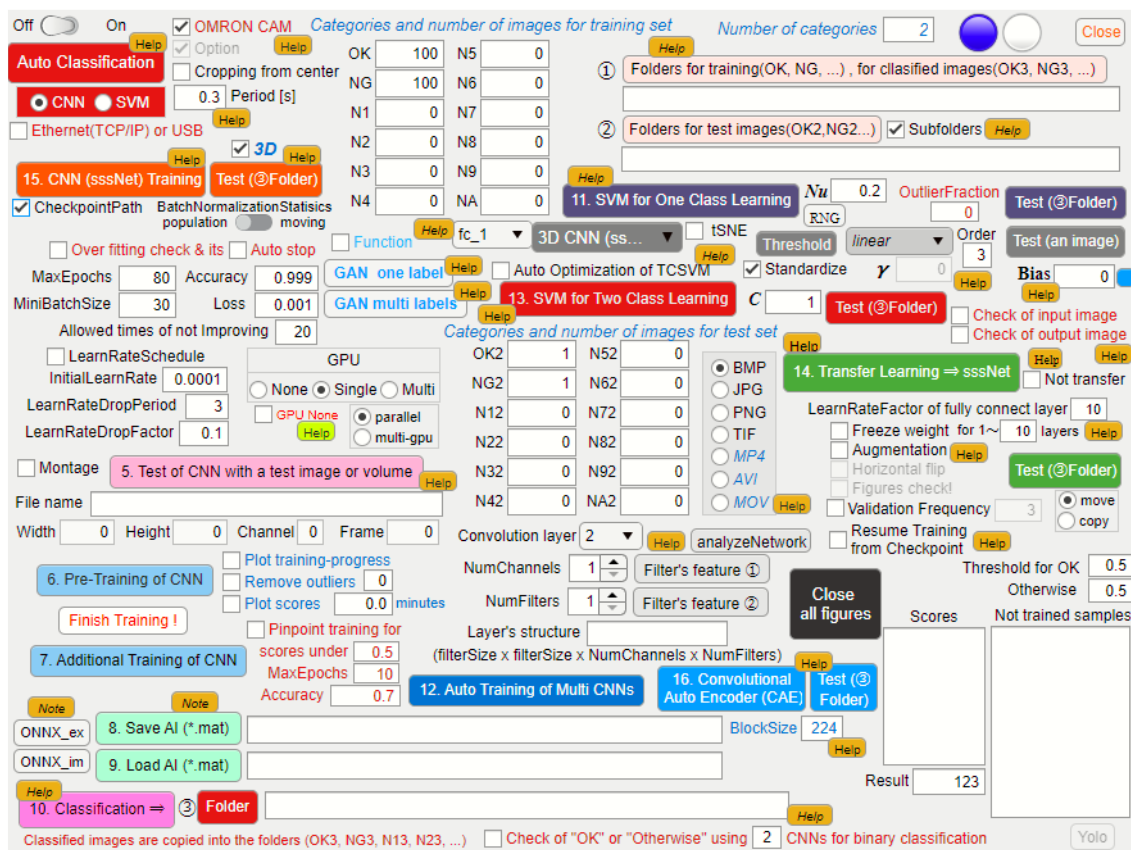


Figure 3 A main dialog developed on MATLAB system to user-friendly design 3D CNNs.

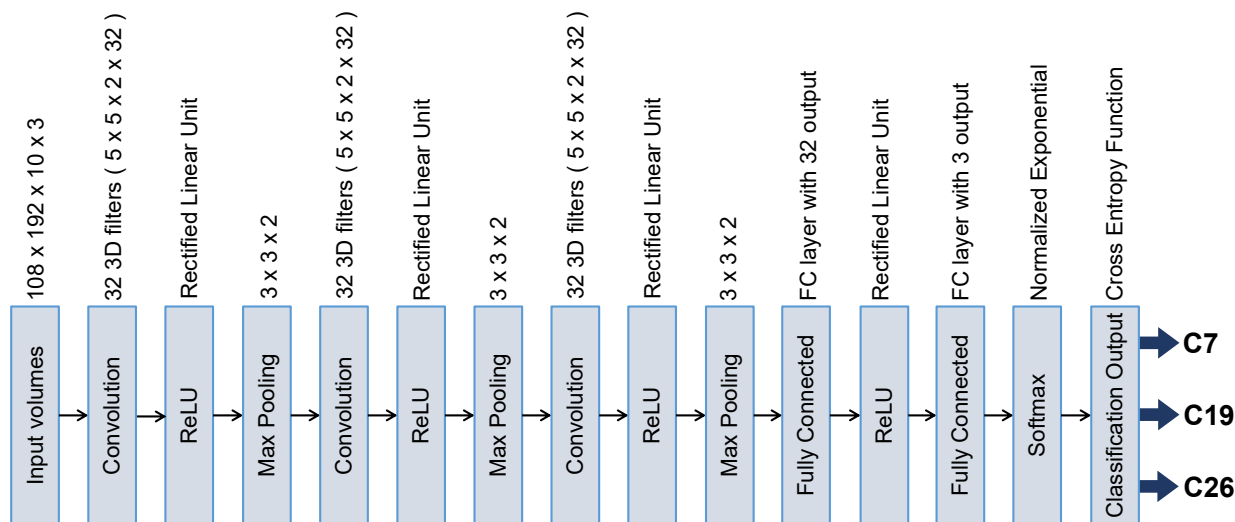


Figure 4 Our original 3D CNN model named sssNet_2A designed using application shown in Fig.3.

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 $108 \times 192 \times 10 \times 3$ のサイズのボリュームが入力され，32 個の $5 \times 5 \times 2 \times 3$ 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

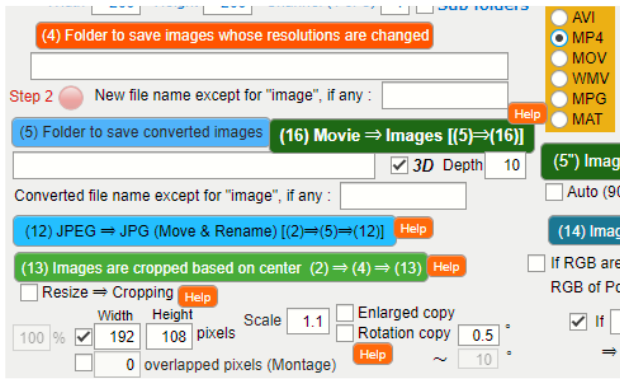


Figure 5 A part of main dialog developed on MATLAB system to extract 3D volumes from videos.

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

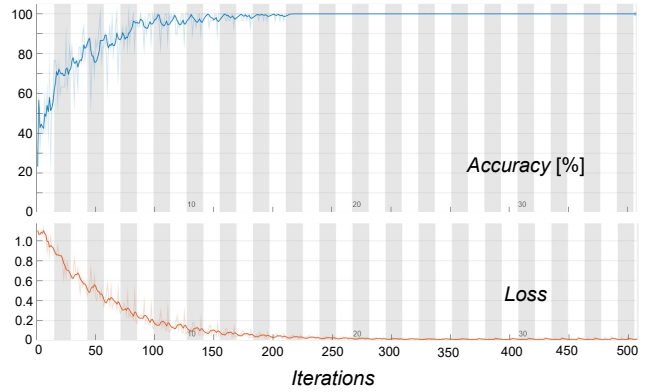


Figure 6 Accuracy and loss versus iterations during mini-batch learning of sssNet_2A.

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 sssNet_2A.

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

Table 2 Confusion matrix classified by sssNet_3A.

True \ Predicted	Predicted		
	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.

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

Table 4 Confusion matrix classified by sss-Net_3B.

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

Table 5 Confusion matrix classified by sss-Net_4B.

True \ Predicted	Predicted		
	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ポリ

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.