

金属の火花試験の自動評価のための 3D CNN と 2D CNN の分類性能の比較

永田研究室 F119055 廣野 聖和

1. 目的

近年、工業製品の検査工程において、深層学習の技術を画像認識に特化させた畳み込みニューラルネットワーク (CNN) を応用した自動化のニーズが高まっており、鋼材の研削時に飛び散る火花の特徴から金属材料の鋼種を判定する火花試験の自動評価にも応用されている。空間上の特徴に加えて時間軸に沿った特徴も抽出することができる 3D CNN の構築には、3D ボリュームと呼ばれるショートムービーが訓練データ、テストデータに用いられるために画像をベースに構築されている 2D CNN に比べて多くの計算コスト、メモリコスト、管理コストが必要になるといった課題がある。本研究では、2D CNN でも 3D CNN と同等の火花の識別性能を発揮させるための動画からの特徴抽出法を提案する。鋼材の火花試験の映像データから訓練データを生成し、設計した 3D CNN モデルと 2D CNN モデルに学習させ、それぞれ 10 クラスの分類実験を行い、提案手法の有効性を検証する。

2. 研究内容

企業より提供のあった 10 種類の炭素含有量の異なる鋼材を用いた火花試験の映像データから 3D CNN 用と 2D CNN 用の訓練データを生成する。各動画から 10 フレームの 3D ボリュームを 1 フレームずつずらしながら抽出し、訓練データ (30,883 枚) とテストデータ (30,870 枚) に分割する。このとき、解像度をオリジナルの 3,840×2,160 から 192×108 にダウンサイジングし、これをデータセット A とする。次に、2D CNN 用の訓練データを生成する。ここでは、指定したフレームの中から火花の活性化度が最も高いフレームのみを抽出する方法を提案する。各動画から 10 フレーム毎に最も活性化度の高いフレームを 1 枚ずつ抽出し、訓練データ (4,982 枚) とテストデータ (2,124 枚) に分割する。このとき、CNN の入力層を考慮して解像度を 224×224 にダウンサイジングし、これをデータセット B とする。学習については、研究室オリジナルのモデルである sssNet を全 25 層に拡張した CNN を設計し、データセット A を用いて追加学習を行う。また、公開されている CNN モデルである VGG19 の最終の全結合層部を転移学習により 10 クラス分類用に書き換えた CNN を設計し、データセット B を用いて追加学習を行う。なお、データセット A および B を用いた訓練で得られた CNN をそれぞれ、sssNet_3D, sssNet_2D とする。訓練後の各 CNN モデルの汎化性能を評価するためにテストデータを用いて分類実験を行う。

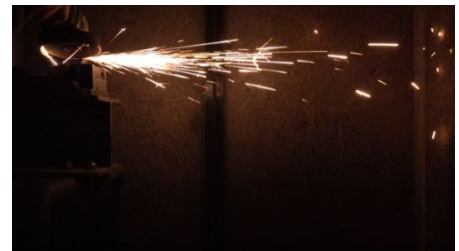


Fig. 1 Video sample of steel spark test.

3. 結果

図 2 には sssNet_2D の分類結果を表す混同行列を示す。縦軸と横軸はそれぞれ真のラベルと予測ラベルを表している。なお、ラベル名にはそれぞれの金属の炭素含有量の数値 (%) を 100 倍し、先頭に C を付けたもので表している。sssNet_2D を用いた場合、誤分類された画像は 11 枚で、テストデータに対する正解率を表す認識率は 99.48 % だった。一方、sssNet_3D の場合、誤分類された画像は 3 枚で、認識率は 99.99 % だった。このように 2D CNN でも 3D CNN に近い高い分類精度で金属の炭素含有量を推定することができており、提案する特徴抽出法により金属の火花試験の自動評価における 2D CNN の有効性を確認できた。今後は、分類クラス数を増やしていくと同時に安定した分類性能を発揮できる CNN を構築し、CNN を利用した火花試験の自動評価の実用化を目指していきたい。

	C4	C5	C6	C7	C8	C9	C10	C12	C13	C15
C4	253									
C5		252	2	1						
C6	1	1	267					1		
C7				300						
C8					158		1			
C9					1	271				
C10				1			236			
C12				1		1		199		
C13									90	
C15										87
	C4	C5	C6	C7	C8	C9	C10	C12	C13	C15

Fig. 2 Classification result of sssNet_2D given by confusion matrix.