

AI 技術を活用した検査工程の省力化・効率化（第1報）

— 深層学習を用いた欠陥画像分類モデルの改良 —

渡辺博己*、生駒晃大*、松原早苗*、伊藤司†、安田敏†、山下泰司†

A study on artificial intelligence for labor savings and efficiency improvements of inspection process (I)

— Improvements of defect classification networks using deep learning —

WATANABE Hiroki*, IKOMA Akihiro*, MATSUBARA Sanae*,

ITO Tsukasa†, YASUDA Satoshi† and YAMASHITA Hiroshi†

本研究では、省力化・効率化を図ることによる生産性向上の支援を目的として、品質保証に絡む重要な作業である検査作業について、AI 技術を利用した画像検査技術を開発している。本稿では、欠陥画像に対する認識性能の向上を目指して、昨年度開発した欠陥画像分類モデルについて、4 つの改良を試みた。実験では、各改良において上位の評価結果となったモデルを順に改良し、最も認識性能の高いモデルを選定した。また、従来モデルと比較する実験を行い、欠陥・正常分類で平均再現率 1.19%、平均適合率 1.28% の認識性能の向上を確認した。

1. はじめに

近年、人口減少による人手不足が懸念されている。1990 年代前半には 8,700 万人だった国内生産年齢人口は、2056 年には 5,000 万人を割り込むことが予想され¹⁾、製造業においては労働力確保が課題となっている。特に人手を要する検査作業は、品質保証に絡む重要な作業であり、検査員の確保が重要な課題となっている。

人による目視検査は、曖昧な違いがあっても経験により検査が可能であったり、検査項目や条件が変わったとしても柔軟に対応が可能であったりするなど、導入のしやすさから広く普及してきた。しかし、労働力確保とともに、熟練検査員のノウハウ継承や、人は体調や環境により判断基準にバラツキが生じるため、安定した検査品質の確保が課題であった。

一方、機械による画像検査は、客観的な評価基準による一定の検査品質が保証されるだけでなく、導入により人手を削減でき、検査速度を含めて経済的な効率化が期待できる。しかし、従来技術では、数値で定義できないものは検査が困難であり、微妙な光の加減で誤判定が起きるなど、自動化できない対象があり、できたとしても精度が低い場合があるなど、解決すべき課題が幾つか存在した。

こうした中、人の持つ柔軟性と機械の持つ効率性を組み合わせ、人工知能（以下、AI）による外観検査技術が注目されている。しかし、データが少量であったり、偏りがあったりすると、性能の確保が困難となるだけでなく、多様な検査項目に応じたアルゴリズムの選定が必要となるなど、技術の導入は容易ではない。

そこで本研究では、AI 技術を適用した画像検査技術を開発し、検査工程の省力化・効率化を図ることにより、生産性向上を支援する。本稿では、欠陥画像に対する認

識性能の向上を目的として、昨年度開発した畳み込みニューラルネットワーク²⁾（以下、CNN）による欠陥画像分類モデル³⁾（以下、モデル）を改良し、欠陥画像に対する分類実験を行ったので、その結果について報告する。

2. 欠陥画像分類技術の概要

2. 1 欠陥画像分類モデル

モデルは、既存検査装置の検出・分析機能に課題があり、AI 技術の利用による装置の改善を要望する企業ニーズにより開発した、CNN に基づいた深層学習モデルである。図 1 に開発したモデル（以下、従来モデル）の構成を示す。従来モデルは、2 個の入力（Input）層と 3 個の出力（Output）層を持ち、13 個の畳み込み（Convolution）層と 1 個の全結合（FullyConnected）層を持つ 2 つのネットワークを統合した構成となっている。各ネットワークで処理されたデータは、最後のプーリング（MaxPooling）層を通過後に分岐され、各ネットワークから分岐された 2 つのデータは、統合（Concatenate）層で連結された後、各 1 個の畳み込み層と全結合層を経て、画像種別に対する確率として出力される。欠陥画像を分類する際には、統合後の出力をモデルの出力結果として使用し、各ネットワークの出力結果は学習時のネットワークの検証用としてのみ使用する。

2. 2 欠陥画像データセット

欠陥画像データは、企業の実際の検査対象物を既存検査装置で検査することにより取得されたデータである。検査対象物は円筒部品であり、検査では、部品を回転させた状態で、回転機構と同期の取れたラインスキャンカメラにより、部品全周の画像が取得される。この時、部品は 4 回転し、各回転において、異なる方向から照明光が照射された状態で画像が取得される。取得された画像については、欠陥の有無が分析され、欠陥があれば、欠陥部分が画像中心となるよう、256×256 の解像度で、

* 情報技術部

† 株式会社前田精工

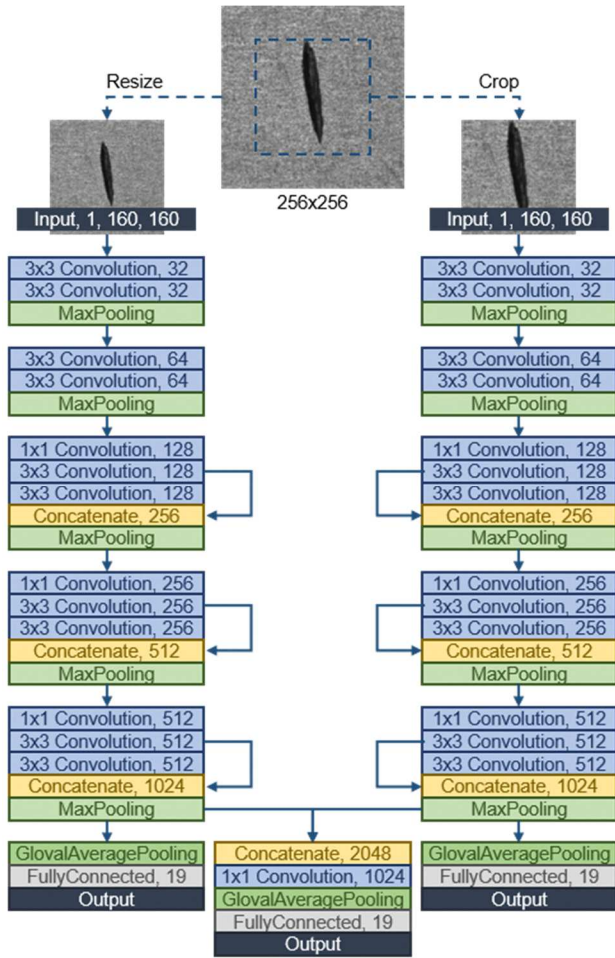


図1 欠陥画像分類モデルの構成

ビットマップ形式の画像ファイルとして保存される。欠陥の種類によっては、全ての照明方向において検出される欠陥もあれば、そうでない欠陥もある。また、欠陥の分布・大きさによっては、複数の画像で同一の欠陥が記録されることもある。そのため、欠陥の状態により取得される画像数が異なり、1つの検査対象物について、複数の欠陥画像ファイルが保存される。図2に保存された欠陥15種類、正常4種類の画像例を、表1に種別毎のデータ数を示す。なお、データセットは昨年度と比較すると、欠陥C、H、O、正常(キズ)の4種類が増加し、分類の難易度が高くなっている。

3. 欠陥画像分類モデルの改良

3.1 畳み込みブロックの構成

モデルを構成する1つのネットワークに着目すると、ネットワークはプーリング層を区切りとして、 3×3 の畳み込み層のみから構成されるブロックと、 1×1 と 3×3 の畳み込み層、及びスキップコネクションによる統合層から構成されるブロックに分けることができる。前者を畳み込みブロックCB1、後者を畳み込みブロックCB2とすると、ネットワークは $(CB1_{ij}, CB2_{kl})$ と一般化することができる。ただし、 i, k はブロックの繰り返

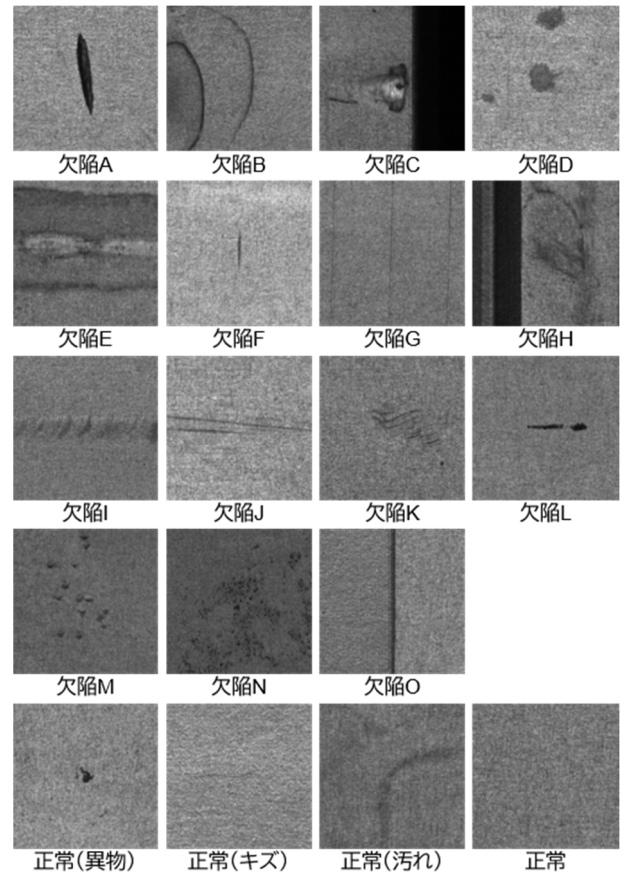


図2 欠陥画像データセットの画像例

表1 欠陥画像データセットのデータ数

種別	データ数	種別	データ数	種別	データ数	種別	データ数
欠陥A	794	欠陥B	1,076	欠陥C	325	欠陥D	795
欠陥E	939	欠陥F	797	欠陥G	923	欠陥H	290
欠陥I	1,461	欠陥J	902	欠陥K	978	欠陥L	728
欠陥M	743	欠陥N	883	欠陥O	362		
正常(異物)	916	正常(キズ)	649	正常(汚れ)	994	正常	762

し数、 j, l はブロック内の畳み込み層数を表しており、 $j \geq 1, l \geq 3$ とする。図3に $j=1, 2, l=3, 4$ の時の各ブロックの構成例を示す。CB1は、単純に 3×3 の畳み込み層を重ねる構成であるが、CB2は、 3×3 の畳み込み層を重ねると同時に、スキップコネクションにより2層目以降の畳み込み層を統合する構成としている。一般化することにより、様々な組み合わせでネットワークを構成することが容易となるが、本稿では $4 \leq i+k \leq 5$ ($i \leq k$)、 $1 \leq j \leq 2, 3 \leq l \leq 4$ として畳み込みブロックを構成し、モデルの認識性能を検証した。なお、従来モデルのネットワークは $(CB1_{2,2}, CB2_{3,3})$ となる。

3.2 Squeeze-and-Excitation ブロックの追加

Squeeze-and-Excitation ブロック⁴⁾ (以下、SEブロック)は、大規模画像認識のコンテストである ImageNet Large Scale Visual Recognition Challenge 2017の画像分類タスクで1位を記録したネットワークを構成する手法で、チャンネル毎の特徴マップを適応的に重み付けするブロックである。SEブロックは、ネットワーク中の1つのコ

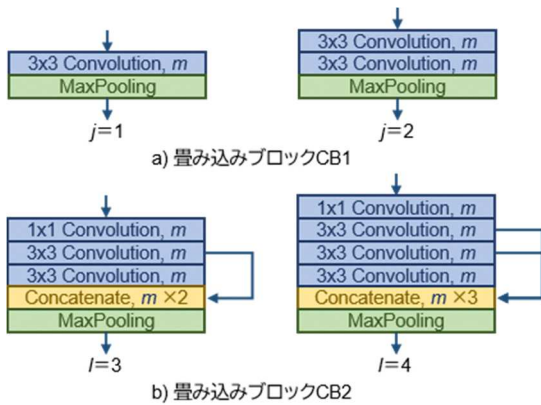


図3 畳み込みブロックの構成例 (m はチャンネル数)

ンポーネントとして振る舞うことが可能であるため、様々なネットワークに適用することが可能という特徴を持つ。本稿では、図4に示すように、各畳み込み層の後にSEブロックを組み込み、SEブロックの有無によるモデルの認識性能を検証した。

3.3 Dropoutの追加

Dropout⁵⁾は、ネットワークの学習時に一定割合のノードを不活性化させながら学習を行うことで、未知のデータに対して汎化性が低い状態である過学習を緩和する効果があると同時に、複数の機械学習結果を利用して判定を行うことで学習性能を上げるアンサンブル学習に近似した効果がある手法である。本稿では、図5に示すように、各プーリング層の後にDropoutを組み込み、Dropoutの有無によるモデルの認識性能を検証した。

3.4 統合後の畳み込み層の構成

本モデルは、異なる画像をそれぞれのネットワークで畳み込んでいくため、各ネットワークにおいては異なる意味を持つ特徴マップが形成されていると考えられる。そのため、統合後の畳み込み層は不要であると考えられるが、仮に畳み込み処理に効果があるのであれば、さらに畳み込み層を繰り返した場合の効果についても検証する必要がある。そこで本稿では、図6に示すように、統合後の畳み込み層数を0から3(以下、L0~L3)の間で変化させ、モデルの認識性能を検証した。なお、従来モデルの畳み込み層数はL1である。

4. 実験と結果

4.1 改良効果の検証

前章の4つの改良に対する効果を検証するために、2.2節のデータセットからランダムにデータを抽出し、学習・分類を試行する実験を行った。モデルの評価にあたっては、正解であるデータのうち実際に正解したデータの割合を示す再現率(recall)と、正解として予測したデータのうち実際に正解したデータの割合を示す適合率(precision)を求め、再現率と適合率との調和平均であるF値(F-measure)により評価した。

実験は、各改良に対して5回試行し、集計した分類結

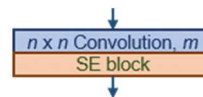


図4 SEブロックの組込例 (n はフィルタサイズ)

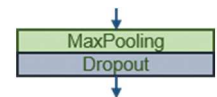


図5 Dropoutの組込例

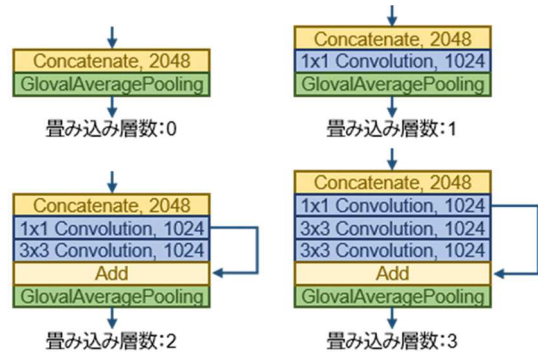


図6 統合後の畳み込み層の構成

果について、画像種別19カテゴリ、欠陥・正常2カテゴリに対するF値を、学習時に使用した検証データ、テスト用の評価データから求めた。また、実験は、3.1から3.4節の改良の順に行い、3.2節の改良については、3.1節の改良による上位5つのモデルを、3.3、3.4節の改良については、3.2、3.3節の改良による上位3つのモデルについて評価した。なお、モデルの優位性については、各F値の調和平均により判定した。各改良に関する実験結果を順に表2から5に示す。

畳み込みブロックの構成に関する実験については、(CB1_{1,2}, CB2_{3,4})、(CB1_{2,2}, CB2_{2,4})、(CB1_{1,1}, CB2_{3,4})、(CB1_{2,1}, CB2_{2,4})、(CB1_{2,2}, CB2_{3,4})の順に上位となり、CB2における畳み込み層数が4のモデルが有効であった(表2)。SEブロックの有無に関する実験については、SEブロックを使用したモデルが上位となり(表3)、SEブロックの有効性が明確となった。Dropoutについては、モデルによっては必ずしも効果があるとは言えないが(表4)、認識性能を向上させる傾向が見られた。統合後の畳み込み層の構成については、(CB1_{2,1}, CB2_{2,4})の2つのモデルでL3 > L0 > (L1 or L2)となったが、(CB1_{2,2}, CB2_{2,4})のモデルではL1が上位となった(表5)。そのため、統合後の畳み込み層数については、モデルの構成により、幾つに設定することが望ましいかを、その都度検証する必要がある。なお、これらの実験結果より、本稿では、表5において最も調和平均の高かったモデルを優位性の高いモデルとして選定した。

4.2 従来モデルとの比較

前節で選定したモデル(以下、選定モデル)について、従来モデルとの認識性能を比較する実験を行った。選定モデルと従来モデルとの構成の違いを表6に示す。

実験は、前節と同様に学習・分類を各モデルについて行い、10回の試行から検証データに対する比較結果を求めた。表7に画像種別に分類した場合、表8に欠陥・

表2 畳み込みブロックの構成に関する実験結果

畳み込み ブロック	検証データ		評価データ		調和 平均
	画像種別	欠陥・正常	画像種別	欠陥・正常	
(CB1 _{1,1} , CB2 _{3,3})	0.8966	0.9531	0.8825	0.9468	0.9187
(CB1 _{1,1} , CB2 _{3,4})	0.9088	0.9582	0.8919	0.9547	0.9275
(CB1 _{1,1} , CB2 _{4,3})	0.9000	0.9530	0.8919	0.9476	0.9223
(CB1 _{1,1} , CB2 _{4,4})	0.9017	0.9516	0.8902	0.9471	0.9219
(CB1 _{1,2} , CB2 _{3,3})	0.9015	0.9549	0.8870	0.9524	0.9230
(CB1 _{1,2} , CB2 _{3,4})	0.9108	0.9601	0.8937	0.9534	0.9287
(CB1 _{1,2} , CB2 _{4,3})	0.9043	0.9541	0.8862	0.9440	0.9213
(CB1 _{1,2} , CB2 _{4,4})	0.9046	0.9558	0.8872	0.9488	0.9232
(CB1 _{2,1} , CB2 _{2,3})	0.9029	0.9541	0.8788	0.9434	0.9188
(CB1 _{2,1} , CB2 _{2,4})	0.9091	0.9581	0.8876	0.9513	0.9256
(CB1 _{2,1} , CB2 _{3,3})	0.9008	0.9522	0.8829	0.9430	0.9188
(CB1 _{2,1} , CB2 _{3,4})	0.9030	0.9531	0.8813	0.9418	0.9189
(CB1 _{2,2} , CB2 _{2,3})	0.9023	0.9558	0.8853	0.9478	0.9218
(CB1 _{2,2} , CB2 _{2,4})	0.9098	0.9606	0.8957	0.9508	0.9284
(CB1 _{2,2} , CB2 _{3,3})	0.9010	0.9542	0.8923	0.9501	0.9235
(CB1 _{2,2} , CB2 _{3,4})	0.9065	0.9562	0.8926	0.9483	0.9251

表3 SEブロックの有無に関する実験結果

畳み込み ブロック	SE ブロック	検証データ		評価データ		調和 平均
		画像種別	欠陥・正常	画像種別	欠陥・正常	
(CB1 _{1,1} , CB2 _{3,4})	無	0.9039	0.9538	0.8738	0.9406	0.9169
	有	0.9123	0.9602	0.8831	0.9477	0.9248
(CB1 _{1,2} , CB2 _{3,4})	無	0.9044	0.9543	0.8793	0.9402	0.9186
	有	0.9140	0.9635	0.8907	0.9466	0.9278
(CB1 _{2,1} , CB2 _{2,4})	無	0.9050	0.9532	0.8744	0.9373	0.9165
	有	0.9153	0.9629	0.8827	0.9462	0.9257
(CB1 _{2,2} , CB2 _{2,4})	無	0.9070	0.9572	0.8817	0.9434	0.9214
	有	0.9170	0.9652	0.8898	0.9501	0.9296
(CB1 _{2,2} , CB2 _{3,4})	無	0.9048	0.9550	0.8700	0.9373	0.9156
	有	0.9100	0.9559	0.8847	0.9473	0.9236

表4 Dropoutの有無に関する実験結果

畳み込み ブロック	SE ブロック	Dropout	検証データ		評価データ		調和 平均
			画像種別	欠陥・正常	画像種別	欠陥・正常	
(CB1 _{1,2} , CB2 _{3,4})	有	無	0.9149	0.9650	0.8925	0.9519	0.9302
		有	0.9161	0.9592	0.8948	0.9458	0.9283
(CB1 _{2,1} , CB2 _{2,4})	有	無	0.9158	0.9638	0.8989	0.9531	0.9322
		有	0.9199	0.9642	0.9041	0.9530	0.9347
(CB1 _{2,2} , CB2 _{2,4})	有	無	0.9148	0.9632	0.9002	0.9518	0.9318
		有	0.9276	0.9689	0.9089	0.9553	0.9396

表5 統合後の畳み込み層の構成に関する実験結果

畳み込み ブロック	SE ブロック	Dropout	畳み込 み層数	検証データ		評価データ		調和 平均
				画像種別	欠陥・正常	画像種別	欠陥・正常	
(CB1 _{2,1} , CB2 _{2,4})	有	無	L0	0.9176	0.9641	0.8939	0.9539	0.9315
			L1	0.9157	0.9640	0.8925	0.9548	0.9308
			L2	0.9156	0.9650	0.8927	0.9516	0.9303
			L3	0.9150	0.9642	0.8953	0.9570	0.9320
			L0	0.9197	0.9645	0.8976	0.9554	0.9335
			L1	0.9191	0.9620	0.8930	0.9481	0.9298
(CB1 _{2,2} , CB2 _{2,4})	有	有	L2	0.9218	0.9659	0.8934	0.9499	0.9319
			L3	0.9203	0.9644	0.8975	0.9561	0.9338
			L0	0.9227	0.9641	0.9056	0.9570	0.9367
			L1	0.9275	0.9692	0.9075	0.9593	0.9402
			L2	0.9244	0.9644	0.8973	0.9510	0.9336
			L3	0.9226	0.9648	0.8966	0.9538	0.9337

正常に分類した場合の平均再現率、平均適合率を示す。

実験より、選定モデルでは従来モデルと比較して、画像種別、欠陥・正常別の順に、平均再現率が1.97%、1.19%、平均適合率が2.53%、1.28%向上した。

表6 選定モデルと従来モデルとの構成の違い

モデル	畳み込みブロック	SEブロック	Dropout	畳み込み層数
選定モデル	(CB1 _{2,2} , CB2 _{2,4})	有	有	L1
従来モデル	(CB1 _{2,2} , CB2 _{3,3})	無	無	L1

表7 画像種別に対する認識性能の比較結果

種別	学習 データ数	検証 データ数	選定モデル		従来モデル	
			平均再現率	平均適合率	平均再現率	平均適合率
欠陥A	600	150	0.8793	0.8627	0.8433	0.8496
欠陥B	800	200	0.9440	0.9787	0.9395	0.9376
欠陥C	200	50	0.9040	0.9113	0.9320	0.8457
欠陥D	600	150	0.9633	0.9438	0.9453	0.9160
欠陥E	700	175	0.9394	0.9821	0.9309	0.9605
欠陥F	600	150	0.9520	0.9389	0.9293	0.9312
欠陥G	700	175	0.9966	0.9892	0.9954	0.9886
欠陥H	200	50	0.9760	0.9859	0.9600	0.9339
欠陥I	1,000	250	0.9460	0.9556	0.9104	0.9539
欠陥J	700	175	0.8920	0.8168	0.8651	0.8053
欠陥K	800	200	0.8270	0.8926	0.8015	0.8464
欠陥L	500	125	0.9776	0.9393	0.9664	0.9243
欠陥M	600	150	0.8747	0.9143	0.8340	0.8730
欠陥N	700	175	0.9446	0.8974	0.9023	0.8876
欠陥O	200	50	1.0000	1.0000	0.9960	1.0000
正常(異物)	700	175	0.9017	0.9371	0.9017	0.8895
正常(キズ)	500	125	0.8072	0.8624	0.7544	0.8243
正常(汚れ)	800	200	0.9470	0.9284	0.8880	0.9193
正常	600	150	0.9580	0.8976	0.9600	0.8664
全体	11,500	2,875	0.9279	0.9281	0.9082	0.9028

表8 欠陥・正常に対する認識性能の比較結果

種別	学習 データ数	検証 データ数	選定モデル		従来モデル	
			平均再現率	平均適合率	平均再現率	平均適合率
欠陥	8,900	2,225	0.9863	0.9861	0.9803	0.9808
正常	2,600	650	0.9523	0.9530	0.9345	0.9327
全体	11,500	2,875	0.9693	0.9696	0.9574	0.9568

5. まとめ

本研究では、認識性能の向上を目的として、従来の欠陥画像分類モデルに対し、畳み込みブロックの構成、SEブロックやDropoutの追加、ネットワーク統合後の畳み込み層の構成について改良を行った。その結果、従来モデルと比較して、欠陥・正常に分類する認識性能が平均再現率で1.19%、平均適合率で1.28%向上した。

今後は、モデルの改良は元より、データセットの構成や学習方法等についても見直しを行い、さらなる認識性能の向上を目指して、深層学習に関する研究開発を実施する予定である。

【参考文献】

- 1) 国立社会保障・人口問題研究所, 人口問題研究資料第336号, 2017
- 2) Y. LeCun, et al., Proc. of the IEEE, vol.86, no.11, pp.2278-2324, 1998
- 3) 渡辺ら, 岐阜県情報技術研究所研究報告第20号, pp.1-6, 2019
- 4) J. Hu, et al., arXiv:1709.01507, 2017
- 5) N. Srivastava, et al., The Journal of Machine Learning Research, 15(1), pp.1929-1958, 2014