画像解析を利用した生産現場の状態監視技術に関する研究（第３報）

生産技術部
中野太郎　大坪昭文　福島章吾

本研究は，生産現場における確認・監視作業に画像解析技術を適用することで，製造工程の状態の把握や管理を自動化する技術開発に取り組むものである．本年度は，深層畳み込みニューラルネットワーク (Deep Convolution Neural Networks) を転移学習し，旋盤加工の工程分類問題に適用する手法について検討した．その結果，複数の製造工程パターンを含む検証用のデータセットに対し，工程分類の正解率が 92.7〜99.8%という高い識別性能が得られた．また，未知のデータに対する汎化性能においても SVM（Suport Vector Machine）よりも大幅に優れていることを確認した．さらに，構築した画像解析モデルを組み込んだ自律型の工程解析エッジデバイスを製作し，近距離無線通信などの IoT 技術と連携させることで，データの収集・蓄積・可視化までをオートメーション化した工程進捗の監視・計測システムを開発した．

1. はじめに

ものづくりの生産現場においては，生産設備の稼働状態や製造品の外観品質などを恒常的に確認する必要がある．また，品質管理や生産性の改善活動に活用するために，加工工程の進捗状況をより細分化して把握することが求められている．しかし，その多くを目視確認や手書き入力等に頼っているため，個人差や見落としなどが生じる問題がある．さらに，検査や確認作業に掛かるコストの大部分は人件費であり，人手不足対策としても，自動機の導入による生産効率の向上・省人化が期待されている．そこで，本研究では，生産現場における確認・監視作業に画像解析技術を適用することで，製造工程の状態の把握や管理を自動化する技術開発に取り組む．

昨年度は，機械学習の一種である Support Vector Machine(SVM) を用いた画像解析により，旋盤加工の工程進捗を 4 工程に細分化して把握することを試み，89.5%という比較的高い正解率で識別が可能であることを確認した [1)]．だたし，学習データに含まれていない未知の製品形状の分類精度（汎化性能）に関する検証が不十分であり，多様な製品が製造されている実際の生産工程への適応に関して懸念が残った．

そこで，本年度は，複数の製品品番の工程撮影画像で構成される，汎化性能評価用のデータセットを用意し，SVM の分類精度に関する追加検証を行った．また，一般的な機械学習モデルよりも画像分類精度が高いとされる深層学習モデルを，旋盤加工の工程分類問題に適用できるよう転移学習し，汎化性能および実行処理速度に関する SVM との比較検証を実施した．さらに，深層学習モデルによる画像解析機能を組み込んだ専用エッジデバイスを製作し，IoT 技術と連携することで工程進捗を自動的に監視・計測するシステムの開発を行った．

2. 深層学習

2.1 深層畳み込みニューラルネットワーク

近年，画像認識の分野では，畳み込みニューラルネットワーク (Convolution Neural Networks) を深層化することにより画像識別精度を飛躍的に向上させた深層畳み込みニューラルネットワーク (Deep Convolution Neural Networks，以下 DCNN) の研究が盛んに行われており，従来の手法を大きく上回る識別性能を達成する事例が数多く報告されている．

DCNN は，対象に適応した特徴量抽出フィルタや各ユニットの結合重みが，学習によって深層化されたネットワーク内に構築されるため，手作業で特徴量を設計することなく，特徴量抽出器および分類器として機能するという特徴がある．ただし，数億にも及ぶパラメータを持つ DCNN を学習させて，高度な識別性能を発揮する深層学習モデルを構築するためには，数百万という規模のサンプル画像を用い，膨大な時間を掛けて学習する必要がある [2)]．

しかしながら，ものづくりの生産現場への画像解析の適用を考慮した場合，多様な撮影条件における画像を実際の生産現場で収集することは容易ではなく，個々の案件に対して個別に巨大な学習データセットを作成し，新規にモデルを構築することは現実的に困難である．そこで，本研究では，学習済みモデルを再利用することで効率的にトレーニングを行う転

移学習を旋盤加工の工程分類問題に適用し，少量の訓練データ（数百〜数千個）で高い分類精度を達成する画像解析モデルの構築方法について検討した．

2.2 転移学習

転移学習 (transfer learning) とは，ある元となるドメイン（source domain）で特定の問題を解くために獲得した学習結果を，別の関連したドメイン (target dmain) における問題に適用する学習手法である[3]．転移学習の概略を図 1 に示す．DCNN における転移学習では，既存の学習済みモデルの最終出力層を，転用先のドメインに対応した出力層に付け替えて利用する．付け替えた出力層およびその手前の幾つかの層の結合パラメータを手元にある少量のデータで再学習させる．元ドメインから転用した学習結果を，関連したドメインにおける知識や情報を用いて補填することで，通常よりも少ないデータ量でも効率よく学習を行えるとされる[4]．

現在，TensorFlow[5] や PyTorch[6] など，一般に公開されている深層学習用のワークフレームが複数存在するが，今回の転移学習では，学習環境やデータセット構築の利便性などを考慮し fastai[7] を用いた．fastai は PyTorch をベースとし，より高水準な API（Application Programming Interface）を提供するライブラリの一つであり，最先端の研究で得られたモデルや手法を手軽に利用できる環境が整備されている．

また，転移学習のベースとなる既存の学習済みの深層学習モデルとして ResNet34 を用いた．ResNet は画像認識に特化した DCNN の一つで，2015 年に行われた大規模な画像認識コンテスト (ILSVRC 2015: ImageNet Large ScaleVisual Recognition Challenge 2015) で優勝をおさめており，加工工程の自動監視においても高い認識精度が期待できる．

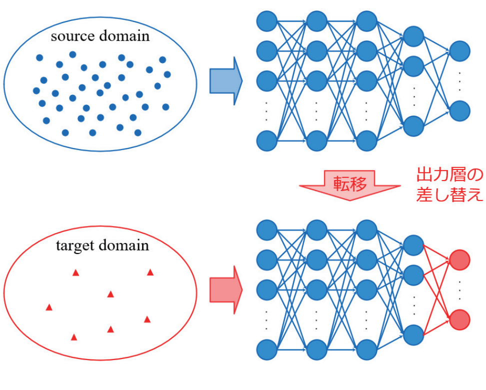


図 1　転移学習の概要

3. データセットの構築

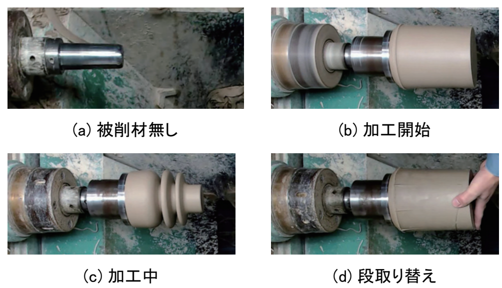
昨年度の研究において，SVM を用いることで特定の製品形状の工程分類において比較的高い分類精度が得られたが，学習データに含まれていない，未知の製品に対する工程分類に関して，正解率が悪化することが懸念された．このような未学習のテストデータに対する識別能力（汎化性能）について検証するために，複数の異なる製品品番の工程画像を含むデータセットを構築し，SVM および転移学習した DCNN の識別性能に関する評価を行う．

画像解析モデルの汎化性能を検証するために用意したデータセットの概要を表 1 に示す．学習用データとして 2 つの製品品番（a，b）の画像合計 1040 枚，評価用画像として 3 つの製品品番（c,d,e）の画像合計 1680 枚を，AR マーカを活用した Data Augmentation[1] により動画像から抽出した．また，旋盤加工工程の識別を多クラス分離問題として設計するために，図 2 に示すような 4 つの作業工程クラスを定義し，学習，評価においてクラス間の差が出ないよう，各製品品番のクラス毎の画像枚数は統一させた．

データセットに含まれる製品品番のうち，学習用データセットに含まれる品番「a」と，評価用データセットに含まれる品番「c」は類似する形状を有しているが，それ以外の製品は大きく異なる外観となっている．これらのデータセットを用い，評価用データの分類精度を各製品品番毎に評価することで，未知のテストデータ（今回のデータセットにおいては「d」「e」が該当）に対する汎化性能を検証する．

表 1　データセットとデータ拡張処理の内容

	学習用データセット		評価用データセット		
製品品番	a	b	c	d	e
製品形状	「c」と類似	類似品なし	「a」と類似	類似品なし	類似品なし
画像枚数	600枚（150×4種）	440枚（110×4種）	480枚（120×4種）	600枚（150×4種）	600枚（150×4種）



(a) 被削材無し　(b) 加工開始　(c) 加工中　(d) 段取り替え

図 2　作業工程（4 クラス）

4. モデルの学習および評価

4.1 最適な学習率の探索

深層学習モデルを訓練する際，学習率を適切に設定することが重要である．学習率が大きすぎると，最適化アルゴリズムが発散し，学習が全く進まない恐れがある．また，学習率が小さすぎると，モデルの訓練に掛かる時間が長くなるだけではなく，限られたデータで繰り返し学習を行う機会が増えることによって，訓練データに対して過学習となる可能性が高くなってしまう．このように，学習率は深層学習において重要でありながら，適切に調整することが難しいハイパーパラメータの一つであったが，2015 年に学習率ファインダ (learning rate finder) という手法が提案[8–10]され，比較的容易に最適な学習率を探索することが可能となった．

学習率ファインダでは，学習率の変化に伴う誤差の推移を小規模な実験単位で調査し，誤差の増減の傾向から最適な学習率を決定する．fastai[7]には学習率ファインダが機能の一部として実装されており，学習率に関する調査および結果の可視化が簡便に実施できる．今回の学習用データセットに対して行った学習率ファインダの実行結果を図 3 に示す．初期が学習率 10^{-7} から開始し，学習率が上昇するにつれて徐々に減少する傾向が見て取れる．特に 10^{-3} から 10^{-1} に掛けて大きく落ち込み，その後急激に増大している．最終的な学習率の決定方法に厳密な決まりはないが，誤差が最小となった地点の少し手前が良いとされているため，今回はその指標に従い 10^{-2} を学習率として採用することにした．

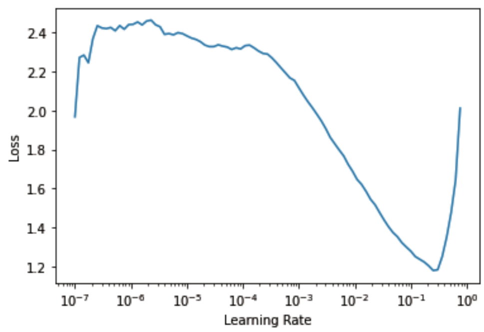


図 3　fastai による学習率ファインダの実行結果

4.2 転移学習の実施と性能評価

学習率を 10^{-2} とし，製品品番「a」と「b」の合計 1040 枚の学習用データセットを用いて DCNN(ResNet34) の転移学習を行った．GPU (NVIDIA GeForceTitanV) で増強したデスクトップ PC で実施したところ，転移学習によるトレーニングは約 120 秒という短時間で完了した．

転移学習済みの DCNN モデルで評価用データセットの分類精度を検証した結果（多クラス混同行列と正解率）を表 2 に示す．また，比較のために，同じデータセットを用いて，昨年度の研究で実施した SVM による手法[1]を適用した分類結果を表 3 に示す．DCNN の分類精度は総じて高く，全ての製品品番において 9 割以上の正解率が得られている．

表 2　DCNN の分類精度検証結果（入力画像サイズ 128×128 画素）

製品品番:c　正解率:0.998

		予測 無し	開始	加工中	段取替
真値	無し	120	0	0	0
	開始	0	120	0	0
	加工中	0	0	120	0
	段取替	0	0	1	119

製品品番:d　正解率:0.965

		予測 無し	開始	加工中	段取替
真値	無し	148	0	1	1
	開始	0	150	0	0
	加工中	2	5	139	4
	段取替	0	6	2	142

製品品番:e　正解率:0.922

		予測 無し	開始	加工中	段取替
真値	無し	148	6	3	0
	開始	0	150	0	0
	加工中	0	0	150	0
	段取替	1	8	29	112

表 3　SVM の分類精度検証結果（入力画像サイズ 300×150 画素）

製品品番:c　正解率:0.895

		予測 無し	開始	加工中	段取替
真値	無し	120	0	0	0
	開始	0	94	0	26
	加工中	1	2	111	6
	段取替	0	2	13	105

製品品番:d　正解率:0.660

		予測 無し	開始	加工中	段取替
真値	無し	149	0	1	0
	開始	0	54	7	89
	加工中	21	1	73	55
	段取替	8	19	3	120

製品品番:e　正解率:0.637

		予測 無し	開始	加工中	段取替
真値	無し	150	0	0	0
	開始	0	84	33	33
	加工中	0	79	36	35
	段取替	17	14	7	112

一方，SVMの正解率に関しては，学習用データセットに類似する製品（品番「a」）が含まれている品番「c」の識別精度は89.5%と比較的高いが，品番「d」「e」は60%台にまで急激に低下しており，学習済みのデータから乖離するテストデータに対して識別性能が著しく低下する傾向が見られた．

4.3 実行処理速度の比較

生産現場への画像処理技術の適用を考慮した場合，小型エッジデバイス等で処理できる実用的な解析手法であることが望ましい．そこで，学習済みのDCNNとSVMの実行処理速度に関する比較検証を行った．

検証用のエッジデバイスとして　Raspberry Pi 4 Model B (4GB) を，比較用の解析デバイスとして前述のGPUで増強したデスクトップPCを用いて，画像処理を10回実施した際の処理時間の平均値と標準偏差を表4に示す．それぞれの画像解析手法によって適切な入力画像サイズが異なるため厳密な比較とは言えないが，概ね同レベルの画素サイズにおける処理速度結果となっている．

デスクトップPCの場合，DCNNによる画像解析時間は平均66.2[ms]であり，動画をリアルタイムに解析できる程度の高い処理速度が得られている．一方，Raspberry Piの場合，SVMの解析時間は平均59.1[ms]と短いが，DCNNは平均1124[ms]を要している．処理能力の低いエッジデバイスなどでDCNNを使用する際は，定期的に撮影した静止画像を処理するなど，連続的な負荷を避けるための運用上の工夫が必要になると考えられる．

表4　処理速度の比較検証

解析手法	DCNN（転移学習モデル）		SVM（機械学習）	
画像サイズ	128×128		300×150	
デバイス	PC(GPU)	RasPi	PC(GPU)	RasPi
平均 [ms]	66.2	1124.9	4.8	59.1
標準偏差 [ms]	3.5	168.8	0.4	17.1

表5　構成機器

使用機器	仕様	
エッジデバイス	型番	Shuttle DH4100
	CPU	Intel Celeron G5905
	メモリ	DDR4 8GB
	HD	SSD 500GB
	OS	Ubuntu 20.04 LTS
Webカメラ	型番	Logicool c920
	解像度	1920 × 1080
無線モジュール	型番	Digi XBee ZB S2C

5. 工程解析システムの開発

5.1 画像解析ユニット

旋盤加工の工程分類に関しては秒単位での作業工程遷移が判別できれば十分であるため，速度よりも分類精度を重視し，転移学習によるDCNNモデルを解析手法として採用することにした．

実際に生産現場に設置するために開発した画像解析ユニットを図4に，構成機器を表5に示す．RaspberryPi4は基本的にSDカードで稼働する設計となっているため，長期間の運用における耐久性に不安がある．そこで，システムドライブがSSDで，さらに処理能力がRaspberryPi4よりも優れているShuttle社製の小型PC(DH4100)をエッジデバイスに採用し，オペレーションシステムはUbuntu(Linux)とした．また，解析結果をデータベースに送信する通信手段として近距離無線通信(ZigBee)を採用した．さらに，全ての外部機器をUSBで簡便に接続できる構成とすることで，生産現場の任意の場所への設置･運用が容易に実施できるよう配慮した．

Webカメラによる画像撮影，DCNNによる工程分類，解析結果の送信処理などを実行するプログラム（スクリプト）は，表6に示す開発環境においてPython言語で作成した．また，こららのすべての処理を自動運用するために，Linuxのsystemd(システム管理デーモン)にサービス起動スクリプトとして登録し，工程監視の実施曜日や時間帯を設定したうえで，工程分類結果を2秒間隔で自製のオンプレミスサーバーに自動送信するよう設定した．

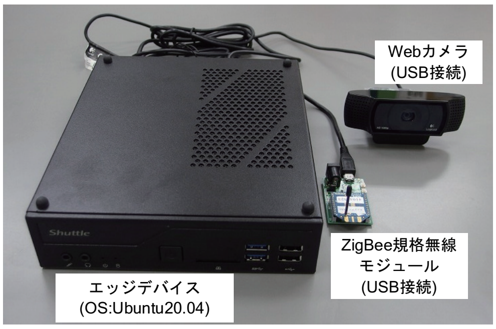


図4　画像解析ユニット

表6　開発環境

プログラミング言語	Python 3.8.6
ライブラリ	opencv-contrib 4.5.1
	digi-xbee 1.3.0
	torch 1.7.1
	torchvision 0.8.2
	fastai 2.1.9

5.2 生産現場への設置・運用

オンプレミスサーバーのデータベースに蓄積した工程遷移の解析結果を可視化したものを図5に示す．サーバ上に構築したWebアプリケーション[11]を介してデータベースにアクセスすることで，同一ネットワーク内にある任意のPC，スマートフォン，タブレット等のWebブラウザから状態の確認や結果のダウンロードができるようになっている．

画像解析システムを実際の加工装置に設置した様子を図6に示す．生産活動の妨げにならないよう，撮影用のWebカメラは作業領域の上方に設置した．工場等の生産現場では，カメラ等の機器を設置する際の制約により，位置や固定方法が制限されることが多い．そのため，理想的な撮影範囲が得られなかったり，カメラの位置ずれにより撮影範囲が変わるといったことが問題となるが，本システムにおいてはARマーカを基準として監視領域を特定[1]しているため，カメラの初期設置が容易であり，長期運用における監視範囲の安定性も比較的高い．

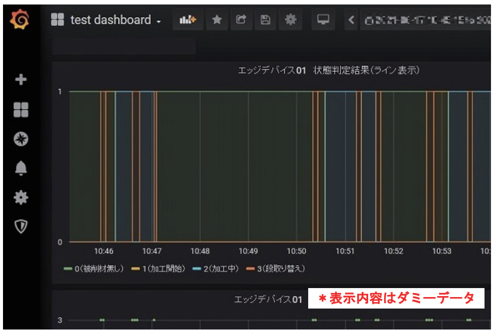


図5　工程解析結果の可視化画面

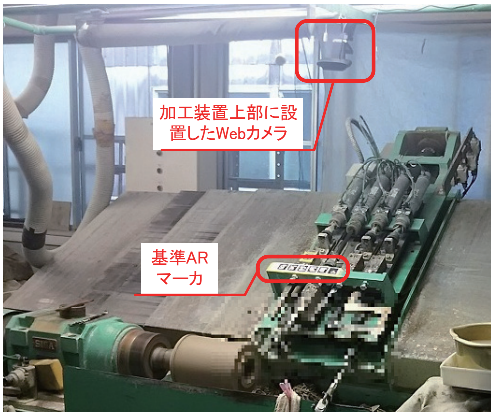


図6　生産現場への設置状況

6. おわりに

本報告では，未知の製品形状に対する作業工程分類の精度向上のため，深層学習の一種であるDCNN（ResNet32）モデルを，学習率ファインダにより最適化した学習率で転移学習させた．そして，機械学習手法のSVMとの性能比較において，未学習の製品形状に対するSVMの正解率が63.7～66.0%まで低下したのに対し，DCNNを用いた解析では92.2～96.5%と高い正解率を維持しており，汎化性能においても大幅に優れていることを確認した．一方，RaspberryPi上での実行速度の比較検証においては，SVMが平均59[ms]であるのに対し，DCNNは平均1124[ms]と処理時間を要することが分かり，動画像解析などの高負荷なタスクにDCNNを適用する場合，低スペックなエッジデバイスでは対応が困難であることが明らかとなった．実務運用においては，精度・処理速度の優先順位を鑑み，用途や目的によってこれらの解析手法を使い分ける必要があると思われる．

本検討においては，定期的な静止画像解析によって工程遷移の計測が可能であったため，処理速度よりも解析精度を重視し，オリジナルのデータセットで転移学習したDCNNモデルを採用した．また，撮影や解析処理を任意のスケジュールで実施する自律型の工程解析用エッジデバイスを，LinuxOSをベースにPython言語を用いて開発した．さらに，近距離無線通信によって自製のオンプレミスサーバと連携させることで，データの収集・蓄積・可視化までをオートメーション化し，実際の生産現場における工程進捗を監視・計測するシステムを構築することができた．

昨今，AI技術の発展により，様々な画像解析手法が一般に公開され比較的容易に利用できる環境が整いつつある．しかし，それらの技術を実務に適用し継続的に運用するためには，核となる先端技術に加え，IoTなどの周辺技術との連携が不可欠と言える．

本検討において，県内企業と共同して課題解決に取り組み，画像解析，エッジデバイス開発，通信技術による連携までを一貫して実地検証したことで，生産現場に先端のAI技術を適用するためのノウハウを蓄積することができた．深層学習を実行する際の処理速度などに課題が残るが，今後は，本システムをベースとしてさらに改善を図り，県内企業への画像解析技術およびIoT技術の普及・導入に向けた応用展開について検討していきたい．

参考文献

1) 中野太郎，大坪昭文，福島章吾，画像解析を利用した生産現場の状態監視技術に関する研究（第２報），令和２年度佐賀県工業技術センター研究報告書，2020，No.29，p.53-57.
2) 岡谷貴之，画像認識のための深層学習の研究動向: 畳込みニューラルネットワークとその利用法の発展，人工知能，2016，31.2，p.169-179
3) 神嶌敏弘， 転移学習，人工知能，2010，25.4，p.572-580.
4) 伊藤雅仁，et al.，転移学習と生成ネットワークの試行事例，Journal of Computer Chemistry, Japan，2019，18.3，p.132-135.
5) TensorFlow，https://www.tensorflow.org/
6) PyTorch，https://pytorch.org/
7) fastai，https://docs.fast.ai/
8) Jeremy Howard，Sylvain Gugger，PyTorch と fastai ではじめるディープラーニング，オライリー・ジャパン，2021，p.193-195.
9) Smith, Leslie N. No more pesky learning rate guessing games. CoRR, abs/1506.01186, 2015, 5.
10) Smith, Leslie N. Cyclical learning rates for training neural networks. In: 2017 IEEE winter conference on applications of computer vision (WACV). IEEE, 2017. p. 464-472.
11) Grafana，https://grafana.com/