

Title	Comparison of Neural Network and k-NN Classification Methods in Medical Image and Voice Recognitions
Author(s)	金, 應珪
Citation	大阪大学, 1993, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/38741
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉 大阪大学の博士論文について <a>〉 をご参照ください。

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	金 應 珪
博士の専攻分野の名称	博 士 (医 学)
学位記番号	第 1 0 9 1 4 号
学位授与年月日	平成 5 年 9 月 17 日
学位授与の要件	学位規則第 4 条第 2 項該当
学位論文名	Comparison of Neural Network and k-NN Classification Methods in Medical Image and Voice Recognitions (医用画像と音声の認識におけるニューラルネットワークと k-NN 分類法の比較)
論文審査委員	(主査) 教授 田村 進一 (副査) 教授 津本 忠治 教授 越智 隆弘

論 文 内 容 の 要 旨

[目 的]

膝関節 CT 画像からの膝蓋骨亜脱臼症の判定をニューラルネットワーク (BPN) で行うと共に、その際のニューラルネットワークの識別能力を従来からのノンパラメトリック学習の代表例である k-最近隣法と比較する。さらに音声データについても同様の比較を行う。これらを通してニューラルネットワークの識別能力が k-NN 法より優れていることを実験的並びに解析的に明らかにする。

[方法及び結果]

入力医用画像データは左膝屈曲角度 15° 状態の CT 画像である。画像の枚数は正常者のもの22枚、異常者 (膝蓋骨亜脱臼) のもの66枚である。この原画像は局所自動 2 値化処理により骨領域が抽出される。次に周辺部不要領域の除去と骨領域の穴埋めを行う。さらに 2 値化された画像に対して位置、大きさについて正規化を行う。形状特徴としては、正規化処理後の 2 値化画像及び 2 値化画像から抽出したエッジ画像を用いる。

入力音声データは、男11人及び女3人の学生を被験者とし、一人につき母音/a/, /i/, /u/, /e/, /o/を各10回発声させ、FFTによる特徴抽出を行った後、対数変換した。この処理のスペックはサンプリング周波数10kHz、量子化ビット数8ビット、フレーム長20ms、FFT サンプル点数128、同特徴数64、FFT ウィンドはハミング窓である。

BPN 方式での学習及び試験時のパラメータの設定に関しては膝画像の場合、その形状特徴が入力されたユニット数は 32×32 、中間層のユニット数は50、出力層のユニット数は2である。音声の場合、入力ユニット数は64、中間層のユニット数は8で、出力ユニット数は5である。

認識実験において、画像の場合には、使用したパターンは正常者22名と異常者66名のものである。この中の正常者10名、異常者15名の左膝画像を学習パターンとして使った。ただし、BPN 方式で学習するとき、これらの各パターンは同じものを2回提示し、この $25 [\text{パターン}] \times 2 [\text{回}] = 50 [\text{回}]$ の学習を1セッションとして、 $r = 35$ セッションの学習を行なった。

音声の場合には、14セットのデータから、男3人、女2人の発声データのうち、各人各母音2回合計 $P = 5 [\text{人}] \times$

5 [母音] × 2 [回] = 50 [発声] を学習データとする。ただし、BPN 方式で学習するとき、これらの各パターンは同じものを5回連続して入力した。この50 [発声] × 5 [回] = 250 [回] の学習を1セッションとして、r=200セッションの学習を行なった。次に未学習パターンを用いて識別テストを行なった。

その結果は学習パターンを除いた場合のものを見ると、膝画像認識の場合、正常者の認識率はBPN, k-NN 両法同じ (83.3%) で、異常者の認識率はBPNのほうが高い (78.4%) ことがわかった。このとき、k-最近隣法の認識率は56.8%であった。

音声認識の場合の神経回路網法 (BPN) の認識率は全体として80.0%であり、k-最近隣法の認識率は72.4%であった。全14パターンの中で4つのパターン (4/14=28.6%) だけは後者のほうが認識率が高かった。

[総括]

- 1) 画像と音声の異なる2種類のデータに対して分類を行った。このとき、BPNの認識率はk-NN法より平均9.2%の高い認識率を得た。これはYuchun Leeが手書き数字の認識の場合について得た結果と同じ傾向を示している。
- 2) 特徴数64個の音声の場合の認識率の差が7.6%であったが、特徴数1024個の膝画像の場合の認識率の差が10.8%であった。これはデータの次元数が大きい場合、k-NN分類法の認識率は相対的に低くなるというBaumの報告を裏付ける結果である。
- 3) k-NN法は多くのメモリを必要とし、分類時間もかなり長くなる。したがって、実用化には困難がある。一方、BPNの学習時間は長い、認識に要する時間はかなり短い、これは実用的なリアルタイム認識に非常に有利である。
- 4) 識別面の点からk-NNとBPNを比較するとつぎのようなことが言える。即ち、現実世界でよくあるように、2つのクラスの学習パターン自体がすでにオーバーラップしているような場合、他クラス領域へ侵入した異端学習パターンはk-NN法にておいは正当学習パターンと同等に扱われ、一定の認識領域を占有してしまう。そのため、誤りが増える。即ち、学習時の雑音・ゆらぎの影響が識別にも残ってしまう。一方、BPNにおいてはそれらに対して平滑化機能が働き、うまく識別面が設定され、異端学習パターンの影響は消える。

論文審査の結果の要旨

膝蓋骨亜脱臼は若年者でとくに問題となり、スクリーニング検査が望まれる疾患であるが、従来その定量的評価や自動診断の研究は行われていなかった。

本研究は膝間接CT画像からの膝蓋骨亜脱臼の判定をニューラルネットワーク (BPN) で行っている。入力画像は左膝屈曲角度15°状態のCT画像である。この際同時に、従来からのノンパラメトリック学習の代表例であるk-最近隣法 (k-NN) と識別能力の比較を行っている。さらに、母音音声のFFT特徴についても同様の検討を行っている。

結果は未学習パターンに関し、BPNのほうがk-NNより認識率が高く、特徴数64個の音声の場合の認識率の差が7.6%であったが、特徴数1024個の膝画像の場合の認識率の差は10.8%であるとしている。これはデータの次元数が大きい場合、k-NN分類法の認識率は相対的に低くなるという推測に一致する。またk-NNにおいては、学習時の雑音・ゆらぎの影響が識別にも残留するのに対して、BPNにおいてはそれらに対し平滑化機能が働き、うまく識別面が設定されるため、認識率が高くなるとの考察を行っている。

以上のように、本研究は初めて膝蓋骨亜脱臼症の計算機による画像診断の可能性を明らかにした点と、その際用いたBPニューラルネットワークの識別能力が従来法であるk-NN法より優れていることを明らかにした点で学位に値するものと認められる。