

帰納学習を用いた医学知識の獲得と分類*

1P-8

溝口 文雄† 大和田 勇人† 古渡 大輔† 白土 城照‡

†東京理科大学 理工学部

‡東京大学医学部眼科学教室

1 はじめに

最近、成人病検診の際に大量の眼底写真を自動処理することが強く望まれ、盛んに緑内障眼底写真の画像解析が行われている。例えば、乾氏は色の空間的变化を利用した色座標変換による血管の抽出、除去に成功している[1]。しかし、大量の眼底写真からその病状を自動診断し、緑内障の早期発見に貢献している例は見られない。それは、計算機自身に病状を学習し分類する機能を持たせる必要があるからであり、実現には多くの診断例に共通する緑内障診断知識の獲得が必要である。

一方、最近の人工知能の分野では、帰納学習[2]に関する研究が盛んである。帰納学習は、背景知識と数多くの事例を与えた時に、背景知識と正事例に矛盾しない、そして負事例を満たさない規則を導出することである。そこで、本稿では眼底画像を処理したデータに対して、この帰納学習の枠組を利用した診断知識の関係獲得のアプローチについて提案し、その例を示す。

学習に利用するデータは、眼底写真スライドと視野計で測定した視野の欠損部を示したデータである。まず、眼底写真スライドを画像処理して特徴抽出を行う。その際に乳頭の周りに広がる血管が病状の選定に邪魔をするので、色の空間的变化を利用した色座標変換による処理[1]により血管像を除去する。次に、眼底画像から帰納学習に利用する正事例を作り出すために、視神経乳頭の陥凹の形状を明確に抽出しなければならない。図2に緑内障性視神経乳頭を示す。

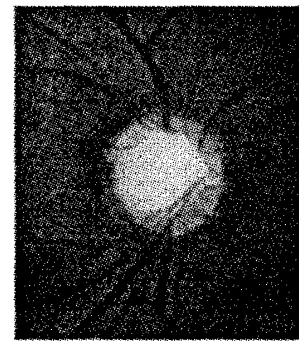


図2: 緑内障性視神経乳頭

2 知識獲得のアプローチ

本システムの処理の流れは、図1のようになる。

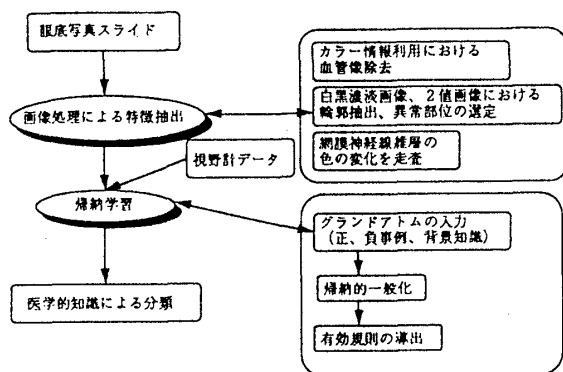


図1: 知識獲得の過程

画像データから正事例を作り出す際に次のことを考慮する。

- 視神経乳頭の輪郭を楕円近似した時のずれの大きさ
- 視神経乳頭における内円と外円の重心のずれ
- 色の急激な変化

これらの点に注意して乳頭を分割し、異常部分のエリアを限定し、病状の属性とエリアの記号で正事例を表す。

計算機に網膜神経線維層欠損の判断をさせるには、血管像除去後の乳頭以外の部分を縦、横に走査することによる濃度の変化を利用する。その際、乳頭辺縁部の異常の判断と同じように異常部分のエリアを記号で求める。

視野計からのデータは、異常部分が明確に示されているので、分割したエリアの記号に基づいてそのまま学習

*Acquisition and classification of medical knowledge using inductive learning

†Fumio Mizoguchi, Hayato Ohwada, Daisuke Kowatari, Faculty of Sci. and Tech., Science University of Tokyo

‡Siroteru Sirato, Department of Ophthalmology, Faculty of Medicine, University of Tokyo, Tokyo

データに変換する。視野の分割の基準は、解剖学的に観察した網膜神経線維を参考にする。

3 帰納学習の利用

本稿では、医療診断データが得られたという仮定のもとに帰納学習システム GOLEM[2] を用いて実際に学習を行ってみる。GOLEM は、多くの正事例間に共通する規則をホーン節によって導出する。データは、正事例、負事例、背景知識として分けて与えられる。例として、表 1 のデータ 10 例を学習させてみた結果を示す。

表中で使用する属性は、次の 5 種類とする。この 5 つについては、実際に医師が診断する際に、特に重点的に吟味するので学習のデータとして取り上げた。

- soucer (皿状の陥凹拡大)
- notch (切痕)
- nflid (網膜神経繊維層欠損)
- dh (乳頭辺縁の出血)
- vfd (視野の欠損)

正事例、負事例、背景知識は以下のように与える。

1. 正事例 (実際の視野の異常部分)
視野の異常を vfd を用いて表す。
2. 負事例 (実際の視野の異常部分以外)
視野の正常なエリアを vfd で表す。
3. 背景知識 (実際の視野の異常部分と病的部分)
正事例と各属性データすべてを与える。

記号 a,b,...,h は視神経乳頭の辺縁を反時計回りに 8 分割したエリアを、A,B,C,D,E は視野を 5 分割したエリアを示し、none は異常がないことを示す。データはすべて右目データである。獲得されたルールを、以下に示す。

$$\begin{aligned} vfd(A, c) &: \text{--soucer}(A, a). \\ vfd(A, b) &: \text{--dh}(A, B). \\ vfd(A, c) &: \text{--notch}(A, d), \text{soucer}(A, d). \\ vfd(A, b) &: \text{--soucer}(A, d), \text{notch}(A, d). \end{aligned}$$

例えば一番上のルールは、人 A の視神経乳頭の a のエリアに soucer がある時、視野 C のエリアに異常があるということを示している。この場合は、視野の異常と各病状との関係を得るために正事例を vfd(*,*) で与えている。このため、得られるルールはヘッド部が vfd(*,*) になっているが、データの与え方を変えることによって各

表 1: データ集合

番号 \ 属性	soucer	notch	nflid	dh	vfd
1	a	none	a	none	C,D
2	a	d	a,d	e	B,C
3	none	none	d	none	B
4	d,e	d	none	a	A,B,C
5	e	d	none	none	A
6	d,e	none	none	none	A
7	d	d	none	none	B,C
8	a	none	a	none	C
9	d	d	c	none	B,C
10	a	none	a	none	C

属性をヘッド部にしたルールも得られる。例えば、次のようである。

$$\text{soucer}(A, B) : \text{--notch}(A, d).$$

これらの獲得したルールを利用して、視神経乳頭の異常部分から視野の欠損部を推測できる。また、乳頭の異常と視野の異常の双方向からの診断のアプローチが可能である。

4 おわりに

本稿では、緑内障診断知識間の関係獲得を行うという目的に沿って、眼底の画像処理により特徴抽出されたデータを視野計データとともに帰納学習させるためのアプローチを述べ、実際に帰納学習を行ってみた。適当な属性データを与えることで、多くの眼底所見に共通する規則が得られることを示した。しかし、まだ診断に使えるような規則が得られたわけではない。また、画像処理部分を実験的なものにするによって、この帰納学習によるアプローチは現実的なものとなるため、さらに画像処理からのアプローチについても研究する必要がある。今後はこれらの問題点についてさらに追求していく予定である。

参考文献

- [1] 乾 成里, カラー情報を利用した眼底写真からの病変検出, 情報処理学会論文誌 Vol.34 No.5 pp.873-882,1993 5
- [2] S.Muggleton, Inductive Logic Programming, New Generation Computing, Vol.8,1991