

3D CNN による金属の火花試験の評価システムの提案

松山 弘樹（山口東京理科大学工学部機械工学科），永田 寅臣（山口東京理科大学大学院）

三木 康平（山口東京理科大学大学院），加藤 博久（山口東京理科大学工学部機械工学科）

渡辺 桂吾（岡山大学大学院）

Evaluation System Using 3D Convolutional Neural Network for Spark Test

Hiroki MATSUYAM (Sanyo-Onoda City University)

Fusaomi NAGATA (Sanyo-Onoda City University), Kohei MIKI (Sanyo-Onoda City University)

Hirohisa KATO (Sanyo-Onoda City University), Keigo WATANABE (Okayama University)

Abstract : This paper proposes a system processed by a 3D Convolutional Neural Network (3D CNN) that adds information in time domain to conventional CNN. The 3D CNN can deal with motion and dynamics of various kinds of physical phenomena. The authors apply it to the evaluation of spark test for steels. In this experiment, the steel spark images are classified into three grades such as C7, C19 and C26, whose number means the carbon content rate. The format of training data for C7, C19 and C26 is 192 width \times 108 height \times 10 frames, which are called short 3D volumes extracted from mp4 videos. The effectiveness and promise of our proposed system are evaluated and proved through classification experiments using the designed 3D CNN.

1. 緒言

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。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 モデルを提案し、その分類性能を報告する。

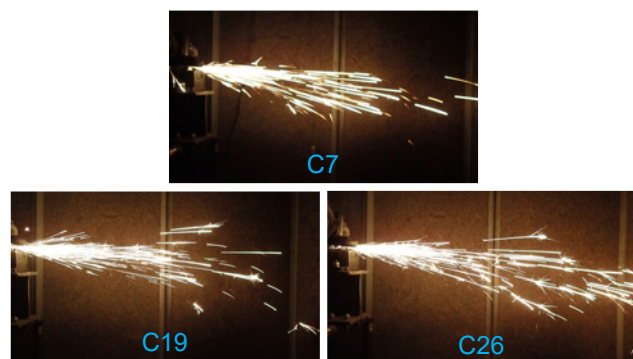


Fig. 1: Image samples used for steel spark test.

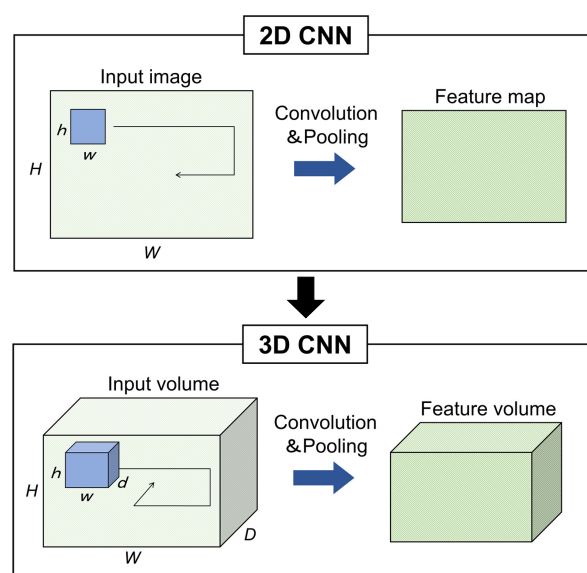


Fig. 2: Conceptual diagrams of 2D CNN and 3D CNN

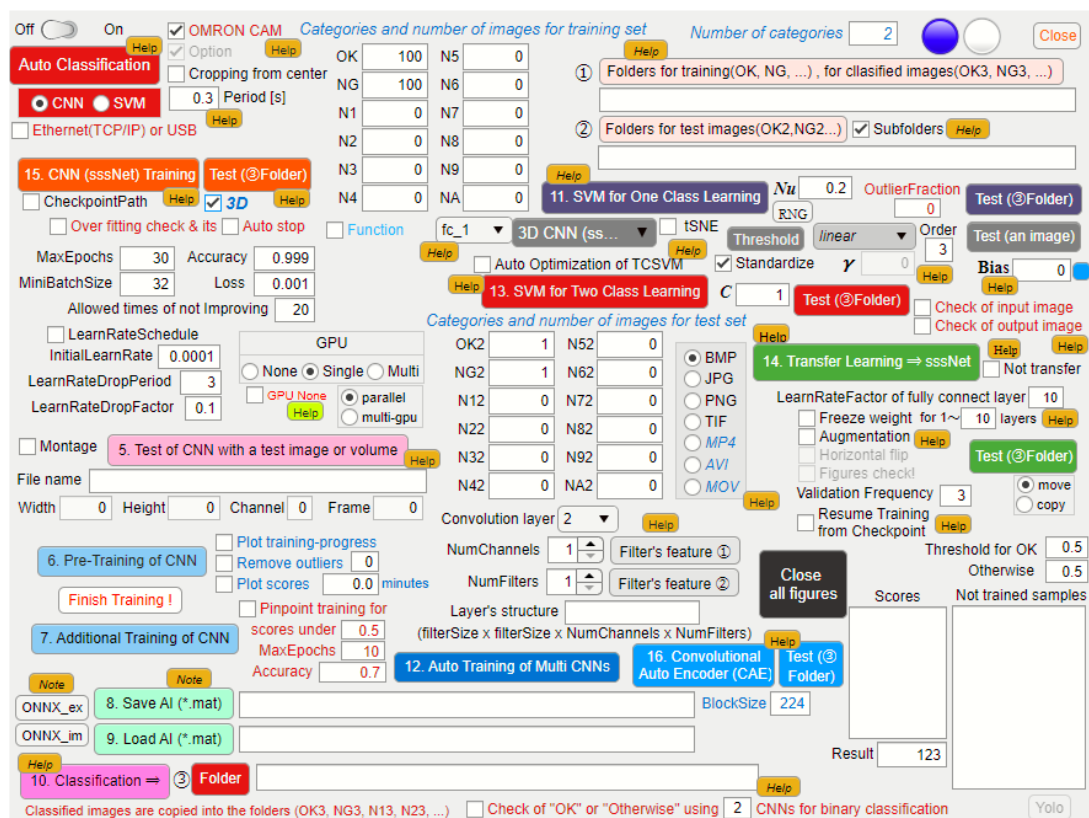


Fig. 3: Main dialog developed on MATLAB system to user-friendly design CNNs.

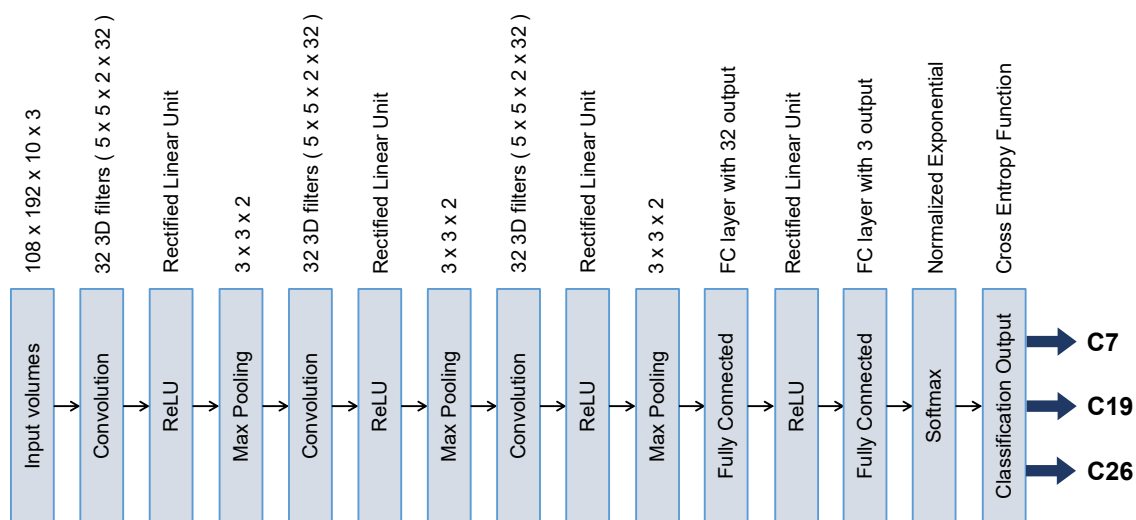


Fig. 4: Our original 3D CNN named 3DsssNet_2 designed using application shown in Fig.3

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．ただし，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．そして，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 [5]．

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2種類設計する．1つは図4に示すように畳み込み層における3D フィルタのDepthが2である3DsssNet_2．もう1つはDepthが3の3D フィルタをもつ3DsssNet_3である．3DsssNet_2, 3DsssNet_3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 ボリュームの生成と3D CNN の訓練

3.1 訓練のための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．

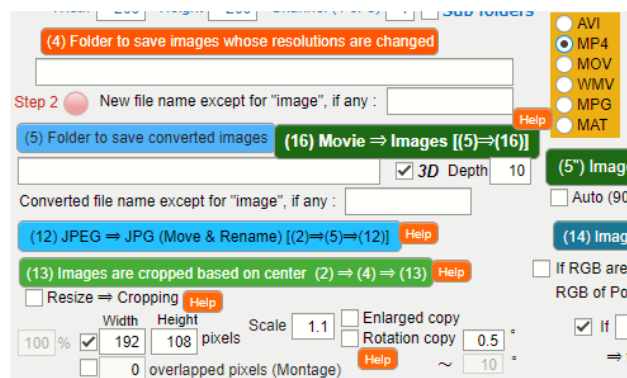


Fig. 5: A part of main dialog developed on MATLAB system to extract 3D volumes from videos.

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Depthを10に設定したところ，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，学習にかかる時間的，容量的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3DsssNet_2と3DsssNet_3の訓練を行った．例として図6に3DsssNet_2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%に達したので学習を終了させている．

4. 3D CNN による分類結果

学習を終えた2つの3DsssNetの汎化性能を評価するためにテスト用データセットとしてC7, C19, C26の3クラスでそれぞれ25個，43個，23個，計91個の3D ボリュームを用いた．表1には3DsssNet_2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3DsssNet_3の混同行列を示す．3DsssNet_2, 3DsssNet_3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

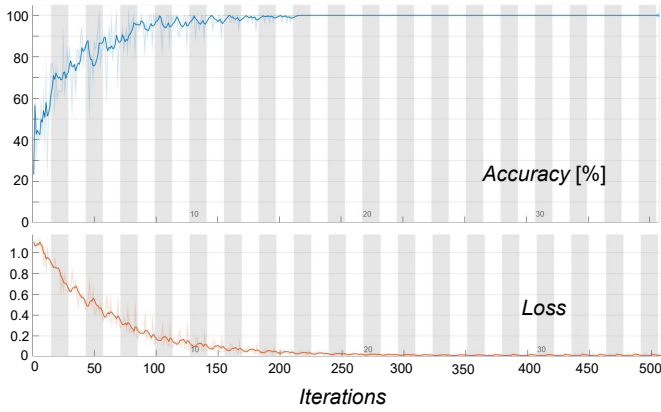


Fig. 6: Accuracy and loss versus iterations for mini-batch learning of 3DsssNet_2.

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 3DsssNet_2.

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

Table 2: Confusion matrix classified by 3DsssNet_3.

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

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試験を撮影した mp4 データから 10 フレームで構成する 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，訓練用データセットを作成した．オリジナルの 3D CNN は畳み込み層の Depth が異なる 2 種類 (3DsssNet_2, 3DsssNet_3) のものを設計し，このデータセットを用いて訓練した．そして，別に用意したテスト用データセットを用いて 2 種類の 3DsssNet を評価，比較した．3DsssNet_

2, 3DsssNet_3 はともに高精度な分類性能を示し，製造現場への適用可能性が確認できた．また，3D フィルタの Depth が異なるものの，それによる分類結果の違いは見受けられなかった．本研究では生成した 3D ボリュームの Depth を 10 に設定したため，3D フィルタの Depth を 4 以上深くすることができなかった．今後は 10 フレームよりも深い 3D ボリュームを訓練データとして生成し，3D フィルタのサイズを深くすることで分類精度にどのような影響を及ぼすのかを評価していく．

参考文献

- [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, pp. 360-366, 2021.
- [5] D. Tran, L. Bourdev, R. Fergus, L. Torresani, M. Paluri, “Learning Spatiotemporal Features With 3D Convolutional Networks,” *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 4489-4497, 2015.