

# 決定木および決定ネットワークによる 画像分類過程の説明の自動生成

崎津 実穂<sup>1,a)</sup> 菅沼 雅徳<sup>1</sup> 土屋 大樹<sup>1</sup> 長尾 智晴<sup>1</sup>

受付日 2015年9月3日, 再受付日 2015年10月19日,  
採録日 2015年11月23日

**概要:** 近年, 機械学習を用いて認識器や分類器を自動構築する研究がさかんに行われ, その有効性が示されている. しかしながら, 分類の過程や使用される特徴量が複雑になり, ブラックボックス化することが多い. 一方で, 計算機を用いた画像の自動分類システムを利用する際, 説明責任が求められる場合が存在する. 筆者らの研究グループでは先に, 人にとって分類過程が理解しやすいといわれている決定木を拡張し, 可読性が高い分類器を自動構築する進化的条件判断ネットワーク (Evolutionary DEcision Network; EDEN) を提案した. これらの手法は if-then ルールを用いており, 他の分類器と比較すると分かりやすい構造をとっている. しかしながら, 分類器の構造が複雑になると, どのような分類過程であるかを人間が把握することは難しくなると考えられる. そこで本論文では, あらかじめ構築された決定木や EDEN の分類過程を説明する文章を自動生成する手法を提案する. 提案手法では, ユーザが説明の粒度を指定した後, 分類に使用する特徴量やしきい値をあらかじめ定義した辞書内の語句に変換することで分類過程の説明文を自動生成する. 提案手法を一般画像と医用画像の分類器それぞれに適用して説明文を生成する実験を行い, その有効性を確認した.

キーワード: 分類過程の解析, 画像分類, 機械学習, 決定木, 進化計算法

## Automatic Generation of Sentences Explaining Image Classification Processes Constructed by Decision Tree and Decision Network

MIHO SAKITSU<sup>1,a)</sup> MASANORI SUGANUMA<sup>1</sup> DAIKI TSUCHIYA<sup>1</sup> TOMOHARU NAGAO<sup>1</sup>

Received: September 3, 2015, Revised: October 19, 2015,  
Accepted: November 23, 2015

**Abstract:** The automatic construction of image classification algorithms using machine learning has been studied and is shown to be effective. Although classifiers constructed through such methods are very effective, these structures and classification processes cannot be easily understood by humans. A decision tree has a readable structure among classifiers. Further, we previously proposed evolutionary decision network (EDEN) that places greater importance on the ease of understanding for humans. These methods are easier to understand than other machine learning methods because they use if-then rules. However, they become more difficult to understand as the structures of classifiers increase in complexity. In this paper, we propose a method that generates sentences to explain classification processes constructed using a decision tree and EDEN in advance. The proposed method converts the classification processes into explanatory sentences using the dictionary of word-feature pairs and word-threshold pairs. The explanatory sentences are generated according to the granularity specified by the user. We apply this method to test problems involving two-class categorization of generic images and two-class categorization of medical images. The results indicate that the proposed method generates appropriate explanatory sentences that are easy to understand for humans.

**Keywords:** analysis of classification processes, image classification, machine learning, decision tree, evolutionarily algorithms

<sup>1</sup> 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

<sup>a)</sup> sakitsu-miho-jx@ynu.jp

## 1. はじめに

近年、ニューラルネットワークや Support Vector Machine などの機械学習を用いて分類器を自動構築する研究がさかんに行われ、その有効性が示されている。これらの分類器の多くは社会や産業の現場で実際に使用されることが前提とされているが、これまでは、分類精度だけが重要視され、機械学習によって構築された分類器を使用する人の立場になって考えることがあまり行われていなかったと考えられる。しかしながら、分類器の信頼性を示すためには分類の根拠や分類過程を提示することが求められる場合があり、分類器の可読性も重要視されている。特に人の健康や命に関わる産業用画像処理、車載の認識器、医用画像診断などではその傾向が強い。このため、機械学習を用いた分類器が今後、社会・産業でさらに利用されるようになるためには、分類器の精度だけでなく、分類器が有用であるかということ人間が容易に確認できるようにするため、分類器の可読性の抜本的改革やアルゴリズムの可視化が求められる。

分類器の中でも比較的読解性があり、人に理解されやすい手法として if-then ルールを用いて分類を行う決定木が知られている。決定木は、条件判断を行うノードが木構造状に配置された構造の分類器である。根から入力されたデータに対して各ノードで特徴量を用いた分岐を繰り返し、最終的に到達した終端ノードに対応するクラスに輸入データを分類する。また、筆者らのグループでは、高精度かつコンパクトで人に理解されやすい構造を持つ分類器を自動構築する手法として、決定木を拡張した進化的条件判断ネットワーク (Evolutionary DEcision Network; EDEN)\*<sup>1</sup>[1] を提案している。EDEN は分類対象の入力データおよび参照入力データに対して単純な条件判断を行うノード群を、進化計算法を用いてネットワーク状に自動構築する。画像分類実験において、EDEN を用いることで決定木に比べてノード数を大幅に削減した構造で分類が可能になることが示されている [1]。

これらの手法は if-then ルールを用いて分類を行うため、分類過程が比較的読解しやすく、可読性があるといえる。しかしながら、分類器の構造が複雑になるにつれ、人間が分類器の構造を提示されるだけですべての分類過程を理解することは困難になると考えられる。

これまで、画像を計算機に入力して理解させ、その画像を言語を用いて説明を行う研究が行われている。人間が画像を認識する過程や画像中のどのような構成要素に注目しているかを風景画を用いて検証し、ファジィ論理と組み合わせることで計算機に言語表現を行わせるといった研究が行われている [2]。その他にも、画像を入力し、画像認識と

説明文の生成を組み合わせた研究 [3], [4], [5] が行われており、近年では deep learning を用いて入力画像に対する説明文を得る研究も行われている [6], [7]。また、画像だけでなく、動画中の人物の動きを抽出し、行動についての説明文を出力する研究 [8] も行われている。これらの手法は画像についての説明文を生成する手法であり、どのような過程で説明文が出力されたのか、すなわち物体の分類過程や認識過程を出力することはできない。また、臨床において安心して機械学習を用いたシステムを利用してもらうために、医用画像の分類結果に対する根拠を提示することで分類器の信頼性に対する課題を解決しようとする研究もなされている [9]。この研究では分類に使用されたファジー if-then ルールを根拠として提示しているが、特徴量の値を数値そのまま提示しており、この特徴量と数値の組合せの持つ意味を言語に変換することなどによって、より直感的に分類過程を理解することができると考えられる。

また、ファジー推論ニューラルネットワークを用いて風景画像や人工物に対して画像認識を行い、学習で得られたファジー if-then ルールを適当な言語を当てはめることで言語形式に変換するといった研究 [10], [11] がある。しかしながら、文献 [10], [11] の手法は具体的な文章への変換方法が示されていない。また、決定木や EDEN のように、if-then ルールを用いて分類を行う分類器がどのような分類過程を経て結果を出力したのかということ、言語によって分かりやすく説明するといった研究は筆者らが知る限りまだなされていない。分類器が高精度であるだけでなく、画像の分類過程を分かりやすく提示することで、分類器が有用であるかを利用者が簡単に検証可能とすることは、今後機械学習がますます発展し、社会において安心して計算機による知能情報処理を利用してもらうために必須であると考えられる。

そこで本論文では、構築された分類器の処理アルゴリズムに対する説明責任を果たし、人間が分類器を使うにあたってその分類過程が有用であるかの確認を容易にすることを目的とし、比較的説明しやすくと考えられる決定木および決定ネットワーク EDEN の 2 つの手法を対象として、画像の分類過程を説明する文章を自動生成する手法を提案する\*<sup>2</sup>。

## 2. 決定木および決定ネットワーク EDEN

### 2.1 決定木

決定木は、機械学習を用いて条件判断を行うノードを木構造状に配置し、分類を行う手法である。入力ノードから分類対象データの特徴量を入力し、中間ノードでの条件判断によって分岐先を決定する。これを終端ノードに到達するまで繰り返し、終端ノードに対応するクラスに輸入デー

\*1 特許番号：5548990 号

\*2 特願 2015-041313

タを分類する．決定木を生成する代表的なアルゴリズムとして，ID3 [12] や C4.5 [13] が提案されている．

ID3 は，空の木から始めて，各ノードにおいて最もデータ集合のエントロピーを削減する，すなわち最大の情報利得を持つ属性を選択し，元のデータ集合を部分集合に分割していく．これを繰り返し，徐々にノードを付け加えてデータ集合を分割させていくことで，最終的な木構造が得られる．

C4.5 は，ID3 のアルゴリズムを拡張した手法である．C4.5 は，連続値を扱うことができる点や，学習データの属性値が欠損している場合でも学習データに用いることができる点が ID3 と異なっている．また，ID3 のように情報利得を用いて分岐ノードに使用する属性を決定する場合，分岐数が多い属性が選択されやすく，良好な結果が得られない場合がある．そのため C4.5 では，情報利得ではなく，情報利得を分割情報量（ある属性で分割を行った際のエントロピー）を用いて正規化した情報利得比を使用して，各ノードで分岐に使用する属性値を決定する．また，過学習を防ぐために木構造の枝刈りを行うことも特徴の一つである．

## 2.2 決定ネットワーク EDEN

筆者らの研究グループでは，決定木をネットワークに拡張し，高精度かつコンパクトで可読性のある構造を持つ分類器を自動構築する手法として進化的条件判断ネットワーク EDEN を提案している [1]．EDEN では，分類対象の入力データに対して単一の特徴量を用いて単純な条件判断を行うノードを，進化計算法によってフィードフォワード型のネットワーク状に自動構築する．各条件判断ノードは，分岐条件とする特徴量とその分岐の判断基準となるしきい値，および出力先で構成され，EDEN はこれらの組合せを最適化する．分類を行う場合は，決定木と同様に入力ノードから分類対象データの特徴量を入力し，条件判断ノードで大小判定を行い分岐先を決定する．最終的に到達した出力ノードは，分類対象データがどのクラスに属するかを表している．このため，EDEN のネットワークの分類過程も if-then ルールで表すことができ，可読性のある構造となっている．

EDEN の表現型とその表現型に対応する遺伝子型をそれぞれ図 1，図 2 に示す．遺伝子型から表現型へは一意に変換することができる．遺伝子型は一次元の数字列で表されており，これに対して遺伝操作を行う．各染色体は，ノードの種類（入力，中間，出力），分岐に使用する特徴量の種類，出力先，およびしきい値となるパラメータを保持している．ノードの出力数はそのノードの種類によって一意に決まっており， $n$  出力の場合は  $n + 1$  番目以降の出力先は無視される．図 2 の場合は  $n = 3$  であるため，4 番目の出力先は無視される．

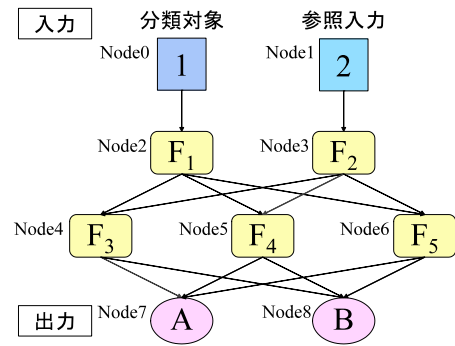


図 1 EDEN の構造例（表現型）

Fig. 1 A phenotype of EDEN.

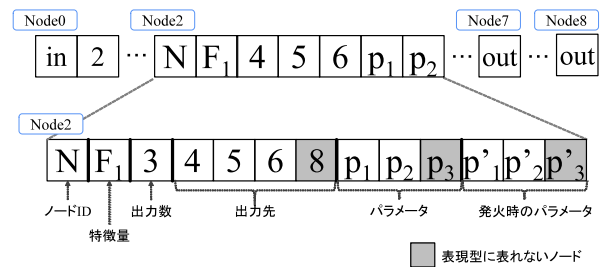
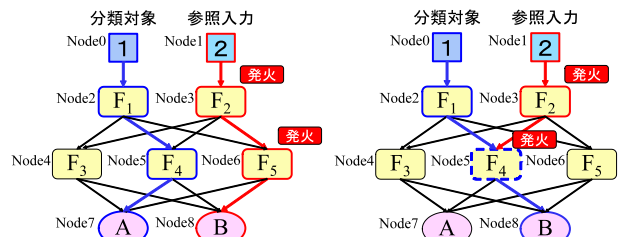


図 2 EDEN の構造例（遺伝子型）

Fig. 2 A genotype of EDEN.



(a) 参照入力による変更がない場合 (b) 参照入力による変更がある場合

図 3 参照入力による処理の変更例

Fig. 3 An example of interactions.

また，EDEN では分類対象となるデータだけでなく関連する参照データを入力することができる．この参照データによって分類対象データの分類処理を変更することで少ないノードでも適切な分類ができるようになることを期待している．参照データの入力が分類対象データの処理に影響を及ぼさない例と影響を及ぼす例を図 3 に示す．まず，参照データを対応する入力ノードから入力し，条件分岐によって通過したノードを発火状態とする．発火状態となったノードでは，分類対象データが通過した際には発火していない状態と異なる処理が行われる．具体的には，分岐の判断をする際のしきい値を，発火前と発火後で異なる値を用いることで処理の変更を行う．図 3 では，参照データが通過した Node5 が発火したことで，Node5 での処理が変更され，分類対象データの分類結果が変化している．



### 3. 提案手法

#### 3.1 提案手法の概要

本論文では、決定木や EDEN のような if-then ルールを用いた分類器の分類過程を説明する文章を自動で生成する手法を提案する。提案手法では、分類に使用する特徴量やしきい値に対応する語句を定義した辞書を用意しておき、あらかじめ構築された分類器や学習データなどの特徴量の種類や値をこの辞書の語句と対応づけることで説明文に変換する。この際、文章の数や文章の長さを指定することでユーザの求める粒度で説明文を生成することができる。また、提案手法を用いることで、分類器自体に対する説明文、学習データの集合に対する説明文、特定の1つのデータに対する説明文の3種類の文章を生成することができる。

#### 3.2 変換辞書

提案手法では、分類に使用する特徴量やしきい値に対応する語句を定義した辞書をあらかじめ用意しておく。具体的には、次のような例が考えられる。

- 色特徴量 HLS の L の平均値：「明るさ」
- 色特徴量 HLS の H の分散：「色数」
- 分類データの値がしきい値 0.4 未満：「少ない」
- 分類データの値がしきい値 0.4 以上：「ある程度多い」
- 分類データの値がしきい値 0.8 以上：「とても多い」
- 分類データの値がしきい値 0.8 未満：「とても多いわけではない」

後に文章が続く場合は、「とても多くて」のように語尾を変化させることで対応させる。

また、EDEN のネットワークでは参照入力によってノードのしきい値が変更される場合がある。この場合は、参照入力の影響を付帯条件として、語句表現に反映させる。具体的には、参照入力によって発火したノードに接続する、1つ前のノードについて説明する語句を追加する。図3の(b)の例では、発火した Node5 について説明する場合に Node3 についての説明を追加する。具体的には、参照入力として分類対象データの周囲の特徴量を使用する場合、付帯条件として追加する語句には次のような例が考えられる。

- 1つ前のノードで L の平均値が値 0.9 で分岐を行った場合：「周囲がとても明るくて」
- 1つ前のノードで H の分散が値 0.1 で分岐を行った場合：「周囲の色数がほとんどなくて」

#### 3.3 説明文の生成

提案手法では、分類データを流した際に通過したパス上に存在するノードで使用されている特徴量およびしきい値を、辞書を参照して対応する語句に変換して説明文を自動生成する。その際、できるだけ詳しく分類過程を説明してほしいと考える利用者から、おおまかにどのような分類を

行っているのかを知りたいという利用者まで様々な利用者が存在すると考えられる。そこで、提案手法ではどの程度分類器について詳細に説明するのかを2つのパラメータで表す。

まず1つ目は、説明するパスの数  $N$  である。本論文では、分類器にデータを流した際に多くのデータが通るパスほど、分類クラスを表すのに重要なパスであると考えられる。このため、構築された決定木や EDEN のネットワークに対して、全分類対象データを流した際のノードおよびパスの通過頻度を計測する。その後、指定された必要なパスの数  $N$  だけ、データの通過確率が高いパスから順に説明文を生成する。

2つ目は文章の長さ  $n_{\max}$  である。決定木ではルートノードからエントロピーが最も削減される順にノードが配置されているため、ルートノードに近いノードの方がデータの分類において重要であると考えられる。そこで、提案手法では説明文を生成する際に、ルートノードから指定された  $n_{\max}$  個のノードを使用する。すべてのノードを使用する場合は、 $n_{\max}$  をノードの最大値  $\max$  とすることで入力ノードから出力ノードまですべてのノードを説明する文章を生成することができる。EDEN においては、必ずしもルートノードに近いノードほど重要であるとは一概にいけない場合もあるが、本論文では決定木と同様にルートノードに近いノードから使用して説明文を生成することとする。

#### 3.4 説明文の種類

提案手法を用いることで、1つの分類器に対して複数の種類の文章を生成することができる。本論文では、3種類の文章の生成について説明する。

まず1種類目として、分類器のしきい値そのものを語句に変換し、説明文を生成する方法がある。構築された分類器における各ノードのしきい値を、辞書と照らし合わせて文章に変換する。この場合は、分類器に利用されているしきい値そのものを語句に変換するため、分類器がどのような構造になっているのか、この分類器を利用してどのようなデータがどのクラスに分類されるのかを説明文として生成することができる。

2種類目として、学習データの特徴量の平均値を語句に変換し、説明文を生成する方法がある。説明文を生成するパスの各ノードにおいて、そのノードを通ったデータの特徴量の平均値を計算し、この平均値と辞書を照らし合わせて文章に変換する。この場合は、特徴量の平均値を語句に変換するため、学習データの集合がどのような性質を持っているのかを説明文として生成することができる。決定木や EDEN では、学習データはそのクラスを表す代表的なデータを選択することが一般的で、未知データは学習データと似た性質を持っていると考えられる。このためこの方法で説明文を生成することで、各クラスに分類されるデー

表 1 EDEN の最適化のパラメータ

Table 1 Parameters of EDEN.

世代交代モデル	MGG [15]
世代数	100000
個体数	100
子個体数	30
交叉率	0.9
一様交叉率	0.1
突然変異率	0.05
入力ノード	2
最大中間ノード数	50
出力ノード	2

タの性質を予測することができると考えられる。

3種類目として、データの集合ではなく、ある1つのデータに対して説明文を生成する方法がある。構築された分類器を使用する際、学習データの中のある1つのデータや特定の未知データの分類過程を確認したいという場合があると考えられる。このような場合は、説明文を生成したいデータを分類器に入力し、分類パス上における各ノードの特微量の種類、値を辞書と照らし合わせて文章に変換する。この方法によって、分類器やデータ集合全体ではなく、特定の1つのデータについてどのような分類が行われているのかを知るための説明文を生成することができる。

## 4. 説明文生成実験

### 4.1 実験概要

提案手法を用いて、一般画像と医用画像を対象とした2種類の2クラス分類問題について実験を行った。

決定木については、オープンソースソフトウェアである Weka (ver. 3.7.9) [14] を用い、C4.5 アルゴリズムによって決定木を生成した。パラメータは、デフォルトである信頼度 0.25、リーフノードにおける最小データ数 2 を用いた。

EDEN では、ネットワークの分類精度を高めると同時に、総ノード数ができる限り少なくなるように遺伝的アルゴリズムを用いて最適化した。最適化における各個体の適応度 fitness は式 (1) を使用した。

$$\text{fitness} = \alpha \times \text{RC} + \beta \times \frac{1}{C}, \quad (1)$$

ここで、RC は各クラスの再現率 recall の積であり、C は構築されたネットワークの条件判断ノードの総数である。また、 $\alpha$  および  $\beta$  は変更可能なパラメータであり、本論文の実験においては  $\alpha = 0.9$ ,  $\beta = 0.005$  とした。再現率 recall は、データの総数を  $D$ 、正しく分類したデータ数を  $D_{\text{correct}}$  とすると、式 (2) で表される。

$$\text{recall} = \frac{D_{\text{correct}}}{D}, \quad (2)$$

また、その他の最適化におけるパラメータを表 1 に示す。

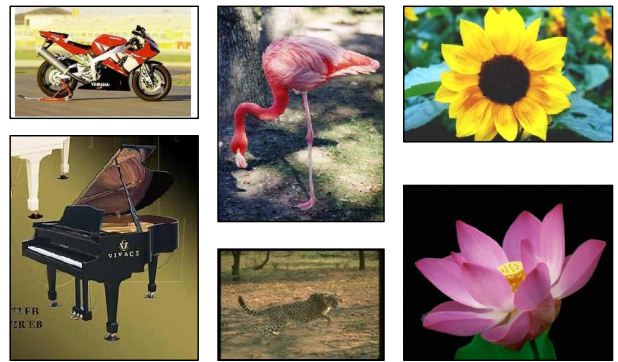


図 4 実験に使用した画像例

Fig. 4 Example images used for the experiments.

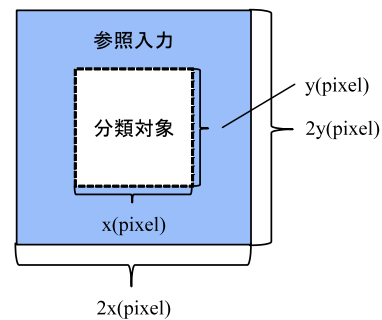


図 5 EDEN の分類対象と参照入力

Fig. 5 Target of classifying and reference input areas of EDEN.

### 4.2 一般画像分類

一般画像の2クラス分類として、バイクとグランドピアノ、フラミンゴと豹、ひまわりと蓮を分類する3種類の実験を行い、学習データに対して説明文の生成を行った。画像セットは Caltech-101 \*3 から、各クラスそれぞれ 50 枚の画像を用い、画像 1 枚単位での分類を行った。使用した画像の例を図 4 に示す。C4.5 については画像全体から特微量を算出し、EDEN では図 5 のように画像の中心から画像サイズの 0.5 倍の部分进行分类対象、それ以外の周囲の部分参照入力として使用した。EDEN の条件判断ノードとしては、入力特微量としきい値による 2 分岐のノードだけを使用した。使用した HLS 色特微量とエッジに関する特微量 15 種類と語句の対応表を表 2 に示す。このとき、各ノードのしきい値やデータの平均値に割り当てる語句は、0.0 以上 0.2 未満、0.2 以上 0.4 未満のように 0.2 ずつの範囲で異なるものを割り当てている。たとえば平均値が 0.0 以上 0.2 未満の場合は「とても暗い」、0.6 以上 0.8 未満の場合は「明るい」ようになる。また、S や L の標準偏差などは「鮮やかさの差がほとんどない」、「明るさの差がとても大きい」などと表 2 の語句の後に副詞や形容詞を用いることでそれぞれの値を表現している。一方で S の平均値などは、このように形容詞や副詞を付加して各値を表現するのではなく、「にぶい色」、「鮮やかな色」といったよう

\*3 [http://www.vision.caltech.edu/Image\\_Datasets/Caltech101/](http://www.vision.caltech.edu/Image_Datasets/Caltech101/)

表 2 一般画像分類実験で用いた特徴量と語句の対応表

Table 2 The dictionary of word-feature pairs for generic image classification.

特徴量	対応する語句
H が 30° 未満または 330° 以上の画素の割合	赤色っぽい部分
H が 30° 以上 90° 未満の画素の割合	黄色っぽい部分
H が 90° 以上 150° 未満の画素の割合	緑色っぽい部分
H が 150° 以上 210° 未満の画素の割合	水色っぽい部分
H が 210° 以上 270° 未満の画素の割合	青色っぽい部分
H が 270° 以上 330° 未満の画素の割合	紫色っぽい部分
L の平均値	明るさ
S の平均値	鮮やかさ
H の標準偏差	色数
L の標準偏差	明るさの差
S の標準偏差	鮮やかさの差
水平方向の ±20° のエッジを持つ画素の割合	横線
垂直方向の ±20° のエッジを持つ画素の割合	縦線
上記以外の方向のエッジを持つ画素の割合	斜線
エッジ強度の平均	輪郭

表 3 一般画像分類の正解率と総ノード数

Table 3 The correct classification rates and the total number of nodes for generic image classification.

	C4.5		EDEN	
	分類精度	ノード数	分類精度	ノード数
バイク・グランドピアノ	98.0%	19	100.0%	9
フラミンゴ・豹	99.0%	7	100.0%	9
ひまわり・蓮	99.0%	19	100.0%	11

に、値によって表 2 の語句に関連する別の語句を対応付けて辞書に登録してある。

また、C4.5 と EDEN による一般画像分類の精度と構築された構造の総ノード数を表 3 に示す。これらの学習結果を用いて、説明文の生成を行った。バイクとグランドピアノに関しては分類器の説明文、フラミンゴと豹に関しては学習データに関する説明文、ひまわりと蓮に関しては図 4 の画像についての説明文を生成した。指定したパラメータは  $N = 1$ ,  $n_{max} = \max$  である。生成された説明文の例を次に示す。また、バイクとグランドピアノを分類する C4.5 と EDEN の構造を図 6、図 7 に示す。

- C4.5 によるバイクとグランドピアノの分類過程の説明文
  - 横線がある程度多くて、輪郭がある程度はっきりしているため、バイクである。
  - 横線が少なく、緑色っぽい部分がほとんどなくて、明るさの差が少なく、鮮やかさの差がほとんどないため、グランドピアノである。
- EDEN によるバイクとグランドピアノの分類過程の説明文
  - 明るさの差が少なく、輪郭がある程度はっきりしているため、バイクである。

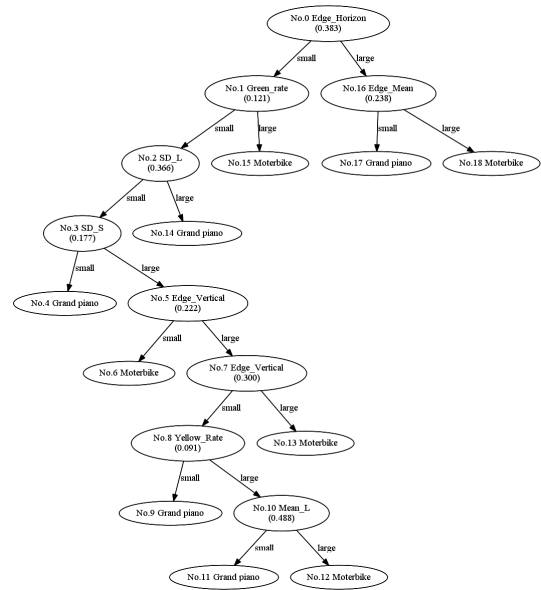


図 6 バイクとグランドピアノを分類する C4.5 の構造

Fig. 6 The C4.5 classifier for moterbike and grand piano images.

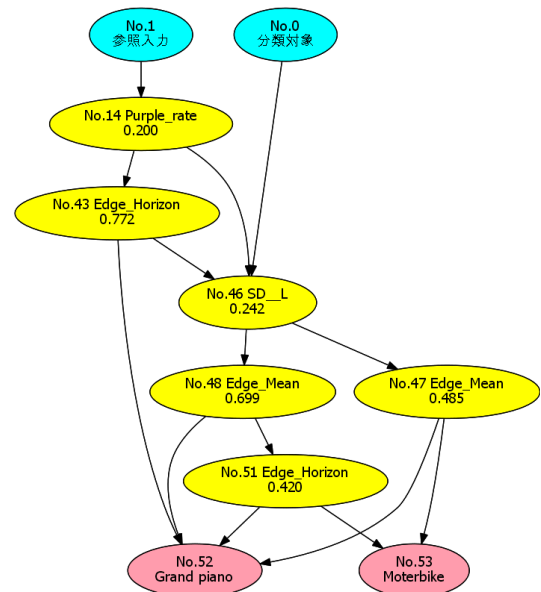


図 7 バイクとグランドピアノを分類する EDEN の構造

Fig. 7 The EDEN classifier for moterbike and grand piano images.

- 明るさの差がある程度多くて、輪郭があまりはっきりしないため、グランドピアノである。
- C4.5 によるフラミンゴと豹の学習データの説明文
  - 色数が少なく、赤色っぽい部分がある程度多いため、フラミンゴである。
  - 色数がとても少ないため、豹である。
- EDEN によるフラミンゴと豹の学習データの説明文
  - 赤色っぽい部分が多くて、色数が少ないので、フラミンゴである。
  - 赤色っぽい部分がほとんどなくて、黄色っぽい部分



がとても多くて、青色っぽい部分がほとんどないため、豹である。

- C4.5 によるひまわりと蓮の分類過程の説明文
  - 黄色っぽい部分がある程度多くて、ある程度鮮やかな色で、ある程度明るくて、斜線がある程度多いため、ひまわりである。
  - 黄色っぽい部分がほとんどないため、蓮である。
- EDEN によるひまわりと蓮の分類過程の説明文
  - 黄色っぽい部分が多くて、ある程度明るくて、色数が少ないため、ひまわりである。
  - 黄色っぽい部分がほとんどなくて、色数が少ないため、蓮である。

提案手法を適用することで、分類器の構造や流したデータの分類パスをネットワーク上に表示するなどの方法よりも理解しやすい表示ができていけるといえる。図 4 の画像と比較しても、特徴を良好にとらえることができている。構築された分類器による分類が妥当であることが分かる。たとえばバイクとグランドピアノの分類では、バイクはグランドピアノとは異なり細かいパーツが多いためエッジがはっきり出る部分が多く、エッジ強度の平均値が高い。そのためバイクの説明文には輪郭がはっきりしているという内容の語句が多く含まれている。また、フラミンゴと豹の分類では豹の画像と比較してフラミンゴの画像で支配的な色であると考えられる赤色に関する語句が文章に含まれている。ひまわりと蓮の分類に関しても同様に、蓮の画像と比較してひまわりの画像で支配的な色であると考えられる黄色に関する語句が文章に含まれている。なお、今回の実験では分類器を構築する際、分類対象となる物体と背景を分離せず特徴量を算出して使用している。そのため、データセットに依存した本来分類に使用すべきでない特徴量、たとえば対象物体以外の特徴量を使用して分類器が構築される可能性がある。例として C4.5 によるグランドピアノの分類過程の説明文では、「緑色っぽい部分がほとんどなくて」という語句が文中で使用されている。今回使用したバイクのデータセットには、背景に草木が写っている画像が一定数あり、この「緑色っぽい部分」という表現は、バイク部分の特徴量でなく背景の特徴量を使用していると考えられる。このように、提案手法を用いて分類器に関する文章を生成することによって、対象物体を分類するのに適切な分類を行っているかどうか確認することが容易になると考えられる。

また、グランドピアノを説明する文章のように、C4.5 で構築した分類器に関する説明文では「明るさの差が少ない」、EDEN で構築した分類器に関する説明文では「明るさの差がある程度多い」と同じ特徴量に対して異なる文章が出てくる場合がある。これは、C4.5 と EDEN では分類器を構築する際の特徴量やしきい値の選択方法が異なるため、各分類器で最もデータが通るパスの性質が異なるもの

になったからであると考えられる。

また、分類精度自体は高いものの、文章を生成すると人間の判断過程とは異なるような説明文が出てくる場合がある。本論文では、分類過程を説明するための文章を生成するために 2 クラス分類問題に対して提案手法を適用している。計算機にとっては「2 つが異なるものである」という分類をすることができればよいから、人間の判断とは異なるプロセスで分類を行っていると考えられ、人間が考えるような特徴を抽出する文章にはなっていない。分類器を構築し、ある 1 つの物体の特徴を抽出したいといった場合には、今回のような 2 クラス分類ではなく、多クラスの中から 1 クラスを抽出するような問題で分類器を構築することで、提案手法によって分類データの特徴を抽出するような文章を生成することも可能であると考えられる。

### 4.3 医用画像分類

医用画像として、カプセル内視鏡によって撮影された小腸画像を用いて異常と正常の 2 クラス分類の実験を行い、学習データに対して説明文の生成を行った。カプセル内視鏡から撮影した画像の例を図 8 [16] に示す。使用した小腸画像のサイズは  $256 \times 256$  pixel であり、専門家によって  $8 \times 8$  pixel のブロックごとに正常、異常のフラグが付与されている。この小腸画像を  $8 \times 8$  に分割し、 $32 \times 32$  pixel のブロック単位で異常 880 ブロックと正常 880 ブロックの分類を行った。 $32 \times 32$  pixel のブロックにおいては、異常ブロックが 2 割以上含まれている場合に異常とした。C4.5 と EDEN の分類対象ノードから入力する特徴量は、 $32 \times 32$  pixel の領域から算出した。また、EDEN の参照入力ノードから入力する特徴量は、図 5 のように、分類対象ブロックを中心とした周囲  $64 \times 64$  pixel の領域から特徴量を算出した。この際、中心の  $32 \times 32$  pixel 部分の特徴量は使用しない。EDEN の条件判断ノードとしては、入力特徴量としきい値による 2 分岐のノードのみを使用した。使用した HLS 色特徴量とエッジに関する特徴量 14 種類と語句の対応表を表 4 に示す。このとき、各ノードの値への語句の割当て方は、4.2 節における実験と同様である。

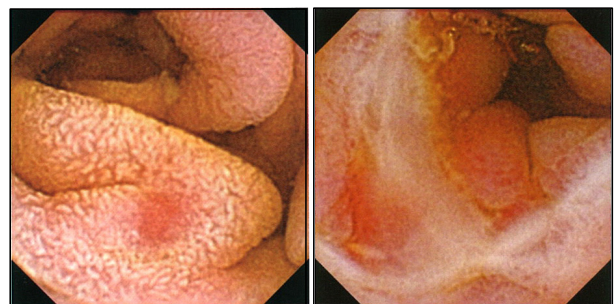


図 8 カプセル内視鏡を用いて撮影した小腸画像の例 [16]  
Fig. 8 Small intestine images taken by wireless capsule endoscopy [16].

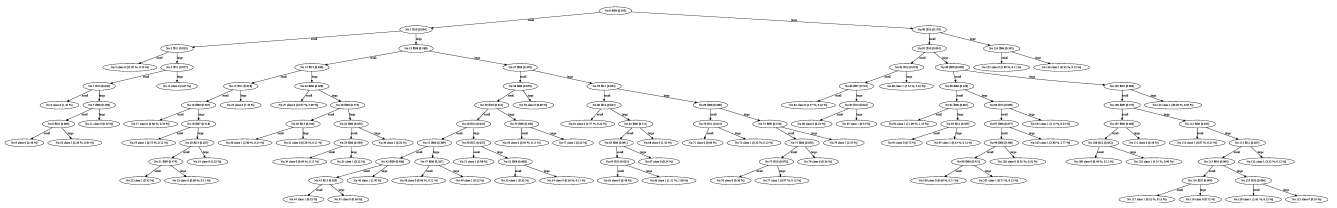


図 9 C4.5 によって獲得された小腸画像の分類構造  
Fig. 9 The C4.5 classifier for small intestine images.

表 4 小腸画像分類実験で用いた特徴量と語句の対応表

Table 4 The dictionary of word-feature pairs for small intestine image classification.

特徴量	対応する語句
H が 30° 未満または 330° 以上の画素の割合	赤色っぽい部分
H が 30° 以上 90° 未満の画素の割合	黄色っぽい部分
H が 90° 以上 150° 未満の画素の割合	緑色っぽい部分
H が 150° 以上 210° 未満の画素の割合	水色っぽい部分
H が 210° 以上 270° 未満の画素の割合	青色っぽい部分
H が 270° 以上 330° 未満の画素の割合	紫色っぽい部分
L の平均値	明るさ
S の平均値	鮮やかさ
H の標準偏差	色数
L の標準偏差	明るさの差
S の標準偏差	鮮やかさの差
L の最大値	最も白っぽい部分
L の最小値	最も黒っぽい部分
エッジ強度の最大値	輪郭

表 5 小腸画像分類の正解率と総ノード数

Table 5 The correct classification rates and the total number of nodes for small intestine image classification.

	C4.5	EDEN
正解率	88.2%	85.5%
総ノード数	127	21

また、C4.5 と EDEN による小腸画像分類の精度と構築された構造の総ノード数を表 5 に示す。これらの学習結果を用いて、分類器に対する説明文と学習データに対する説明文の生成を行った。得られた分類器の構造を図 9, 図 10 に示す。指定したパラメータは  $N = 1$ ,  $n_{max} = \max$  である。生成された説明文の例を次に示す。

- C4.5 による小腸画像の分類過程の説明文
  - 白っぽい部分があって、輪郭がはっきりしてなくて、色数がとても少ないわけではないため、異常である。
  - 白っぽい部分がないで、色数がとても少なく、鮮やかさの差がほとんどないため、正常である。
- EDEN による小腸画像の分類過程の説明文
  - ある程度鮮やかな色で、白っぽい部分があって、周囲の赤っぽい部分がある程度多くて、色数がとても少なく、明るさの差がほとんどないわけではない

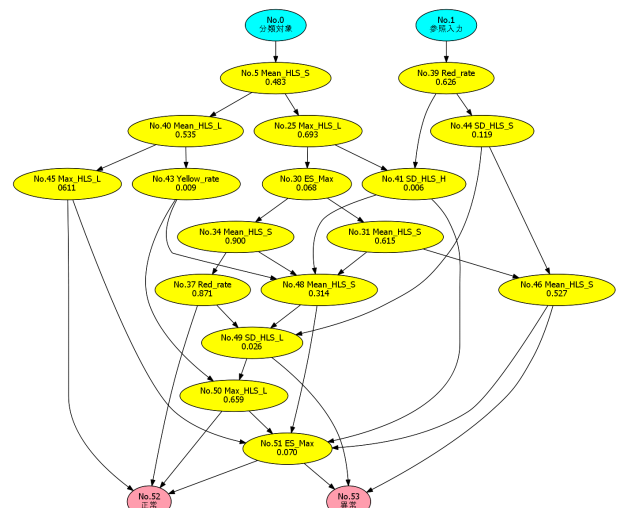


図 10 EDEN によって獲得された小腸画像の分類構造  
Fig. 10 The EDEN classifier for small intestine images.

ため、異常である。

- におい色で、暗くて、白っぽい部分がないため、正常である。
- C4.5 による小腸画像の学習データの説明文
  - 白っぽい部分があって、輪郭がはっきりしてなくて、色数がとても少ないため、異常である。
  - 白っぽい部分があって、色数がとても少なく、鮮やかさの差がほとんどないため、正常である。
- EDEN による小腸画像の学習データの説明文
  - 鮮やかな色で、白っぽい部分があって、周囲の赤っぽい部分がとても多くて、色数がとても少なく、明るさの差がほとんどないため、異常である。
  - におい色で、ある程度明るくて、白っぽい部分がないため、正常である。

こちらも一般画像分類と同様に、提案手法を適用することで説明文という新たな方法で結果の表示を行うことができた。特に、C4.5 で構築された決定木は、総ノード数が 127 の非常に複雑な構造になっており、流したデータの分類パスを追跡することや、分類器の構造を見てなぜこのような分類結果となったのかを把握することは容易ではないと考えられる。このように複雑な構造でどのように分類を行っているかが理解しにくい場合でも、提案手法を適用することで分類過程を分かりやすく表示することができてい



るといえる。しかし、まだ冗長な表現や自然でない表現が散見されるため、より自然な文章を作るために改善が必要であると考えられる。また、現在は1つのパスを1文で表現をしているが、長くなりすぎた場合は2文にするなどの改善も合わせて必要であると考えられる。

## 5. まとめと今後の課題

本論文では、あらかじめ構築された決定木および決定ネットワーク EDEN の分類過程を説明する文章を自動で生成する手法を提案した。一般画像分類と医用画像の分類を行う決定木と EDEN のネットワークを構築し、これらに対して提案手法を適用して説明文を生成する実験を行った。結果として説明分類過程の説明文を生成することができ、分類器の構造全体を提示するよりも分類過程を分かりやすく提示することができたといえる。これにより、人間が分類器が有用であるかをこれまでより容易に確認することができるようになったといえる。

今後の課題としては、言語モデルとの組合せなどによってより直感的に理解しやすい文章を生成できるよう手法を改良することや、より複雑な特徴量を使用した際にどのような文章を生成するかといったことがあげられる。また、本論文では2クラスの分類問題を対象として実験を行ったが、今後多クラスの分類問題に適用する場合、2クラスの分類に比べて分類過程を理解しにくい文章が生成される可能性がある。これは、クラス数に対し文章の長さ  $n$  を小さく設定すると、複数のクラスで同一の文章が生成される可能性が高くなるからである。そのため、今後は説明文を生成する際に重要とするノードの要素を、ルートノードから近いという条件以外に、各クラスを表すパスが分岐し、異なる文章が生成されるノードといった条件を加えるなどして多クラス分類の際にも本手法がより有効に働くよう改良することも課題である。また、現時点では提案手法を適用して得られた説明文から、人が意図していないが分類に有効に働く特徴量といった新たな知見は得られていない。しかし、提案手法を適用することで、前述のような新たな知見が単に分類器の構造を提示されるより発見しやすくなることが期待される。今後さらに実験を行い、このような点で有効性を示すことも課題としてあげられる。また、客観的な評価によって手法の有効性を評価することも今後の課題である。

## 参考文献

- [1] 中山史朗, 穂積知佐, 矢田紀子, 長尾智晴: 進化的条件判断ネットワーク EDEN による画像分類, 映像情報メディア学会誌, Vol.67, No.7, pp.J278–J285 (2013).
- [2] 増井重弘, 寺野寿郎, 関口俊宏: ファジィ論理を用いた風景画の画像理解 (特集) ファジィ画像情報処理, 日本ファジィ学会誌, Vol.6, No.5, pp.1000–1011 (1994).
- [3] Farhadi, A., Hejrati, M., Sadeghi, M.A., Young, P.,

Rashtchian, C., Hockenmaier, J. and Forsyth, D.: Every Picture Tells a Story: Generating Sentences from Images, *Computer Vision-ECCV 2010*, pp.15–29, Springer (2010).

- [4] Ushiku, Y., Harada, T. and Kuniyoshi, Y.: Understanding images with natural sentences, *Proc. the 19th ACM international conference on Multimedia*, pp.679–682, ACM (2011).
- [5] Ushiku, Y., Harada, T. and Kuniyoshi, Y.: Efficient Image Annotation for Automatic Sentence Generation, *Proc. the 20th ACM international conference on Multimedia*, pp.549–558, ACM (2012).
- [6] Srivastava, N. and Salakhutdinov, R.: Multimodal Learning with Deep Boltzman Machines, *Proc. Advances in Neural Information Processing Systems 25 (NIPS 2012)*, pp.2222–2230, Curran Associates, Inc. (2012).
- [7] Vinyals, O., Toshev, A., Bengio, S. and Erhan, D.: Show and Tell: A Neural Image Caption Generator, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015)*, IEEE (2015).
- [8] 小島篤博, 田原典枝, 田村武志, 福永邦雄: 動画像における人物行動の自然言語による説明の生成, 電子情報通信学会論文誌 D-II, Vol.J81-D-II, No.8, pp.1867–1875 (1998).
- [9] Ikuma, Y. and Iyatomi, H.: Production of the Grounds for Melanoma Classification Using Adaptive Fuzzy Inference Neural Network, *Proc. 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp.2570–2575, IEEE (2013).
- [10] 彌富 仁, 萩原将文: ファジィ推論ニューラルネットワークを用いた風景画像からの知識抽出と認識, 電子情報通信学会論文誌 D-II, Vol.J82-D-II, No.4, pp.685–693 (1999).
- [11] 寺島照雄, 萩原将文: 多段階ファジィ推論ニューラルネットワークを用いた画像認識システム, 電子情報通信学会論文誌 D, Vol.J90-D, No.6, pp.1623–1631 (2007).
- [12] Quinlan, J.R.: Induction of decision trees, *Machine learning*, Vol.1, No.1, pp.81–106 (1986).
- [13] Quinlan, J.R.: *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers Inc. (1993).
- [14] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H.: The WEKA Data Mining Software: An Update, *SIGKDD Explor. Newsl.*, Vol.11, No.1, pp.10–18 (2009).
- [15] 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).
- [16] 日比紀文, 緒方晴彦, 大塚和朗: カプセル内視鏡カラーアトラス, 診断と治療社 (2007).



崎津 実穂

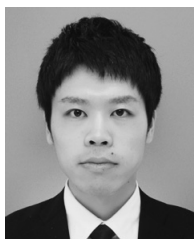
2015年横浜国立大学理工学部数物・電子情報系学科情報工学 EP 卒業。現在、同大学大学院環境情報学府情報メディア環境学専攻博士課程前期在学中。医用画像処理、パターン認識を含む知能情報学の研究に従事。電子情報通信学会, 人工知能学会各会員。



菅沼 雅徳 (学生会員)

2013年横浜国立大学工学部電子情報工学科を飛び級のため中退。現在、同大大学院環境情報学府情報メディア環境学専攻博士課程後期在学中。画像処理，パターン認識を含む知能情報学の研究に従事。情報処理学会，人工知能

学会，日本医用画像工学会各会員。



土屋 大樹 (学生会員)

2014年横浜国立大学工学部電子情報工学科卒業。現在，同大大学院環境情報学府情報メディア環境学専攻在学中。進化計算の研究に従事。電子情報通信学会，人工知能学会各会員。



長尾 智晴 (正会員)

1985年東京工業大学大学院総合理工学研究科博士課程後期中退。同年同大学助手。同大学助教授を経て2001年横浜国立大学大学院環境情報研究院教授。工学博士。画像処理，進化計算法等の知能情報処理学の研究に従事。電

子情報通信学会，人工知能学会，電気学会，進化計算学会，IEEE等各会員。