

F-026

Swing による視覚化を利用した肝細胞癌再発予測支援

Support for Prediction of Recurrence of Hepatocarcinoma using Visualization by Swing

中田 和宏†

西山 裕之‡

大和田 勇人‡

Kazuhiro Nakada

Hiroyuki Nishiyama

Hayato Ohwada

1. はじめに

1981年から現在まで日本人の死因第1位は癌(悪性新生物)である。癌という疾病は早期発見が最も重要であり、それが延命効果につながると言われている。そこで、本研究では再発時における早期発見の問題、中でも再発性の高い肝細胞がんについて扱う。

肝細胞がんの主な診断方法には腫瘍マーカー(癌の進行とともに増加する生体因子)と画像診断の2つがある。しかし、現状の画像診断技術では直径1cm以上の腫瘍しか検出できないという問題がある。そのため、マイクロレベルのがん細胞を早期発見するための技術が必要である。

そこで、医療機関には膨大な患者データがあるところに注目する。膨大な情報の中には重要な要素が隠れていることが多く、これを見つけだすことで新たな予防法や治療法につながると考えられる。これを実現するためのツールとして機械学習がある。実際、機械学習を用いた癌の研究[1][2]はいくつか行われている。しかし、これらの研究の大半は画像診断に機械学習を用いて、画像からの情報だけで癌の有無を判別させているものであり、マイクロレベルの癌を早期に発見する(再発予測)ために機械学習を用いている研究はあまりない。そこで、本研究では機械学習を用いてマイクロレベルの肝細胞がんの再発予測を行うことを目的とする。

2. AFP 値

肝細胞がんの診断方法の1つである腫瘍マーカーには、AFP値というものがある。このAFP値が上昇したことにより肝細胞がんを発見した事例[3]がいくつかあり、肝細胞がんを発見する際には有効であるとされている。また、再発者と非再発者のAFP値と日数の関係を示した図1を見ると、再発者(青色)と非再発者(緑色)の違いが60日目までは現れないが、120日目には違いが現れることが分かる。このようにAFP値をグラフ化することで、視覚的に再発を見分けることができる。

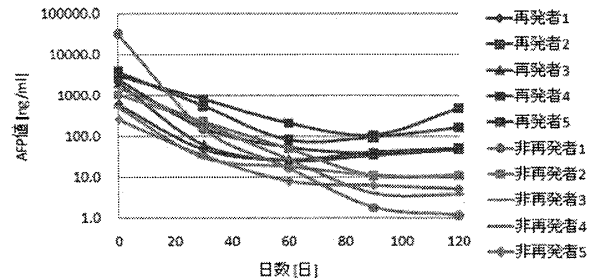


図1. 再発者と非再発者のAFP 値の推移

3. 提案手法

本研究では、早期発見に重点を置いている。そのため、術前、もしくは、手術が完了した日にわかっているデータで見分ける事が望ましい。そこで、AFP値の有効性を考慮し、本研究では、AFP値を予測し、それをグラフ化することで、肝細胞がんの再発予測を視覚的に行うことを支援するシステムを提案する。図2に提案手法の概要を示す。

3.1 AFP 値の予測

まず、肝細胞がん再発予測に有効であるAFP値の推移を、SVMを用いて予測する。(30日間隔で1500日目まで)その際、手術前の段階での予測が前提条件であるため、用いるデータは手術前のデータと予測によって得られたAFP値で、次期のAFP値を予測する。また、AFP値の予測においては「データの補間」の作業が必要である。

[データの補間]

AFP値の推移を予測する前に、まずデータの補間を行う必要がある。それは、現状では癌の検診間隔は確立されていないため、本研究で用いるデータも検診間隔が一定でないためである。そのため、日数を考慮した予測をするためには、手術をしてからの検査までの日数を固定し、固定した日数のAFP値を得る必要がある。

そこで、本研究では線形補間を用いて、検査日数を固定した際のAFP値を補間することにした。(30日間隔で固定)

3.2 グラフ化

AFP値の予測は可能になったが、情報の分野の方以外には、結果がわかりづらく、熟練したもの以外に操作が難しいと

†東京理科大学大学院理工学研究科経営工学専攻

大和田研究室 j7410629@ed.noda.tus.ac.jp

‡東京理科大学理工学部経営工学科

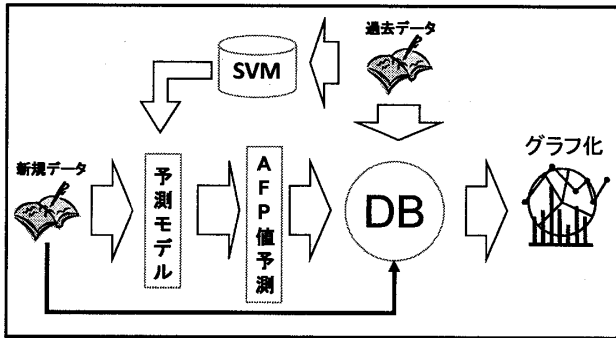


図2 提案手法の概要

考えられる。ところが、今回のシステムは、医者もしくは患者にとって優位性のあるシステムである。そのため、誰にでも操作のできるGUIを作る必要があると考える。

本研究ではGUIの中でも、非常に柔軟なカスタマイズが可能なSwingを用いる。SwingとはJ2SEが提供しているUI用のライブラリーである。

4. 実験結果

4.1 実験データ

本研究で用いたデータには計154名の患者数のデータ(最長5年分)が含まれている。再発者は100名、非再発者は54名である。再発者累計人数の日数に伴う変化を図3に示した。

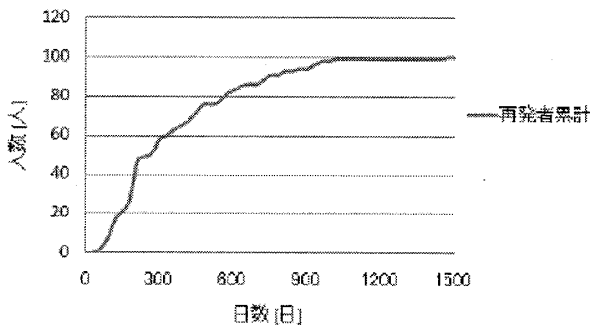


図3 再発者累計

4.2 予測結果

本研究では、用いたデータが154件と少ないため、データ数が少ない時に有効な評価法”leave-one-out法”を用いる。AFP値の予測に関しては、正しくクラス分けができたかどうかの方が重要になるため、精度(accuracy)のみを評価に用いた。表1は360日目までのAFP値予測結果である。これより、高い精度でAFP値を分類できていることが分かり、SVMでの分類は有効性を示した。また、日数が経つにつれ、精度が徐々に低くなっていることが分かる。

表1.AFP値予測結果

日数	30日目	60日目	90日目	120日目	150日目	180日目
accuracy	80.5%	74.5%	71.5%	81.4%	84.6%	84.2%
日数	210日目	240日目	270日目	300日目	330日目	360日目
accuracy	75.2%	71.9%	77.7%	75.2%	73.2%	72.3%

4.3 考察

AFP値の予測において、日数が経つにつれ徐々に精度が下がっているが、この理由は2つ考えられる。1つは、トレーニングデータ数が少ないことである。初めは154件あったデータも360日目には既に約60名が再発しているため、トレーニングデータ数が100件を下回っている。もう1つは、予測した結果を用いて予測をしていることである。予測した結果を用いて予測をするということは、長期的になればなるほど予測データの割合が増えてしまうため、精度が悪くなると考えられる。そこで、全て予測結果を使うのではなく、情報をその都度用いることで、使用する予測データ数を減らし、精度を上げることができると考える。

5. 結論と今後の展望

本論文では、機械学習SVMを用いて手術前の段階での肝細胞がんの再発予測を支援するシステムを提案した。その際に、AFP値の推移を予測し、その結果を利用してグラフ化する手法を提案した。本システムは、医者もしくは患者を対象とするシステムであり、操作性を考慮したGUIを提案する必要があるため、その点に関してはまだ不完全である。どのような情報が求められるのかを柔軟にカスタマイズする必要があるため、Swingを利用したGUIを提案している。今後の展望としては、GUIを利用した再発者のAFP推移のグラフとの比較などが考えられる。

謝辞

資料を提供して頂いた 国立がんセンター東病院 中面先生、信岡先生に感謝致します。

参考文献

- [1] J.A. Cruz and D.S. Wishart. Applications of machine learning in cancer prediction and prognosis cancer informatics.vol.2,p.59,2006
- [2] 本谷秀堅. FDG-PET/CT 画像に基づいた悪性腫瘍診断支援システム. 日本放射技術学会雑誌, Vol. 63, NO. 12, pp1460-1463, 2007
- [3] 蒔田富士雄, 鴨下憲和, 小林光伸, 三ツ木禎尚, 索引用語他. 肝細胞癌切除後の肝外転移例の検討. 日消外会誌, Vol. 32, Np. 9, pp. 2219- 2223