

遺伝的プログラミングを用いた階層的な特徴構築による 画像分類

菅沼 雅徳^{1,2,a)} 土屋 大樹¹ 白川 真一¹ 長尾 智晴¹

概要: 本論文では、画像分類のための階層的な特徴構築手法を提案する。提案手法では入力画像に対して、(1) 既存の画像処理フィルタの組み合わせ、(2) 遺伝的プログラミングで構築したフィルタ、の2層の画像変換処理を行い、変換後の画像の各画素値を分類のための特徴量として扱う。1層目の画像処理フィルタの組み合わせと2層目でのフィルタは遺伝的プログラミングを用いて段階的に構築する。本論文では、カプセル内視鏡から撮影された小腸画像における異常画像と正常画像の分類問題に提案手法を適用し、従来手法との比較と性能の検証を行う。

キーワード: 遺伝的プログラミング, 画像分類, 特徴構築, 多層構造

Image Classification Based on Hierarchical Feature Construction Using Genetic Programming

MASANORI SUGANUMA^{1,2,a)} DAIKI TSUCHIYA¹ SHINICHI SHIRAKAWA¹ TOMOHARU NAGAO¹

Abstract: In this paper, we propose a hierarchical feature construction method for image classification. Our method has two feature construction stages: (1) feature construction by a combination of existing image processing filters, and (2) feature construction by evolved filters. The combination of image filters and evolved filters are constructed step-by-step using genetic programming. We verify the classification performance of the proposed method on the small bowel images taken from a capsule endoscope.

Keywords: genetic programming, image classification, feature construction, multi-layer architecture

1. はじめに

画像分類や物体認識などの技術は幅広い分野で必要とされているため、盛んに研究が行われている。特に、local binary pattern (LBP) [1], histogram of oriented gradients (HOG) [2], scale-invariant feature transform (SIFT) [3] や Gabor bank [4] などの人手によって考案された特徴量を用いた分類手法がこれまでに高い分類精度を示してきた。

しかし、これらの特徴量は特定の種類の画像分類問題では有効に働くが、ほかの種類の画像分類問題では有効に働かない可能性が考えられる。例えば、LBP 特徴はテクスチャ画像分類では高い分類精度を示すが、シーン画像分類では有効に働かないと考えられる。対象となる分類問題は目的によって多種多様であり、それら対象となる分類問題に応じて新たに特徴量を設計することは膨大な労力を必要とする。そこで計算機による特徴量の自動構築手法が求められている。

近年では、画像分類のための特徴構築に深層学習 [5] が数多く適用されており、多くの分類問題で成功を収めている。深層学習では、複数の演算処理が段階的に続く多層構造によって分類に有効な特徴量を自動構築している。多層

¹ 横浜国立大学大学院環境情報学府
Graduate School of Environment and Information Sciences,
Yokohama National University, Yokohama, Kanagawa 240-
8501, Japan

² 日本学術振興会特別研究員 DC
JSPS Research Fellow

a) suganuma-masanori-hf@ynu.jp

構造による特徴構築には, greedy layer-wise training[6] と呼ばれる方法が提案されており, 単層構造で構築された特徴量より分類に有効な特徴量が構築されることが報告されている [7].

また, 遺伝的プログラミング (Genetic Programming; GP) [8] を用いて特徴量を自動構築する手法も数多く提案されており, 画像分類問題において成功を取めている. GP を用いた手法では, (1) 事前に定義した特徴量の組み合わせから新たに特徴量を構築する手法, (2) 生の画素値の組み合わせから新たに特徴量を構築する手法が主に挙げられる.

前者の手法では, GP の入力画像全体の標準偏差や画像中心部分の平均画素値などの事前に定義した統計特徴量で構成され, それら統計特徴量の演算の組み合わせによって分類に有効な特徴量を構築している [9][10]. しかし, 事前に定義した統計特徴量の組み合わせで分類することが難しい問題では新しい統計特徴量を追加する必要がある. このため, 後者の手法のように入力画像中の画素値から新たに特徴量を構築する手法の実現が望まれる. 本論文では, 後者の画素値から特徴量を構築する新しい手法を提案する.

後者の手法では, GP の入力は入力画像中の画素値または画像全体で構成され, それらに対する演算の組み合わせによって特徴量を構築する. Al-Sahaf らは, 画像内から特徴量を抽出する領域と, その領域内の画素値から算出した統計特徴量の組み合わせによって特徴量を構築する構造を提案しており, GP を用いて構造の構築を行っている [11]. Kowaliw らは, Cartesian Genetic Programming (CGP) [12][13] を用いて入力画像に対して画像変換を行い, 変換後の画像から複数の統計特徴量を算出することで分類のための特徴量を構築している [14]. 文献 [15] では入力画像からの特徴構築と分類を同時に最適化する Genetic Image Network for Image Classification (GIN-IC) を提案している. GIN-IC では, 既存の画像処理フィルタによる画像変換後の画像から統計特徴量を算出し, 算出した統計特徴量の演算結果を用いて分類まで行う構造を最適化する. 平野らも同様に画像処理フィルタによる画像変換に基づく特徴構築手法を提案している [16]. これら画像変換に基づく特徴構築手法は, 画像分類問題において有効であることが示されている.

一方, これら GP による特徴構築手法の多くは一段階の最適化によって特徴量を構築している. 近年の深層学習の研究成果で示されているように, 特徴量を構築する構造の多層化を行うことで単一の構造よりも分類に有効な特徴量を構築できることが GP の手法においても期待できる. 実際, Agapitos らは特徴量の最適化を行う構造の多層化を行うことで, 一段階の最適化よりも分類に有効な特徴量を構築できることを示している [17]. Agapitos らの手法は, (1) ランダムフィルタによる画像変換, (2) GP による画

像変換, (3) 平均プーリング, (4) ロジスティック回帰による分類, の 4 層構造を用いている. しかし, Agapitos らの手法では処理の始めにランダムフィルタによる画像変換を行っており, ランダムフィルタによる変換では変換後の画像から効率的に分類に有効な特徴量が抽出されることは難しいと考えられる.

そこで本論文では, (1) 既存の画像処理フィルタの組み合わせによる画像変換, (2) GP で構築したフィルタ処理による画像変換, の 2 種類の処理を二段階の進化計算によって最適化することで特徴構築を実現する手法を提案する. ランダムフィルタではなく, 既存の画像処理フィルタの組み合わせによる画像変換処理を進化過程に組み込むことでより分類に有効な特徴量の構築が可能になると考えられる. 本論文では, データセット内の多様性が高いと考えられるカプセル内視鏡から撮影された小腸画像における異常画像と正常画像の分類問題に提案手法を適用し, 性能の検証を行う.

2. 画像分類のための階層的な特徴構築

2.1 処理の流れ

まず, 特徴構築の第一段階では, 既存の画像処理フィルタの組み合わせによって入力画像を変換し, 変換後の画像に対してプーリング処理を行う. そして, プーリング処理後の各画素値を特徴量として分類器に入力し, 分類を行う. このときの検証画像セットに対する分類精度が高くなるように, 画像処理フィルタの組み合わせを GP によって構築する. 提案手法における特徴構築の第一段階の処理の流れは次に示す通りである.

- (1) 入力画像に対して既存の画像処理フィルタを用いて画像変換 (フィルタリング層)
- (2) プーリング処理 (プーリング層)
- (3) 分類 (分類層)

次に, 特徴構築の第二段階では, 第一段階で得られた変換画像をさらに GP で構築したフィルタ処理によって画像変換を行う. そして第一段階と同様に, 変換後の画像に対してプーリング処理を行い, プーリング処理後の各画素値を特徴量として分類器に入力する. このときの検証画像セットに対する分類精度が高くなるように, フィルタを GP によって構築する. 提案手法における特徴構築の第二段階の処理の流れは次に示す通りである.

- (1) フィルタリング層で変換された画像に対して GP で構築したフィルタを用いて画像変換 (変換層)
- (2) プーリング処理 (プーリング層)
- (3) 分類 (分類層)

2.2 フィルタリング層 (F)

フィルタリング層では, 入力画像に対して既存の画像処理フィルタを適用することで, 分類に有効な特徴構築

表 1 フィルタリング層で使用する画像処理フィルタ.

Table 1 Image processing filters used in the filtering layer

Function	Description
Ave3, 5	Averaging filter with 3×3 , 5×5 window
Max3, 5	Maximum filter with 3×3 , 5×5 window
Min3, 5	Minimum filter with 3×3 , 5×5 window
Sob3, 5	Sobel filter with 3×3 , 5×5 window
Lap3, 5	Laplacian filter with 3×3 , 5×5 window
Gau3, 5	Gaussian smooth filter with 3×3 , 5×5 window
LoG3, 5	Laplacian of gaussian filter with 3×3 , 5×5 window
Exp	Expansion processing
Con	Contraction processing
Gab0	Gabor filter with 7×7 , 11×11 window with orientation of 0 degree
Gab45	Gabor filter with 7×7 , 11×11 window with orientation of 45 degree
Gab90	Gabor filter with 7×7 , 11×11 window with orientation of 90 degree
Gab135	Gabor filter with 7×7 , 11×11 window with orientation of 135 degree
Add	Add input two images pixel by pixel
Sub	Subtract input two images pixel by pixel
Mul	Multiply input two images pixel by pixel
Div	Divide input two images pixel by pixel
Abs	Absolute subtraction of input two images pixel by pixel

を行う。フィルタリング層の構造は、Cartesian Genetic Programming (CGP) と同様のフィードフォワードのグラフ構造である。各ノードは、入力ノード、変換ノード、出力ノードの3種類に分けられ、構造内の各ノードは入力ノードもしくは自身より入力ノードに近い1つまたは2つのノードとの接続関係を有する。入力ノードは入力画像の赤成分画像、緑成分画像、青成分画像、グレースケール画像の4つで構成される。変換ノードでは表1で定義された画像処理フィルタを用いて画像変換を行う。出力ノードは入力画像に対して変換ノードを用いて画像変換を行った画像をもつ。本論文では、この画像変換を行った出力画像を特徴マップと呼び、フィルタリング層では出力ノード数分の特徴マップを出力する。ノードの種類と接続関係は進化計算を用いて最適化を行う。

2.3 プーリング層 (P)

プーリング層ではフィルタリング層から出力された $w \times w \times n$ の特徴マップを入力とする。ここで、 $w \times w \times n$ の特徴マップという表現は $w \times w$ 画素の n 枚の特徴マップのことを表す。提案手法では、プーリング処理は $w \times w$ 画素の各特徴マップ上を縦横方向に p 画素ずつ $p \times p$ 画素のウィンドウをずらしながら行われる。このウィンドウの適用間隔をストライド (stride) と呼ぶ。結果として、 $w \times w \times n$ の特徴マップは $(w/p) \times (w/p) \times n$ の特徴マッ

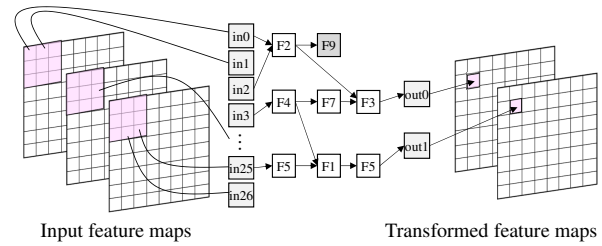


図 1 GP で構築されたフィルタを用いて 8×8 画素の 3 枚の特徴マップを 8×8 画素の 2 枚の特徴マップに変換する変換層の例。入力は注目画素の $3 \times 3 \times 3$ 近傍の画素値である。

Fig. 1 Example of the transformation layer that transforms $8 \times 8 \times 3$ feature maps into $8 \times 8 \times 2$ feature maps using the evolved program. The inputs are the $3 \times 3 \times 3$ pixel intensity values surrounding a target pixel.

表 2 Trasformation layer で使用する関数セット.

Table 2 Function set used in the transformation layer

Function	# Inputs	Description
+	2	Add two inputs
-	2	Subtract two inputs
×	2	Multiply two inputs
÷	2	Divide two inputs
Max	4	The largest value of inputs
Min	4	The smallest value of inputs
Ave	4	The average value of inputs
log	1	Take the natural logarithm for a input
Sqrt	1	Extract a square root of a input
$\times 2.0$	1	Multiply a input by 2.0
$\times 0.5$	1	Multiply a input by 0.5
$\times 0.1$	1	Multiply a input by 0.1

プに変換される。本論文ではプーリング層におけるプーリング処理は平均プーリングを用いる。

2.4 変換層 (T)

変換層ではプーリング層から出力された n 枚の特徴マップと 4 枚の入力画像 (赤成分、緑成分、青成分、グレースケール) を $p \times p$ 画素のウィンドウでそれぞれ平均プーリングした画像を入力とする、すなわち、変換層では $(w/p) \times (w/p) \times (n+4)$ の特徴マップを入力とする。入力画像はフィルタリング層における処理で分類に有効な情報が失われることを防ぐために用いる。変換層の構造はフィルタリング層と同様のフィードフォワードのグラフ構造であり、入力特徴マップにおける注目画素の $s \times s \times (n+4)$ の局所領域の画素値を入力することで変換後の特徴マップの出力値が算出される。グラフ構造の出力ノード数が d の場合、 $(w/p) \times (w/p) \times (n+4)$ の特徴マップは $(w/p) \times (w/p) \times d$ の特徴マップに変換される。図1に $8 \times 8 \times 3$ の特徴マップを $8 \times 8 \times 2$ の特徴マップに変換する例を示す。図1の例では、入力は注目画素の $3 \times 3 \times 3$ 近傍の画素値である。

また、変換層の変換ノードで使用する関数は表 2 に示す通りである。フィルタリング層と同様に、構造中のノードの種類と接続関係を進化計算によって最適化する。

2.5 分類層 (C)

分類層は $h \times h \times m$ の特徴マップを入力とする。 $h \times h \times m$ の特徴マップを $h \times h \times m$ 次元の特徴ベクトルに変換することで、分類に用いる最終的な特徴量とする。本論文では、 k 近傍法を分類器として使用し、近傍数 k は 9 に設定した。

2.6 学習手順

提案手法における特徴構築は二段階の最適化によって行われる。第一段階の最適化では、 $I \rightarrow F \rightarrow P_1 \rightarrow C$ の構造でフィルタリング層 F の最適化を行う。 I は入力画像を表す。フィルタリング層は $w \times w \times n$ の特徴マップを生成し、生成された特徴マップはプーリング層によって $(w/p_1) \times (w/p_1) \times n$ の特徴マップに変換される。 p_1 はプーリング層 P_1 におけるウィンドウサイズを表す。そして、 $(w/p_1) \times (w/p_1) \times n$ の特徴マップを特徴ベクトルに変換し、 k 近傍法に入力することで分類を行う。このときの検証画像セットに対する分類精度を各個体の適応度として用いる。第一段階の最適化によって、最良の画像処理フィルタの組み合わせ F_{best} を得る。

第二段階の最適化では、 $I \rightarrow F_{\text{best}} \rightarrow P_1 \rightarrow T \rightarrow P_2 \rightarrow C$ の構造を用いて変換層 T の最適化を行う。変換層 T では $(w/p_1) \times (w/p_1) \times (n+4)$ の特徴マップが入力され、 $(w/p_1) \times (w/p_1) \times d$ の特徴マップが生成される。そして、生成された特徴マップはプーリング層によって $(w/(p_1 p_2)) \times (w/(p_1 p_2)) \times d$ の特徴マップに変換され、 k 近傍法に入力される。 p_2 はプーリング層 P_2 におけるウィンドウサイズを表す。第二段階目における最適化も第一段階目での最適化と同様に、検証画像セットに対する分類精度を各個体の適応度とする。

3. 小腸画像の分類実験

3.1 実験に用いるデータセット

分類実験にはカプセル内視鏡から撮影された小腸画像を使用した。分類実験では小腸画像 220 枚の中央部 256×256 画素から 64×64 画素のサイズで切り出したブロック画像単位での異常か正常かの 2 クラス分類を行った。切り出す前の小腸画像には専門家によって 8×8 画素のブロックごとに異常、正常のフラグが付与されているため、 64×64 画素のサイズで切り出したブロック画像内に占める異常フラグの割合を用いて異常ブロック画像、正常ブロック画像を作成した。本論文では、異常フラグが占める割合が 0.1 以上のブロック画像を異常ブロック画像、それ以外のブロック画像を正常ブロック画像とし、作成した各クラスからランダムに 1000 枚ずつのブロック画像を選び、学習画像セッ

ト、検証画像セット、テスト画像セットを作成した。本論文で対象とした異常例は、粘膜性病変（白色を呈す）および腫瘍性病変（隆起形状を呈す）である。

3.2 実験設定

提案手法で用いた GP では、個体数は 50、子個体生成数は 10、世代交代モデルは Minimal Generation Gap[18] を用いた。遺伝操作として遺伝子型に対する一様交叉と突然変異を用い、交叉率と突然変異率はそれぞれ 0.8, 0.1 である。第一段階目と第二段階目の最適化における最大世代数はともに 5000 に設定した。

提案手法の有効性を検証するため、次に示す 5 つの手法による小腸画像の分類実験を行った。

- (1) $I \rightarrow F \rightarrow P_1 \rightarrow T \rightarrow P_2 \rightarrow C$ の構造を用いた提案手法 (Proposal 1)。フィルタリング層 F における最大変換ノード数は 40、出力ノード数は 8、変換層 T における入力ノード数は $108 (= 3 \times 3 \times 12)$ 、最大変換ノード数は 60、出力ノード数は 16、プーリング層 P_1, P_2 におけるウィンドウサイズはそれぞれ $p_1 = 2, p_2 = 2$ に設定した。
- (2) Proposal 1 におけるプーリング層 P_1, P_2 のウィンドウサイズをそれぞれ $p_1 = 2, p_2 = 32$ に設定した構造 (Proposal 2)。
- (3) Agapitos らが提案している構造を用いた分類 [17]。ここでは、 $I \rightarrow F_{\text{random}} \rightarrow P_1 \rightarrow T \rightarrow P_2 \rightarrow C$ の構造を用いる。 F_{random} はランダムフィルタリング層を表し、50 個のランダムフィルタによって構成される。各ランダムフィルタは 5×5 画素の受容野をもち、5 つの分布 $U(-1.0, 1.0), U(-5.0, 5.0), UD(1, 5), N(1.0), N(5.0)$ からそれぞれ 10 個のランダムフィルタを生成した。 $U(a, b)$ は $[a, b]$ の範囲の実数値の一様乱数、 $UD(a, b)$ は $[a, b]$ の範囲の整数値の一様乱数、 $N(a)$ は平均 0、標準偏差 a の正規乱数を表す。結果としてランダムフィルタリング層 F_{random} は 50 枚の特徴マップを出力する。変換層 T における入力ノード数は $450 (= 3 \times 3 \times 50)$ 、最大変換ノード数は 200、出力ノード数は 16 に設定した。変換層 T の構造は提案手法と同様のフィードフォワードのグラフ構造を用いた。また、プーリング層 P_1, P_2 におけるウィンドウサイズはそれぞれ $p_1 = 2, p_2 = 2$ に設定した。
- (4) Convolutional neural network (CNN) [19] を用いた分類。CNN は深層学習の一手法であり、近年では画像分類問題において高い分類性能を示している。本論文で用いた CNN の構造を表 3 に示す。畳込み層 (convolution) における活性化関数は ReLU を用いた。バッチサイズは 100、epoch 数は 100 に設定した。学習誤差の算出には交差エントロピー誤差関数を使用し、関数の最小化には Adam 法 [20] を用いた。

表 3 CNN の構造

Table 3 Structure of CNN.

# layer	type	window size	stride	output size
0	data			64 × 64 × 3
1	convolution	5 × 5	1	64 × 64 × 8
2	average pool	2 × 2	2	32 × 32 × 8
3	convolution	3 × 3	1	32 × 32 × 32
4	average pool	2 × 2	2	16 × 16 × 32
5	convolution	3 × 3	1	16 × 16 × 64
6	average pool	2 × 2	2	8 × 8 × 64
7	fully connected			1 × 1 × 256
8	softmax			1 × 1 × 2

(5) 統計特徴量と uniform local binary pattern (ULBP) 特徴 [21] を用いた分類. ここでは画像全体の画素値の平均値, 最大値, 最小値, 中央値, 第一四分位数, 第三四分位数, 標準偏差, 最頻値, 第一 σ 内確率, 第三 σ 外確率, レンジ (最大画素値と最小画素値の差分値) の 11 種類の統計量と 10 次元の ULBP 特徴を RGB カラー画像の各成分からそれぞれ算出した計 63 次元の特徴量を k 近傍法に入力することで分類を行う. カプセル内視鏡は消化器官内を回転しながら画像を撮影するため, 撮影される小腸画像は向きが一意に定まらないことから方向依存性のない特徴量を使用する必要がある. そこで本論文では, 画像全体からの統計特徴量と LBP 特徴を回転不変に拡張した ULBP 特徴を用いる.

3.3 実験結果

3.3.1 従来研究との比較

提案手法と従来手法のテスト画像セットに対する分類結果を表 4 に示す. 提案手法および Agapitos らの手法は 5 試行中の平均分類精度と最良の分類精度結果を示している. CNN は最大 epoch 数 100 のうち, テスト画像セットに対して最も高い分類精度を示した結果を示している.

表 4 の結果から, Proposal 2 が平均分類精度, 最良分類精度ともに最も優れた結果を示していることがわかる. Proposal 1 も従来手法と比べると良好な結果を示している. Agapitos らの手法は提案手法と比べると分類精度が大きく低下してしまっている. これはフィルタリング層 F 以外は Proposal 1 と同一の構造であることから, ランダムフィルタによる変換が小腸画像分類には有効に働いていないことが原因だと考えられる. 小腸画像では照明条件の変化が生じたり, 異常部位や正常部位の種類が多様であることから, 非常に多様性が高い画像となっている. そのため, ランダムフィルタのみではこのような多様性が高い小腸画像から分類に有効な特徴量を構築することは困難であると考えられる. 一方, 提案手法におけるフィルタリ

表 4 テスト画像セットに対する各手法の分類精度

Table 4 Accuracy results of each method on test images.

	Best	Average
Proposal 1	0.816	0.807
Proposal 2	0.842	0.828
Agapitos [17]	0.658	0.632
CNN	0.788	-
統計特徴量+ULBP	0.759	-

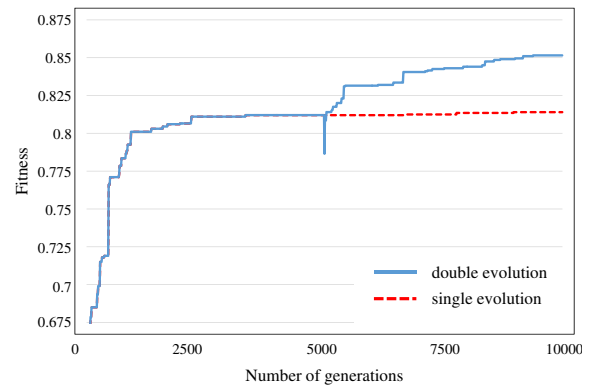


図 2 提案手法の適応度推移.

Fig. 2 Transition of the fitness of the proposed method.

ング層 F では既存の画像処理フィルタを多段に適用することで, 単一のランダムフィルタでは構築することが難しい特徴量を構築することを可能とし, 分類精度で優れた結果を示したと考えられる. また, 提案手法は CNN と比べても優れた性能を示した. これは CNN のような深層学習の手法では一般的に大規模な学習画像を必要とするため, 各クラス 2000 枚ずつの計 4000 枚の学習画像では十分な特徴構築を行うことができなかったと考えられる. 本論文で扱っている小腸画像のように大規模な学習画像セットの用意が難しい問題では, 提案手法のように少ない学習画像セットから分類に有効な特徴構築が行える手法は有望であると考えられる. さらに, 方向依存性のない特徴量である統計特徴量と ULBP 特徴を用いた分類よりも提案手法は優れた性能を示した. これは画像処理フィルタおよび GP フィルタによる画像変換とプーリング処理によって, 方向依存性のない特徴量を自動で構築できたためであると考えられる. このことから, 提案手法は対象となる画像に応じて適切に特徴量を自動構築できることが確認できた.

3.3.2 提案手法に関する考察

5 試行中における提案手法の最良個体の適応度推移の結果を図 2 に示す. 図 2 には, $I \rightarrow F \rightarrow P \rightarrow C$ の構造で 10000 世代の最適化を行った結果 (single evolution) と, 先に述べた構造で 5000 世代の時点での最良個体を固定し $I \rightarrow F_{\text{best}} \rightarrow P \rightarrow T \rightarrow P \rightarrow C$ の構造に変更してさらに 5000 世代の最適化を行った結果 (double evolution) を示している. F_{best} は 5000 世代の時点での single evolution

の最良個体を表す。5000 世代までは同一の個体を表しているため、適応度推移は一致している。図 2 で示されているように、既存の画像処理フィルタのみを用いた構造である $I \rightarrow F \rightarrow P \rightarrow C$ よりも 2 種類の画像変換を用いた構造である $I \rightarrow F_{\text{best}} \rightarrow P \rightarrow T \rightarrow P \rightarrow C$ の方が高い適応度を示していることがわかる。特に、既存の画像処理フィルタのみを用いた構造では 5000 世代を過ぎると進捗が停滞してしまっているが、2 種類の画像変換を用いた構造では 5000 世代を過ぎても適応度が上昇していることがわかる。これは、変換層 T では GP を用いて新たなフィルタを構築しているため、フィルタリング層 F で使用している既存のエッジ検出や平滑化処理などを行う画像処理フィルタだけでは表現できないフィルタを二段階目の最適化で構築できているためであると考えられる。

4. まとめ

本論文では、画像分類のための階層的な特徴構築手法を提案した。提案手法は (1) 既存の画像処理フィルタの組み合わせによる画像変換、(2) GP で構築したフィルタによる画像変換、の 2 種類の処理を二段階の進化計算によって最適化する。提案手法をカプセル内視鏡から撮影された小腸画像における異常画像と正常画像の分類問題に適用し、従来手法との比較と性能の検証を行った。実験結果から、従来手法よりも優れた分類精度を示すことが確認できた。また、2 種類の処理を二段階の進化計算によって最適化することで 1 種類の処理を一段階の進化計算によって最適化するよりも優れた性能を示したことから、異なる処理の構造を多層化することの有用性を示した。

今後の課題として、フィルタリング層や変換層の構造に関するパラメータも同時に最適化することで、より汎用性の高い特徴構築手法の確立が挙げられる。さらに、提案手法の拡張として様々な画像データセットに対する分類実験や、手法の信頼性と精度向上のために、処理内容の解析を行う必要があると考えている。

参考文献

[1] Ahonen, T., Hadid, A. and Pietikäinen, M.: Face recognition with local binary patterns, *Proc. European Conference on Computer Vision*, pp. 469–481 (2004).
 [2] Dalal, N. and Triggs, B.: Histograms of oriented gradients for human detection, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 886–893 (2005).
 [3] Lowe, D. G.: Distinctive image features from scale-invariant keypoints, *International journal of computer vision*, Vol. 60, No. 2, pp. 91–110 (2004).
 [4] Xiao, J., Hays, J., Ehinger, K. A., Oliva, A. and Torralba, A.: Sun database: Large-scale scene recognition from abbey to zoo, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 3485–3492 (2010).
 [5] Bengio, Y., Courville, A. and Vincent, P.: Repre-

sentation learning: A review and new perspectives, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 35, No. 8, pp. 1798–1828 (2013).
 [6] Hinton, G. E., Osindero, S. and Teh, Y.-W.: A fast learning algorithm for deep belief nets, *Neural computation*, Vol. 18, No. 7, pp. 1527–1554 (2006).
 [7] Larochelle, H., Bengio, Y., Louradour, J. and Lamblin, P.: Exploring strategies for training deep neural networks, *The Journal of Machine Learning Research*, Vol. 10, pp. 1–40 (2009).
 [8] Koza, J. R. and Poli, R.: Genetic programming, *Search Methodologies*, Springer, New York, USA, pp. 127–164 (2005).
 [9] Zhang, M. and Smart, W.: Multiclass object classification using genetic programming, *Applications of Evolutionary Computing*, pp. 369–378 (2004).
 [10] Zhang, M., Gao, X. and Lou, W.: A new crossover operator in genetic programming for object classification, *Vol. 37, No. 5*, pp. 1332–1343 (2007).
 [11] Al-Sahaf, H., Song, A., Neshatian, K. and Zhang, M.: Two-Tier genetic programming: towards raw pixel-based image classification, Vol. 39, No. 16, Elsevier, pp. 12291–12301 (2012).
 [12] Miller, J. F. and Thomson, P.: Cartesian Genetic Programming, *European Conference, EuroGP 2000* (Poli, R., Banzhaf, W., Langdon, W. B., Miller, J., Nordin, P. and Fogarty, T. C., eds.), Springer-Verlag, pp. 121–132 (2000).
 [13] Harding, S.: Evolution of image filters on graphics processor units using cartesian genetic programming, *IEEE Congress on Evolutionary Computation*, pp. 1921–1928 (2008).
 [14] Kowaliw, T., Banzhaf, W., Kharma, N. and Harding, S.: Evolving novel image features using genetic programming-based image transforms, *IEEE Congress on Evolutionary Computation*, pp. 2502–2507 (2009).
 [15] Shirakawa, S., Nakayama, S. and Nagao, T.: Genetic image network for image classification, *Applications of Evolutionary Computing*, pp. 395–404 (2009).
 [16] Hirano, Y. and Nagao, T.: Feature transformation using filter array for automatic construction of image classification, *IEEE 7th International Workshop on Computational Intelligence and Applications*, pp. 59–64 (2014).
 [17] Agapitos, A., O’Neill, M., Nicolau, M., Fagan, D., Kattan, A., Brabazon, A. and Curran, K.: Deep evolution of image representations for handwritten digit recognition, *IEEE Congress on Evolutionary Computation*, pp. 2452–2459 (2015).
 [18] 佐藤浩, 小野功, 小林重信: 遺伝的アルゴリズムにおける世代交代モデルの提案と評価, *人工知能学会誌*, Vol. 12, No. 5, pp. 734–744 (1997).
 [19] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278–2324 (1998).
 [20] Diederik, K. and Jimmy, B.: Adam: A method for stochastic optimization, *International Conference on Learning Representations*, pp. 2452–2459 arXiv:1412.6980 (2015).
 [21] Ojala, T., Pietikäinen, M. and Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp. 971–987 (2002).