

## 機械学習手法の脳機能画像診断への応用

### Application of Machine Learning to Diagnosis based on Brain SPECT Image

酒井大輔\* , 渡辺邦彦\* , 高橋信\* , 北村正晴\*  
川島隆太\*\* , 福田寛\*\*

Daisuke Sakai\* , Kunihiro Watanabe\* , Makoto Takahashi\* , Masaharu Kitamura\*  
Ryuuta Kawashima\*\* , Hiroshi Fukuda\*\*

\*東北大学 , \*\*東北大学加齢医学研究所

\*Tohoku University , \*\*Institute of Development, Aging and Cancer, Tohoku University

キーワード : SPECT , 痴呆症 (dementia) , 帰納推論 (inductive inference) , C4.5

連絡先 : 〒980-77 仙台市青葉区荒巻字青葉 東北大学工学部大学院工学研究科 量子エネルギー工学専攻  
北村研究室 , Tel/Fax : (022) 217-7907 , E-mail : sakai@luke.qse.tohoku.ac.jp

## 1. 背景

近年PET (Positron Emission Tomography, ポジトロン断層撮像) やSPECT (Single Photon Emission Computed Tomography, シングルフォトン断層撮像) など, 血流量や代謝機能などの脳の活動を放射性物質を用いて非侵襲的に視覚化する技術がめざましく発達し, その研究や臨床での応用が盛んに行われている. 高齢化社会において大きな問題となっている痴呆症は, 大脳皮質や大脳基底核, 辺縁系の機能と密接にかかわっていると考えられており, これらの計測に適したPET, SPECTは痴呆症の診断や病態の評価において重要な役割を担っている<sup>1, 2)</sup>.

1980年代前半からPETを用いた痴呆症の研究が行われてきた結果, 痴呆の症状によって局所的な代謝や血流量の低下が見られること, そして局所

的な血流代謝の低下は症状とある程度相関することが確認されている.

脳機能画像を用いた診断では, これらの知見をもとにして症例の種類毎に応じた局所的な脳機能の低下を観測し, その相対的な脳血流分布画像のパターンによって診断が行われる. このようにPETによる痴呆症の診断は大きな可能性を秘めてはいるが, PETは線源を準備するためのサイクロトロン施設を必要とするため, 撮影のコストが非常に高い<sup>3)</sup>. そのため一般病院ではあまり普及しておらず, 国内に25施設, 世界的にも約150施設あるだけである. 一方, 本研究で対象にしているSPECTは, 装置の価格もPETに比べて安価で, 比較的容易に脳機能画像を得ることが可能である. ただし, SPECTによる画像は, その撮影原理上の空間分解能がPETに比べて低く, また, 正常例における局所脳血流量やグルコース代謝の絶対値の変動が大

きく、絶対値のみで正常と異常の区別を行うことは重症でもない限り困難であるといわれている。また、PETとSPECTでは撮影のために用いる放射性核種の種類が異なり、標識できる薬剤も異なるので、同じ脳機能画像でもPETとSPECTでは得られる知識が異なる。このためSPECTでは医師による診断においても以上の2つの点が問題となり、症例毎の診断は多くの経験と熟練を要する。この問題点を解決するために、医師の診断を支援するためのシステムの開発が望まれている。

次に、本研究で知識獲得の手法として用いる帰納推論について述べる。帰納推論とは外部教師あるいは環境から得られる事実に基づいて知識を獲得するプロセスのことであり、人工知能における機械学習手法の一つである。帰納推論の例としては、経済の動向から株価を予測したり、気象観測データから天気予報をしたりすることが挙げられる。帰納推論によって得られる仮説は、人間にとって理解しやすいという透明性 (transparency) を持つので、得られた仮説が直接適用できることが期待できる。

## 2. 目的

本研究では、脳機能画像診断の支援システムの開発を最終的な目的として、帰納推論によるSPECT画像弁別のための診断規則導出の可能性について検討する。痴呆症にも多くの種類があるが、この中の代表的な症例であるdementia (痴呆症)、depression (鬱病)、parkinson (パーキンソン病)は診断が困難なことが多く、最終的に臨床的知見が同一になるため、直接的に脳機能を可視化できるPETやSPECTでしか診断できないとされている。本研究では、以上のdementia、depression、parkinson、そしてnormal (正常)のSPECT画像から、診断のための知識を導出することを目的とした。

## 3. 解析手法

### 3.1 SPECT画像データ

サンプルとして用いたSPECT画像は、東北大医学加齢医学研究所において撮影・データベース化された、症例別の画像である。画像は $35 \times 46 \times 37$ ボクセルであり、ボクセルは一辺2mmである。このとき使用された薬剤は $^{99m}\text{Tc}$  (半減期約6時間,  $E=150\text{keV}$ )を標識した $^{99m}\text{Tc}$ -HMPAO ( $^{99m}\text{-Technesium-hexamethyl-propylenamine oxime}$ )、撮像装置は日立メディコSPECT2000Hが用いられた。サンプル画像は、前処理として正規化済みのものである。症例はnormal (サンプル数41)、dementia (サンプル数28)、depression (サンプル数29)、parkinson (軽症と重症、サンプル数45)の4つである。この分野において、症例別にデータベース化された画像としてはかなり大規模なサンプルであることを付記する。

### 3.2 帰納推論

事例を、それがもつ性質から判断して、それが属するクラスを判定する問題を分類問題という。判断あるいは診断といわれていることは、大部分この分類問題の一種である。一般に分類問題は、分類しようとする対象についてその属性値が得られたとき、その対象が属するクラス (分類カテゴリ) を判定しうるような分類規則を発見することを課題とするものである。分類問題には、データに基づく統計的アプローチと機械学習に代表される人工知能アプローチとがあるが、本研究では統計的アプローチほど多量のデータを必要とせず、かつ例外的な事例も考慮に入れた分類規則を見出しうるなど、優れた特長をもつ人工知能アプローチを採用し、その手法の一つであるC4.5を用いることにした。

C4.5は、決定木生成アルゴリズムとして知られ

ているID3<sup>4)</sup>をID3開発者であるJ.R.Quinlan自らが改良した、ID3の後継アルゴリズムである<sup>5)</sup>。エキスパート(対象分野の専門家)の経験的知識は、過去に経験した事例を元に、それに適合するルールが概念形成され、その後追加的に体験された事例情報を加えて、ルールが修正され、徐々に精緻化されたものと考えられる。C4.5はそのような事例からのルールの概念形成に関し、情報理論的なモデルを想定するものである。それは、以下のようなエキスパートの判断の特徴の観察結果をもとにしている。

「一般にエキスパートと初心者の違いは「ツボ」を心得ているところにあるといわれている。すなわち、データを見るとき、初心者は何が重要であるかわからないので、端から順に克明に見ていく。このため、時間がかかり、判断結果も曖昧で、判断理由も他人にうまく説明できない。一方、エキスパートは、キーとなる属性をいち早く見つけ、その属性値によって、次にキーとなる属性を判断し、検査する。このため時間がかからず、判断理由も簡潔明瞭である。」

このような事実をモデル化するためには、属性の重要度を測る尺度が必要になる。C4.5では、その尺度を情報理論的に定義し、クラスの分類・弁別を行っている。

### 3.3 決定木生成ツールSee5

See5はRuleQuest Research Pty Incがデータマイニングのために開発した市販の決定木生成ツールである。属性とその属性値をもつ事例の集合をSee5に入力して実行すると、それを訓練事例集合としてC4.5アルゴリズムに基づいて事例を弁別するのに最適な判別式を導出し、それを決定木形式で出力する。See5はC4.5よりも計算が高速かつ効率的に行えるようにプログラムを改良・実現しているが、決定木生成のための考え方は基本的にC4.5

と同じものである。本研究における機械学習は全てこのSee5の実行結果に基づいている。

## 4. 解析および考察

### 4.1 機能領域間の有意差の検定

研究の第一段階として、ここでは画像データに基づく症例間の弁別が可能かどうかを検証した結果について述べる。具体的には、画像から得られた特徴量に有意な差があるかどうかを症例間毎に統計検定することで弁別可能性を評価した。用いたデータは、37枚のスライスのSPECT画像の中で脳のおよそ中心である17番目のスライス画像である。ここでは最も機能領域が集中していると考えられる17番スライスのみによる弁別の可能性について検討する。

ヒトの脳は機能局在、すなわち脳の各領域にそれぞれ特定の中枢があると考えられている。そこでまず、脳画像読図の専門家により標準脳画像上の機能領域を示す図をスライス上に描いてもらい、それをスキャンして機能領域毎の特徴量を抽出するマスク行列を作成した。マスク行列は関心領域が1、その他は0となるようなバイナリ行列である。この作成を支援するためのインタフェースをFig.1に示しておく。

このようにして得られたマスク行列と画像のスライスデータのそれぞれの要素を掛け合わせることで画像中の機能領域毎にデータを抽出し、それぞれの領域に対して右脳部、左脳部の画素値平均を求め、これを特徴量とした。この量は、その特定の機能領域内の血流量に対応すると考えられる。

最初にF検定(F test)を行い、各症例間の特徴量の分散が等しいかどうかを検定したところ、予想されることではあるが分散は症例によって等しくないことがわかった。そこで次に、ウェルチのt検定(Welch's t-test)を行った。これは「デー

タ数が少なく、2群の分散が等しいと仮定できない場合」に使う検定方法である。検定は帰無仮説「2つの症例について、特徴量の母平均は一致する」について行い、有意水準は1%と5%の両方で行った。その結果はTable1のようになった。表中で、はどちらの有意水準でも帰無仮説を棄却（つまり「差がある」）、×はどちらの有意水準でも帰無仮説を採択（つまり「差がない」）、△は有意水準5%では帰無仮説を採択したが有意水準1%では棄却された（つまり「厳密に見ると差がない」）ことを表している。

この結果より、normalとdementia、depression、parkinsonのそれぞれの間には特徴量に差がある領域と差がない領域があることがわかった。特に、左右の傍シルビウス皮質という領域は全て差があるという結果が出ており、ここに注目すればnormalかそうでないかを判断できると考えられる。これは各症例毎の特徴量の平均値グラフであるFig.2からも直感的に導かれる結論である。

この結果について専門医に確認を求めたところ、傍シルビウス皮質は脳の収縮の影響を最も受けやすい領域の1つであり、この領域に大きな変化が現れているということは、臨床的知見とも一致して

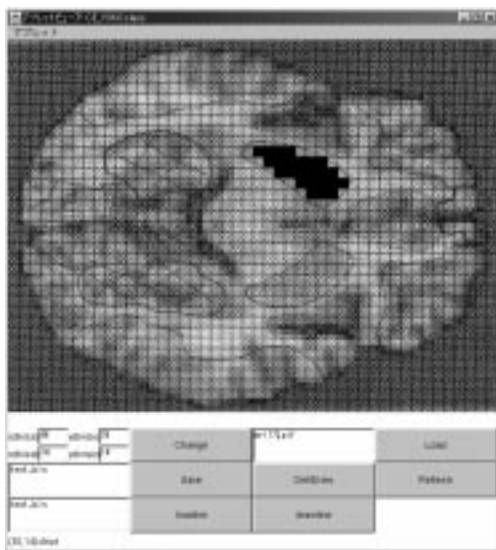


Fig. 1 特徴量抽出のためのインタフェース。

いるとのコメントを得ることができた。また、同じ領域でも右脳と左脳で結果が一致していない場合についても、脳の障害は左脳にある機能中枢の異常による場合が多く、左脳で特徴量に差があるという結果も臨床的に一致しているとのコメントを得ることができた。

以上の結果から、症例間には母平均に有意な差が見られる機能領域があることが裏付けられ、この機能領域に着目して弁別を行うことの可能性が示されたと考える。

Table 1 Normal (サンプル数41) と各症例間の特徴量の母平均に有意な差があるかの検定結果。

機能領域名		領域番号	dementia 28	depression 29	parkinson 45
基底核	右脳	M01	○	○	○
	左脳	M02	○	△	○
帯状回	右脳	M03	○	○	○
	左脳	M04	△	○	○
側頭葉内側皮質	右脳	M05	△	×	×
	左脳	M06	△	○	×
外側後頭葉	右脳	M07	×	×	×
	左脳	M08	△	×	×
舌状回	右脳	M09	×	×	×
	左脳	M10	×	×	×
前頭葉内側面皮質	右脳	M11	○	△	○
	左脳	M12	×	△	△
傍シルビウス皮質	右脳	M13	○	○	○
	左脳	M14	○	○	○
前頭前野	右脳	M15	×	×	×
	左脳	M16	×	×	△
側頭葉	右脳	M17	○	○	×
	左脳	M18	○	○	×

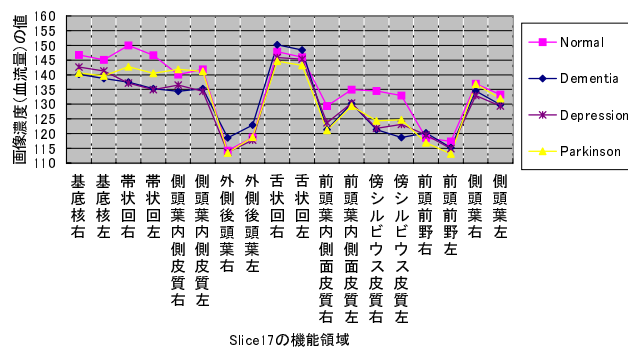


Fig. 2 各症例の特徴量の母平均のグラフ。

Decision tree:

```
M14 > 125.175:
...M01 <= 136.9355: Deme (4)
...M01 > 136.9355: Norm (37/1)
M14 <= 125.175:
...M11 <= 128: Deme (16)
...M11 > 128:
...M04 <= 135.44: Norm (3)
...M04 > 135.44:
...M13 <= 124.6667: Deme (7)
...M13 > 124.6667: Norm (2)
```

Fig. 3 normalとdementiaを弁別する決定木。

## 4.2 機能領域に基づく弁別へのC4.5適用

次に、この特徴量を属性値、機能領域を属性としてSee5を実行し、決定木を生成した。ここではnormal-dementia間の弁別について検証する。

Fig.3はSee5を実行して得られたnormalとdementiaを弁別する決定木である。決定木は左上の判別式を始点 (root, 根) としてIf-Then形式で記述されており、いくつかの判別式 (branch, 枝) を辿っていくと分類されるべきクラス (leaf, 葉) に到達する。また、leafの後の括弧内の数値は正分類された事例数および誤分類された事例数を表しているものである。

決定木中の番号はTable1の機能領域番号と対応しており、すなわち決定木の判別式に選ばれた領域は全て検定結果で「差がある」領域であった。See5は母集団の平均値で弁別を実行しているのではないので、結果が一致しているから正しいとはすぐにはいえないが、左右の傍シルビウス皮質 (領域M13, M14) が判別式に選ばれたということは、統計的事実と整合している。また、誤分類率は全体の1.4% (1/69) という高い精度を示している。

これより、See5によって判別式に選ばれる領域は決して無意味なものではなく、診断に有効な部分だけが得られたと考えられる。この結果は、See5による診断知識の獲得が十分可能性を有するものであることを示している。

## 4.3 機械的分割による症例間の差異発見

ここまでは事前知識として脳機能地図を専門医にマークしてもらい、これに基づいて特徴量の抽出を行い、有意差検定及びSee5の評価を行ってきた。ここではこの専門医による事前知識を用いずに、機械的にデータを分割・抽出し平均値をとって特徴量とし、画像データの中から弁別上有効な部分を発見することを検討する。

機械的に画像を分割し、その一つ一つを「属性」、その分割領域内の画素値の平均値を「属性値」としてSee5を実行するわけであるが、そのまま実行すれば洗練された決定木ができるであろう。しかし、脳機能画像データとしては大規模なデータとはいっても、See5のデータとしてみると扱えるサンプル数は高々数十個という少数であり、そのままSee5を実行して得られた結果はここでのデータのみで特化した結果になってしまい、本研究の目的とする「汎用性のある」診断規則獲得にはならない。

そこで本研究では、より一般性をもつ弁別上有効な部分を発見する手法を考えた。具体的には以下の手順で解析する。

1. 画像データから脳の形だけを取り出す。これは上と同様に脳のマスク行列を作ることによって行う。
2. 画像を機械的に分割する。分割の大きさは一辺1, 2, 3, 4ピクセルの正方形とし、4つの場合それぞれについて以下の2~7の作業を繰り返す。
3. 脳の部分に含まれる分割領域だけを取り出す。これは特徴量の値に脳以外の部分 (0) を影響させないためである。
4. 分割領域毎に属性、その領域内の画素値の平均値を属性値とする。

5. 弁別を行う2つの症例からランダムに半数だけ取り出し、それを訓練事例集合としてSee5を実行し、決定木を生成する。
6. 決定木の判別式に選ばれた領域を記録する。
7. 5~6の作業を30回繰り返す。
8. 分割の大きさを変えて得られた4つの結果を足しあわせる。そうすることで弁別上有効な領域の実際の大きさや分布がわかるようになり、さらに、その部分だけが際立って現れてくるようになる。

決定木生成のために用いる訓練事例としてのサンプルはランダムに全体の半数だけ用いているので、一回一回生成される決定木は当然毎回違うものであり、訓練事例集合だけに特化した決定木である。しかし得られた30回の記録を全て足しあわせれば、一回一回は特化した決定木であっても、特に何度も選ばれている領域、つまり全てに共通して差異があると考えられる領域が現れるはずであり、それこそが弁別上有効な領域、すなわち診断上一般性を持つ注目領域であると考えられる。

まず、dementiaとnormal、depressionとnormal、parkinsonとnormalを弁別する上で有効であると考えられる部分、つまり差異が見られる部分を上の方

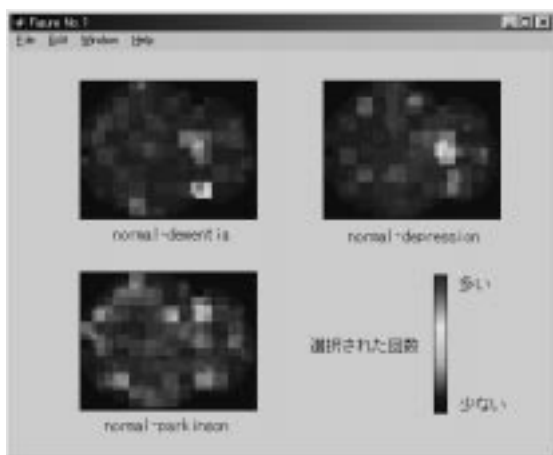


Fig. 4 17番スライス上の差異ある領域。

法で特定した。繰り返し実行回数は30回であり、その結果はFig.4のようになった。normalとdementia、depressionの場合はある部分だけが際立って多く選ばれていることがわかり、画像上ではその部分に差異があると仮説できる。一方、parkinsonの場合は全体的に選ばれた領域が分布しており、この画像(17番スライス)では差異があると思われる部分を見つけるのは困難であると考えられる。

#### 4.4 他スライスでの差異発見

上では17番スライスを用いて実行したが、ここでは17番スライスより下部にある13番スライス、上部にある22番スライスについて同様の解析を行い、検討する。繰り返し実行回数は30回で、結果はFig.5のようになった。

normal-dementiaの場合、17番スライスの時のようにある一箇所が集中して選ばれることはなかったが、13番スライスの中央部分にはまとまった分布が見られる。このようなある程度まとまった分布で

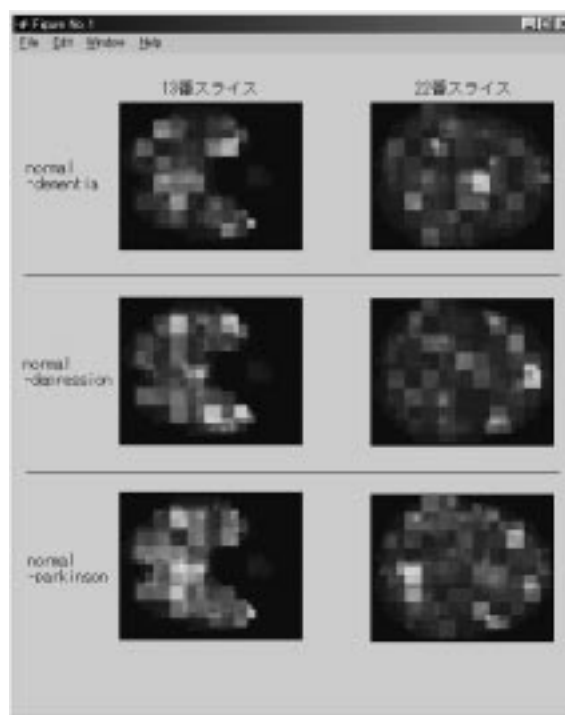


Fig. 5 13, 22番スライス上の差異ある領域。

は、その分布一帯が弁別上有効な領域であると考えられる。normal-depressionの場合、22番スライスに特に何度も選ばれる部分があることがわかる。また、normal-parkinsonの場合、13番スライスの中央部分にまとまった分布が見られるが、normal-parkinsonの場合は選ばれる領域が17番スライスの時と同様に全体に分布している傾向が見られ、これらのスライスを用いてもnormal-parkinsonの相違は見つけにくいといえる。

#### 4.5 構造化ニューラルネットワークによる結果との比較

本研究グループでは、本研究で試みている帰納推論とは違った考え方である構造化ニューラルネットワーク<sup>6,7)</sup> (Structured Neural Network) を用いたSPECT画像からの症例弁別の研究も行われている<sup>8)</sup>。ここでは以上の6つの場合それぞれで検出した差異ある部分と、構造化ニューラルネットワーク(以下SNN)を用いて得られた差異ある部分とを比較し、それぞれの結果の一貫性を検討する。

SNNによる結果は、専門医の知識に基づく脳機能地図から機能領域を抽出し、それを属性、その濃度平均値を属性値としてSNNに学習させ、学習の結果、属性にかかる重み係数の絶対値が大きい

ものが弁別上有効な領域であるという考えに基づくものである。また、SNNでは症例間の弁別を第一の目的としているので、3枚のスライス(13, 17, 22番スライス)を同時に学習させている。

SNNによって弁別上有効な領域であると判断された領域に色付けをするとFig.6のようになる(図中の黒い部分)。このSNNによる結果と、Fig.4, Fig.5を比較する。

以上より、機械的分割の結果はSNNの結果とは完全には一致していないことがわかる。これは、SNNによる解析では脳の特定の部分しか抽出しておらず、また、その抽出した部分の大きさが機械的分割の大きさや形と異なるためであると考えられる。さらに、SNNは3枚のスライスのデータを用いて解析しているが、本研究では3枚のスライスを別々にして解析をしており、そのことも結果の違いにに影響していると考えられる。機能領域の差異を調べるにはSNNが用いたような手法の方が有効であるが、医師による事前知識は人間の先入観を含んでいる恐れがあり、脳全体から差異ある部分を抽出するには機械的分割による手法のほうが有効であると考えられ、今後、SNNも同様な機械的分割による特徴量を用いて弁別に有効な部分の特定を行い、今回行った解析と比較検討を行う必要がある。

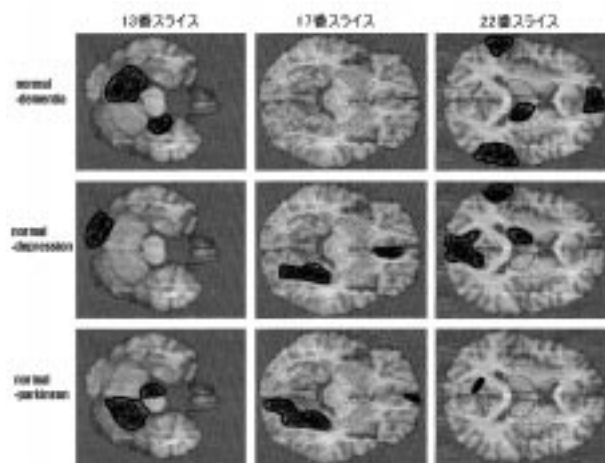


Fig. 6 SNNにより差異があると判断された機能領域。

## 5. まとめ

本研究では、C4.5という情報理論的な手法で症例を弁別し、そこから症例間で差異のある脳の部分の特定を目的として解析を行ってきた。ここでは本研究で得られた知見をまとめる。

- 1) SPECT画像から医師の事前知識に基づく機能領域毎に抽出した特徴量を統計的手法により、各症例毎の母平均に有意な差があるのかを検定し、症例によって差がある領域を特定した。

- 2) 医師の事前知識を用いず，SPECT画像を機械的に分割し，C4.5アルゴリズムに基づいた決定木生成ツールであるSee5を用いて，その分割領域のなかでどこが症例間で差異をもつかを特定した．その結果，normalとdementia，normalとdepression，normalとparkinson間の差異ある部分を特定できた．

以上のことから，帰納推論による診断規則の導出は十分可能であり，大きな可能性を持っていることを示すことができたと考える．

今後は，脳を三次元的に分割したデータに基づく診断知識の獲得および診断システムとしての知識の整理，統合を行う予定である．

## 参考文献

- 1) 「脳の科学」編集委員会：脳機能のイメージング基礎から応用まで，星和書店（1998）
- 2) 鳥塚莞爾：クリニカルPET：臨床応用のためのガイドブック，先端医療技術研究所（1997）
- 3) 久田欣一：SPECTの臨床，金原出版（1996）
- 4) Quinlan, J.R.: Induction of Decision Trees, Machine Learning (1985)
- 5) Quinlan, J.R.: C4.5: Programs for machine learning, Morgan Kaufmann Pub (1993)
- 6) R.Kozma, M.Sakuma, Y.Yokoyama, M.Kitamura: On the accuracy of mapping by neural networks trained by backpropagation with forgetting, Neurocomputing 13, 295/311 (1996)
- 7) Masumi Ishikawa: Learning of modular structured networks, Artificial Intelligence 75, 51/62 (1995)
- 8) 渡辺邦彦, 高橋信, 北村正晴: SPECT脳機能画像診断へのニューラルネットワーク適用, 計測自動制御学会東北支部会第182回研究集会, 資料番号182-6 (1999)