

PET 画像データベースの構築と検索による診断支援手法の提案

石江 哲也[†] 砂子 一徳[‡] 富井 尚志^{*}

[†]横浜国立大学大学院環境情報学府情報メディア環境学専攻

[‡]横浜国立大学工学部電子情報工学科

^{*}横浜国立大学大学院環境情報研究院

〒240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: { †d03hc006, ‡b0044076, *tommy }@ynu.ac.jp

あらまし 近年, 全身 PET(Positron Emission Tomography)画像による診断がガンの有効な診断法として注目されている。従来の読影診断時に医師が行う事柄は, 解剖学的な臓器の位置や生理的集積といった医学的な知識と画像中の特定の位置とを, 頭の中で対応づけてその所見を記述することである。その際, 画像中のどの画素群が臓器に対応しているかという意味情報は明示的に記録されないため, 読影時の診断の情報の再利用は困難である。本稿では, DB を利用した診断支援を目的とし, 医師が読影時に抽出した臓器や集積に対応する領域を明示的に定義する概念設計手法を提案する。本手法では, 医学的知識と画像データを関連付けて蓄積し, 検索が可能な, 「知識」 - 「読影時の概念的存在」 「画像データ」の3層によって表現されるスキームを導入する。また, DBMS 上に実装し, 実際に検索を行うことで提案スキームの有用性と本システムの実用性を示した。

キーワード 医療情報処理, 医用画像データベース, PET, 陽電子放射断層撮影, 診断支援

Design of PET Image Database and Retrieval for Diagnosis Support

Tetsuya ISHIE[†] Kazunori SUNAKO[‡] Takashi TOMII^{*}

[†]Department of Information Media and Environment Sciences, Graduate School of Environment and Information Sciences,
Yokohama National University

[‡]Division of Electrical and Computer Engineering, School of Engineering, Yokohama National University

^{*}Faculty of Environment and Information Sciences, Yokohama National University

79-7 Tokiwadai, Hodogaya-ku, Yokohama 240-8501 Japan

E-mail: { †d03hc006, ‡b0044076, *tommy }@ynu.ac.jp

Abstract In recent years, diagnosis by the whole body PET (Positron Emission Tomography) Images attracts attention as an effective diagnostic method of cancer. The diagnostic method, which a doctor performs at the time of the conventional interpretation of radiogram, is matching the medical knowledge of the position of anatomical internal organs and physiological accumulation and the specific position in images in the doctor's head, and describing the report. The semantic information which pixel group in images corresponds to internal organs is not clearly recorded in that case. Reuse of the diagnostic information at the time of interpretation of radiogram is difficult. In this paper, it aims at the diagnostic support which used DB. Medical knowledge and image data are associated by defining clearly the domain corresponding to internal organs and accumulation which the doctor extracted at the time of interpretation of radiogram, and accumulation and reference are possible. For that purpose, we design the schema, which has 3 layers--knowledge, notional existence at the time of interpretation of radiogram and image data. Moreover, we have implemented this schema on practical DBMS -- the usefulness of this schema and the practicality of this system were shown by actually inquiring the database .

Keyword Medical Information Processing, Medical Image Database, PET, Positron Emission Tomography, Diagnosis Support

1. はじめに

近年, 医療機器の進歩により画像を用いた検査は増加傾向にあり, 病気の兆候を発見する上で非常に有効な手段となっている。画像診断のひとつである全身 PET(Positron Emission Tomography : 陽電子放射断層撮影)画像による診断は, ガンを数ミリオーダーの初期の段階から発見できることや, ガン腫瘍の悪性・良性の判別が可能であるといったことから, 有効な診断方法として注目されている[1]。一方, PET による画像診断には固有の問題も多く, 有効な診断方法で

ありながら, そのポテンシャルを十分に発揮できていないのが現状である。例えば, PET 画像では体軸方向に約 3 mm 間隔で断層撮影が行われるため, 全身(大腿部から頭頂部まで)を撮影すると患者 1 人につきの 1 回の検査で約 300 枚の大量のスライス画像を生成する。このため, 読影診断を行う医師にかかる負担は大きい。さらに, MRI や CT が身体の中を形態的(形)に見るのに対し, PET はブドウ糖の集積等の生理的な状況を見るため, 画像から病気を読み取るとは容易に行えることではなく, 専門的知識があり経験を積んだ読影専門の医師が行う必要がある。しかし, 実際の医療現場において十分な読影医がいるわけではないため, 一度に多くの患者の読影診

断を行うことが難しい。全身 PET 画像検査の集団検診への適用が計画されているが、前述の理由により、多くの患者を診断することの妨げにもなっている。こうした PET 画像診断の現状から、コンピュータによる診断支援が強く望まれている。

ところで、従来の PET 画像の読影診断において、医師は解剖学的な臓器の位置や生理学的集積といった医学的な知識と、画像のコントラストを基に視覚で抽出した画素データ群を頭の中で人体の 3 次元モデルに結びつけることで診断を行っている。その際、どの画素データ群を抽出し、どの部位に対応させたかという情報は画像側には記録されないため、医師の知識情報と画像データは分離している。その為、今後画像データを用いた定量的診断を行ったり、自動診断システムの開発を行う際に、読影時に医師が行った診断の情報を共有したり、再利用することが困難な状況にあるといえる。

また、逆にこれらの読影診断情報を検索できるように蓄積することができれば、読影診断情報の再利用性が向上し、診断の支援となりうる。例えば多くの読影経験をもつ医師が診断した過去のデータを利用して、効果的な診断を行うことができる。このために、読影時に医師の頭の中にある解剖学や生理学、病理学といった医学的知識と PET 画像中の画素データを関連付けてデータベースに蓄積しておく必要がある。

過去に提案された画像に対する索引付け機能を実現する CBIR (Content based Image Retrieval) は、画像データから抽出した形状、色、テキストが表す物理的特徴量をもとに自動的に索引付けを行い、類似度計量系により、検索対象の類似度 (特徴量間の距離) を決定する [2]。しかし、CBIR はデータを数列の集合として扱い、索引付けを生成付けするものであり、医師が用いるようなメタデータ (例: “肝臓”) を直接与えるものではない。一般的には、画像から人間が認識しうる形態のメタデータを自動生成することは困難である [3]。そこで、的確かつ精確な検索分類を必要とする医用画像検索システムにおいては、CBIR の効率的な画像特徴量抽出機能を維持したまま、検索画像の索引付けにメタデータとしての性質を付与するため、医用画像に対して、患者情報や疾病情報を統合する機能を提供する専門知識を所与とした知識モデルベース (Conceptual data model) を導入する試みがある [4][5]。しかし、医療情報は多岐に渡るため、特に人体の生理的な状況を対象とする PET 画像を対象とした有効な手法はいまだ提案されていないのが現状である。

本稿では、すでに提案されている「高度コミュニティ空間」の 3 層スキーマ [6] を PET 画像診断に特化させた概念設計手法の導入を試みる。具体的には、医師が読影時に画像から抽出した集積や臓器に対応する領域を画像上で明示的に定義することで医学的知識と画像データを関連付けて記録することが可能な、「知識」-「読影時の概念的存在」-「画像データ」の 3 層によって表現されるスキーマの設計を行った。さらに、本スキーマの実現性と有用性を示すために、実際に DBMS 上にシステムを実装し、PET 画像と読影時に得られる領域情報、および臓器の種類など医学的知識のデータベース化を試みた。その検索結果の妥当性と有効性を評価することで本スキーマおよびデータベースの有用性を示す。以下、本論文で提案するデータベースを PET 画像データベースと呼び、そのシステムの概観を図 1 に示す。

2. 陽電子放射断層撮影

(Positron Emission Tomography : PET)

ここで PET の原理について簡単に説明する。PET は放射線フッ素 (F-18) で標識した FDG (フルオロデオキシグルコース) などのブドウ糖によく似た放射性的薬剤を注射で患者の体内に投入し、その体内分布を撮影するというものである。FDG の分布の度合いを示す定量値として「SUV」(Standardized uptake value) [7] という値が用いられる。その SUV 値が周囲に比べて高いときその領域を「集積」と呼ぶ。ガン細胞はブドウ糖を好む性質があるため、ガン腫瘍がある部位には FDG がより多く集積する、つまり SUV 値が高くなる。SUV 値を画素の濃淡として画像を構成すると、画像中ではコントラストの差となって表示される。この特性を活かし、従来の CT や MRI の画像では得られなかった病理的特長を PET 画像はより顕著に表すため、ガンの画像診断の有効な方法として近年注目を集めている。この点で PET は CT や MRI より優れているといわれている、具体的には、遠隔転移の診断、治療効果および再発の診断に有効である [1]。

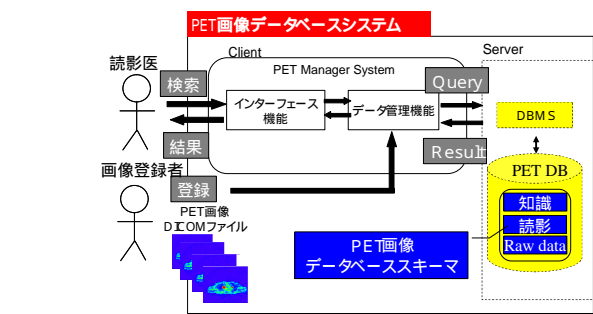


図 1. PET 画像データベースの概観



図 2. PET 画像例 (肝臓ガン)

しかし欠点として、炎症が起きている部位や生理的に FDG をより多く取り込んでしまう臓器・組織 (腎臓・膀胱・肝臓など) にはガンがない場合でも FDG はより多く集積してしまう。このため、読影にはあらかじめ生理的集積に対する知識が必要である。計算機上で PET 画像を取り扱う場合には PET 画像は DICOM 形式 [8] で取得される。機器の特性上、現在利用できる PET 画像の解像度は 128*128 (Pixel) で階調は 16bit グレースケールであり、決して高解像度であるとはいえないため、診断には知識と経験とを必要とするのが現状である。この画像データをコンソール上に表示したものやフィルムに現像したものを医師は読影診断を行っている。図 2 に肝臓ガンの PET 画像例を示す。

3. PET 画像データベースを用いた診断支援

本章では現状での読影診断について説明し、その後で我々の考える PET 画像データベースを用いた診断支援の方法について説明する。

3.1 現行の読影作業と読影診断

実際の臨床での読影時には医師は過去に得た様々な知識や画像の見方の経験と照らし合わせながら診断を進めていく。横浜市立大学医学部放射線科に協力してもらい、実際の PET 画像読影作業の様子を詳細に記録し、医師の診断の手順を調査した。

PET 画像の読影作業は以下の手順で行われている。

- (1) 一人の患者の 1 回の検査につき大腿部から頭頂部まで

- を撮影し、約300枚のスライス画像を得る。
- (2) 技師が連続するスライス画像3枚を1セットとして重ね合わせ、1枚のフィルムに焼く。その際、3枚中の画像の最大値を選択してフィルムの画像を構成する。
 - (3) 読影医が検査依頼書、カルテ情報を参照する。
 - (4) 読影医がフィルム(約100枚)全体に渡って読影し、ガンの可能性がありそうなFDGの集積を特定する。この際、必要に応じてCT画像と照らし合わせながら、フィルム中の画像と体の位置関係を対応させる。
 - (5) (4)で特定した集積について、コンソールと呼ばれる計算機の画面上にスライス画像を表示する。画面の輝度、コントラストを調節、画像の構成方法を変えて違った角度から画像を見る等して詳細に調べる。
 - (6) 読影結果に基づき所見レポートを記述し、検査依頼先の医師に提出する。

以上の読影作業において、(4)と(5)が読影診断を行っている過程に当たる。読影診断時において特徴的な事項は以下のとおりである。

- () フィルム中の画像において、SUV値の生理的集積を持つ臓器(腎臓・膀胱・肝臓など)に対応する領域を体内のランドマークとして画像の大体の「位置関係を把握する」。
- () 臓器ごとの標準的なSUV値と照らし合わせながら、それと著しく外れるSUV値の集積領域を「異常として抽出する」。その際、集積領域にROI(Region of Interest:関心領域)を設定してSUV値の最大値、平均値を算出する等を行い、定量的にガンの可能性があるか判断する。

「位置関係の把握」や「(異常)集積の抽出」には、視覚的な情報(画像のコントラストの差からFDGが集積している場所を特定する)、スライス画像の情報(連続する前後のスライス画像との比較)、解剖学的な情報(人体の臓器の位置関係)、生理学的な情報(生理的にFDGが集積する臓器)、患者の情報(病歴、別の検査情報)などさまざまな情報を組み合わせ、それらの総合的な情報に基づいて、部位や臓器ごとに異なるSUV値の標準値を適用しながら診断を進めている。

しかしながら、これらの事項は、全身PET画像検査自体がまだ新しい検査方法であることに加えて、医師個人のノウハウ・経験としてとらえられているため明示的なデータとして蓄積されていない。また、SUV値には個人差があるため、一定のしきい値による画一的なフィルタを用いた画像処理では異常を正確に診断することは困難である。従って、診断(あるいは自動診断プログラムによる機械診断)を効果的に行うためには、すでに得られたデータをいかに有効に活用し次回以降の診断に利用可能な状況で蓄積しておくかが重要な役割を担う。

3.2 データベースを用いた診断支援方法

3.1節での調査に基づいて、読影時のデータを蓄積する際に本システムにおいて読影医が行う操作を次のように想定する。

Process-1 PET画像データの登録

読影医(あるいは画像登録者)がDB serverにPET画像をスキーマに基づいて蓄積する。

Process-2 臓器、集積領域の抽出

Process-1の画像データに対し読影医(あるいは自動診断プログラム)が読影を行い、抽出した集積や臓器の領域をインタフェース上で明示的に示す(あるいは領域抽出データを渡す)。

Process-3 抽出領域に対して診断

読影医が抽出された領域をどの部位と対応させたか、その集積は生理的かどうかといった診断情報を示す。

Process-4 診断情報の蓄積

読影医がProcess-2、Process-3より示された診断情報をスキーマに基づいて画像データと関連付けDB serverに蓄積する。

以上の操作により、読影医は従来の読影診断のプロセスを変えずに、診断の記録を画像データに関連付けて蓄積することが可能である。このように正確な臨床データが蓄積され、多くの事実関係が集積してくると、「肺におけるSUV値の平均値は1.5程度で、それより大きい値を持つ患者は肺ガンの可能性が高い」という医科学的法則が確立する。これらは「事実の集合の関係」として定量的に表現される[9]。これを臨床の知識として診断時に利用できれば効果的な診断支援が行えると考えられる。

そこで本稿では、この診断プロセスと、その際に生じるデータをモデル化し、検索することのできるデータベース構築を試みる。臨床の知識を利用する方法としてデータベースに対して検索を行う。PET画像データベースでは次のような検索が可能となる。

スライス画像毎に対してしか行えなかった検索が、「集積」や「臓器」という診断時の抽象的なレベルで可能となる。

基本情報(患者IDや検査日時など)やカルテにより患者単位でしか行えなかった検索が、複数患者に対して横断的に可能となる。

今まで2次元のスライス画像上でしか行えなかったROIでのSUV値算出も、スライス画像全体を1つの3次元画像データとして扱うことにより、3次元空間領域に対して行うことが可能となる。

また今後、画像処理によるPET画像からの領域自動抽出[10]、ルールベースによるがんの自動診断方法[11]が確立した際には、本データベースをプラットフォームとして連携を図ることによりデータの再利用性を維持したまま、PET画像のがん自動診断システムを構築することが可能になると考える。

4. モデリングとスキーマ

前章での考察を基に、本章では読影診断支援のためのPET画像データベースのスキーマ設計を行う。

4.1 読影診断における情報のモデリング

我々は「知識」-「読影時の概念的な存在」「raw data」の3層によって表現されるスキーマの提案を行う。以下それぞれの層について具体的に述べる。

[医学的知識(知識層)]

まず読影を行う上で必要な共通知識である解剖学や生理学、病理学といった医学的知識を前もって明示的にしておく。例えば人間には「脳」「肺」「肝臓」などの臓器があり、その位置や形状は読影のプロセスで画像と体の部位を対応付けるランドマークの役割を果たしている。各臓器のSUV値の範囲は個人により差はあるが、「肺」なら一般的に1.5前後の値をとることが知られている。一方、SUV値が周囲と比較して高い領域である「集積」には「生理的集積」と「異常な集積」があることが知られている。

我々のスキーマでは

臓器定義：人体に存在する臓器とそのSUV値の範囲を示す

集積定義：集積の種類を示す

を知識層として記述する。

しかし実際の医学的知識はより複雑な構造を持ち、それらが互いに関連しあっている。これを整理するためにオントロジー[12][13]と呼ばれる概念を導入し、より詳細かつ更新可能な状態で知識情報を管理することができる。医療分野において独自のオントロジーを構築することを目的としたプロジェクトも存在する[14]が、本稿では詳細なオントロジー作成は今後の課題とし、今回必要最小限の医学的知識をスキーマとして表すこととした。

【画像データ(raw data 層)】

検査時に得られる一次データ、すなわち患者、検査、画像に関するデータを蓄積する。PET 画像は DICOM 形式ファイルで提供されるため、患者や検査の情報はファイルのタグ部分から読み取り、画像データとは切り離して保存する。

一般的に医用画像では、画像における1画素ごとの画素値に意味がある。従来の医用画像管理ソフト[16]などに見られる画像単位で管理する手法では、意味情報は画像データ全体の追加属性値として表現される。この方法では集積、臓器に対応する領域および画素単位での検索は、不可能である。本研究では1画素単位で扱うという単純なデータモデルを用いた。そのため本スキーマでは画像データは、各スライス画像の画素を格子点データとして3次元ブロック状に蓄積する。ここでPET 画像の解像度が128*128(Pixel)とCTやMRIに比べて大変小さいため画素単位で蓄積、操作することが可能である。3次元画像データをいかに扱うか(蓄積と検索の効率化)に関して、データモデルの改良の提案[16]もあるが、本研究では今後の課題とする。

【抽出された臓器、集積の領域(読影層)】

前述の2層を関連付ける中間子として「読影時に抽出された集積、臓器」に関する情報を記述することが可能な「読影時の概念的存在」エンティティを導入することを提案する。この要素は raw data 層においては個別の画像内における特定領域を表し、属性値として抽出された Pixel 群を持たせることができる。知識層においては読影医が診断を行った特定の臓器や集積を持たせることができる。また、SUV 値の最大値や平均値といった診断時に固有な値を持たせることができる。このことにより、3.2節で挙げたPET 画像データベースにおける3つの検索が可能となる。

具体的には、図3に示されるとおり、どの臓器なのかを示すパスや、その領域が画像内のどの Pixel 群に対応するのかを示すパスを張ることによって固有情報と一般情報、およびメディアデータ(PET 画像)との対応関係を明示的に定義する。また、特定の臓器と集積の間にパスをはることで、図3のように「肺(領域1)にふくまれる異常集積(集積a)」のような関連を示す。このような「読影時の概念的存在」エンティティと知識層、raw data 層との1つの関連は、所見レポートにおける「右肺門部に結節状の異常な集積を認める。」といった1つの記述に相当すると考えられる。

以上の3層を用いて実際に概念スキーマをUMLを用いて記述した(図4)。クラス間の関連により相互参照可能であることを表す。これにより例えば読影層を経由して raw data 層の画素と知識層の医学的知識を相互に関連させることができる、すなわち、このスキーマ構造によって3層が互いに参照可能である。現段階で、前述したように所見を3層間の1つの関連として

扱う場合には、()知識層に新たに「所見定義」を作り「読影時の概念的存在」エンティティとの間にパスを張る、()「読影時の概念的存在」エンティティ固有の属性値とするといった方法が考えられる。しかし、前者では考えうるすべての所見を記述しなければならず現実的ではない。後者では所見ごとに固有な情報を個別に蓄積できるが、複数の所見にまたがる横断的な検索が困難である。これは所見が、読影医が複数の情報を組み合わせ合わせて総合的に診断した結果導き出した情報であるためと考えられる。従って、所見は読影後に発生するものとして捉えられるが、本稿でのモデルでは読影までを扱うため、所見情報のモデル化については今後の課題とする。

4.2 PET 画像データベースの実装

本研究の目的は図4で示した3層スキーマの有用性を示すために、実際にDBMSを用いてシステムを実装し評価することにある。よって一般的な機能を持つDBMSであれば、本データベースの中核をなすDB serverの役割を実現できると考え、Microsoft社製SQLserver2000を用いて実装を行った。

【関係データベーススキーマへの変換】

本節では前節で示した概念スキーマを基に関係データベーススキーマの構築を行い、SQLserver2000上に実装する事を試みる。そこでUMLで書かれた概念スキーマを関係データベーススキーマに変換しなければならない。その際、以下のような変換ルールに従ってUMLにより記述されたスキーマを関係データベーススキーマに変換した。(1)テーブル：1つのクラスに対し1つのテーブルを作成する。(2)主キー：UMLではオブジェクトの識別子はオブジェクトが暗黙的かつ自動的に備えている基本的なプロパティであるという仮説に立っていて、オブジェクト識別を表現する標準的な手段は用意されていない。よってオブジ

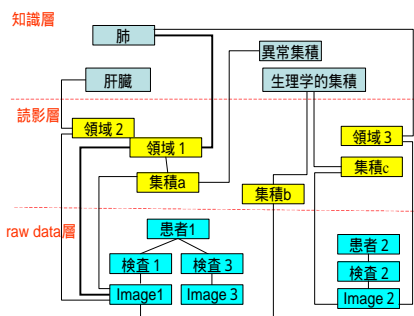


図3. 「読影時の概念的存在」エンティティの導入

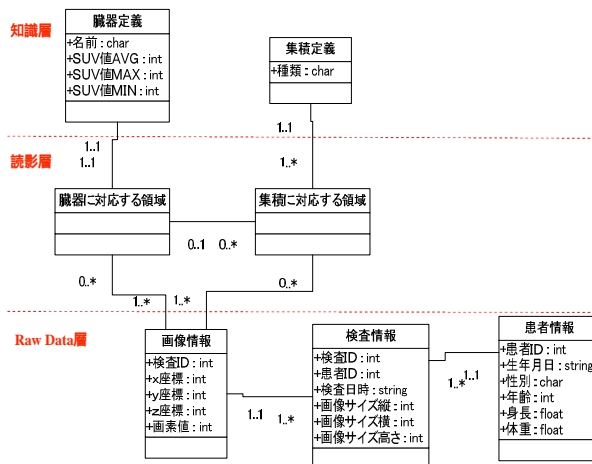


図4. PET 画像データベースの概念スキーマ

エクトを一意に決定する識別子として主キーを作成した。(3)外部キー：クラス間の関連については相互参照を可能にするため新たにテーブルを作成し(Raw Data 層におけるクラス間の関連は除く)、参照先の主キーを外部キーとした。得られた関係データベーススキーマを図5に示す。

[知識層へのインスタンスの挿入]

知識層へのインスタンスの挿入について説明する。知識層は医師の医学的知識にあたるため、読影前にあらかじめインスタンスを挿入しておく必要がある。この「Ont_Region」テーブルには解剖学的知識に基づいた臓器がインスタンスとして格納され、その臓器がとりうる SUV 値の範囲や平均値がその属性値となる。「Ont_Accumulation」テーブルのインスタンスは集積の種類として生理的集積、異常な集積が入る。

[Raw data 層へのインスタンスの挿入]

次に、raw data 層へのインスタンスの挿入について説明する。この層への挿入は Process-1 に対応する。DICOM ファイルには患者や検査、画像のメタデータがタグ情報として埋め込まれている。タグ情報から患者、検査に関するデータを読み込みそれぞれのテーブルに挿入する。この方法により患者、検査と画素データとを分離して格納する。4.1 節より1画素は関係データベースのテーブルにおける1タプルに相当する。以上のRaw Data 層へのデータ挿入操作は図1のデータ管理機能により自動化される様、実装した。

[読影層へのインスタンスの挿入]

読影層へのインスタンスの挿入は医師が実際に読影診断をしながら行うことになる。Process-4 における診断情報の蓄積では臓器、集積に対応する領域の挿入に伴い、3つのテーブルへの挿入を行うこととなる。

- 臓器：Region, Segmentation, Categorization テーブルへの挿入
 - 集積：Accumulation, Identification, Definition テーブルへの挿入
- 「肺にある集積」のような臓器と集積との関係を表す診断情報の蓄積には「Inclusion」テーブルへの挿入を行う。

領域抽出のためのコストは GUI を用いたインタフェース画面上で行うことにより軽減することが可能である。領域自動抽出の方法[11]が確立した際には自動抽出を行うことが可能と考える。画像からの特徴量抽出と索引付けのプロセスを切り離しておくことで、将来より正確な領域抽出が可能になった場合、そのアルゴリズムのみを変更するだけで検索の精度の向上が見込める[17]。

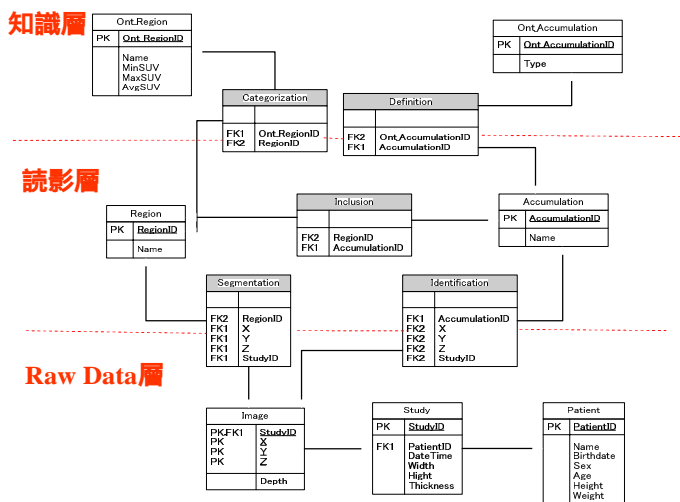


図5. 関係データベース図

5. 実験と考察

本稿で提案するスキーマを導入したシステムを実際に DBMS 上に実装し、PET 画像と読影時に得られる領域情報、および臓器の種類など医学的知識のデータベース化を試みた。その検索結果の妥当性と有効性を評価することで本スキーマおよびデータベースの実現性と有用性を示す。実験に用いたシステムは4.2節より関係データベース上に実装したPET画像データベースである。

5.1 実験環境

実験に用いた計算機環境

DBserver の実装に使用したコンピュータは Dell Computer 製 Dell Precision530 (Xeon2.4GHz デュアル構成、実装メモリ 1GB)である。

検索対象データ

知識層

解剖学的知識と集積の種類を蓄積した。表1に示す。

Raw Data 層

検索の対象とするためのメディアデータ (PET 画像) として横浜市立大学医学部放射線科から提供された断層画像 10 名分を蓄積した。10 名分の患者とその診断の情報を表2に示す。

読影層

図2のインタフェース機能により画像を表示し、医師の読影記録を基に臓器、集積に対応する領域を画面上でマウスを用いて手動で抽出、入力した。

データの内容を概念スキーマにインスタンスとして挿入した例として、患者 ID: 「0」の例を図6に示す。

表1 知識層のインスタンス例

臓器	ID	臓器名	集積	ID	種類
	R000	未定義		A000	未定義
	R001	脳		A001	生理的集積
	R002	右肺		A002	異常集積
	R003	左肺			
	R004	肝臓			
	R005	右腎臓			
	R006	左腎臓			
	R007	膀胱			

表2 検索対象診断データ

患者ID	患者名	読影の結果
0	A	正常
1	B	肝臓に異常集積
2	C	右肺に異常集積
3	D	正常
4	E	正常
5	F	正常
6	G	正常
7	H	正常
8	I	左肺に異常集積
9	J	右肺、肝臓に異常集積

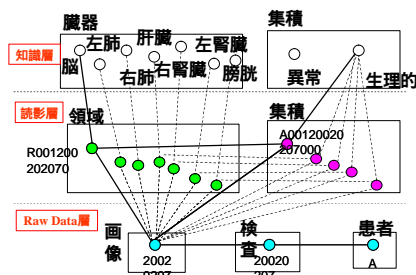


図6. 3層構造スキーマとインスタンス例

5.2. 実験1:さまざまなクエリの実現と有用性の検証

本実験ではシステム内部での検索の流れを示すと共に、検索によって本システムの有用性を実際に例として3つのQueryを行うことで検証する。今回のQueryはスキーマを熟知したDB管理者が作成した。

[実験方法]

(1) 読影により抽出された領域、集積に対する検索

Query 1:異常と診断された集積を示せ。

ここでは読影層の集積のIDを示すことがQ1の答えとなる。なぜなら図5より読影層のインスタンスが読影時に抽出された集積の存在を表しているからである。

(2) 画像データに対する集約演算的な検索

Query 2:肝臓領域のSUV値のMAX, MIN, 平均値を求めよ。

ここでは「肝臓」に対応する領域の画像データに対する集約演算を用いた検索を行う。この検索は知識層、読影層、raw data層のパスを順にたどることで可能である。

(3) 領域と集積を組み合わせた検索

Query3:異常な集積のある肝臓を示し、SUV値の平均を求めよ。

ここでは、集積と領域を組み合わせることで知識層、読影層、raw data層の全体をまたぐ検索を行う。この検索は、異常な集積のある臓器を求める、その臓器のSUV値の平均を求めるという2つの検索を組み合わせたものである。

[実験結果]

図1におけるデータ管理機能によりSQLを用いて、DBserverへQueryを実行したときのインタフェース画面を示す。Query1、2、3:実行したSQLとその検索結果を図7、8、9に示す。

実験考察]

図7の結果より「異常と診断された集積」は AccumulationID: 「A00220020129000100」、「A00220011227000200」、「A002200201224000900」、「A00220020129000901」の4つである。AccumulationIDとは図5の集積に対応する領域の存在エンティティに一意に割り振られたIDである。従って表2と比較して正しく

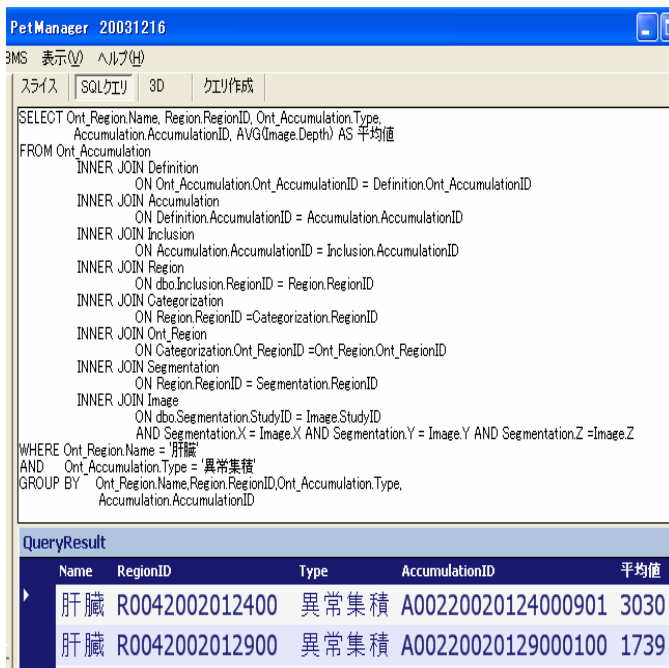


図9. Query3のSQLとその結果

検索できていることがわかる。さらにRawdata層までパスをたどって検索を行うことにより集積が画像内のどこに位置しているかを求めることができる。位置情報が得られることにより、読影時に医師が抽出した集積や臓器を表す領域を明示的に表示して確認することが可能になる。このことにより、例えば集積の見落としの防止につながる。また、熟練した医師の記録を研修医が画像から検索して学習することが可能になり、PET画像診断における診断レベルの向上にもつながる。

図8において、読影時に医師が「肝臓」に対応する領域として画像から抽出した空間領域のSUV値に対して集約演算を用いた検索結果(最大値, 最小値, 平均値)が求められた。このことにより従来は断面画像という2次元平面上でしか行えなかったROIのSUV値の



図7. Query1のSQLとその結果

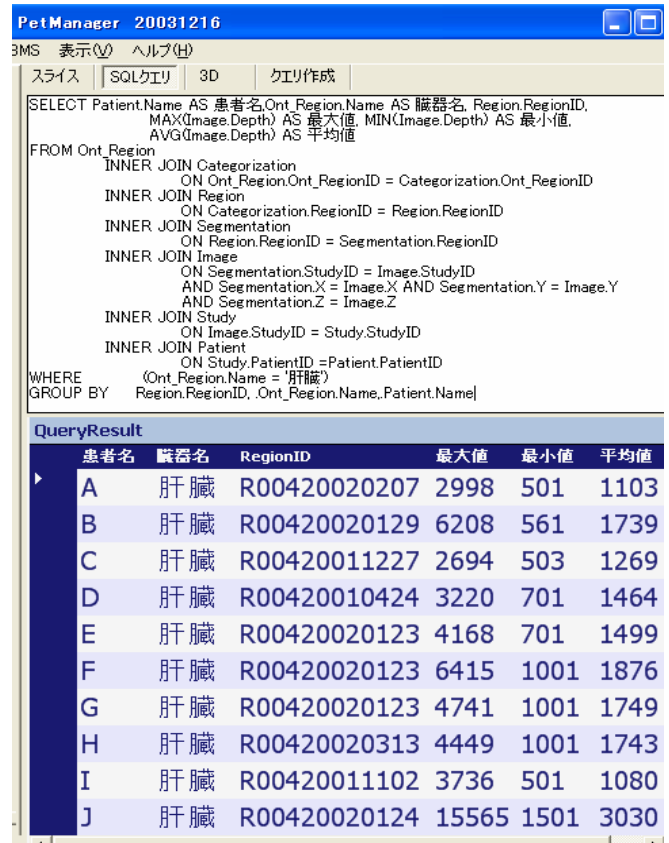


図8. Query2のSQLとその結果

算出が、本スキーマを用いたことにより空間的に求めることができることを示した。

実際に求めた値を考察すると、最大値では B, F, J の 3 名が 6208, 6415, 15565 と他に比べて高いことがわかる。平均値では J が最も高く 3030, B, F, G, H も 1739, 1876, 1749, 1743 と高い。2 章で説明した PET 画像におけるガン腫瘍の性質より、最大値、平均値共に高値である B, F, J がガンの可能性が高いと考えられる。そこで、実際の診断結果と比較を行ったところ、患者 B, J に対しては読影により発見された異常集積はガン腫瘍であることが診断された。よって B, J の 2 名は検索結果と診断結果が一致していた。

検索結果で高値が認められた患者 F は異常な集積は認められず、正常であると診断された。この理由を調べるために、実際の画像を確認したところ肝臓に対応する領域中に集積が認められた。しかし、読影時に医師は生理的なものと判断した。この集積のように SUV 値の値だけでは生理的なものか異常なものか判断が難しい画像もある。患者 F について異常肝臓以外の臓器に対しても同様の検索を行ったところ、他の正常例と比較して最大値、平均値ともに高い値を示すことが確認された。

PET 画像ではこのように個人差により SUV 値が変化するので、患者ごとに専門家（医師）の判断が必要になる場合もある。そのため、一定のしきい値による異常、正常の判断はなかなか適用が難しい場合もある。しかし、そのような特別な場合を除けば、正常な臓器の SUV 値は臓器ごとにある範囲に決まってくる。そして、その範囲は 3.1 節で述べたように「事実の集合」として定量的に表現されるので、データベースによる集約演算を用いた検索が有効であるといえる。例えば、図 8 の結果から、肝臓の正常値は平均値で 1500、最大値は 5000 以下といったことが読み取れる。この情報は、読影時に異常な集積かどうかを判断するしきい値として用いる直接的な診断支援として有効である。また、画像をコンピュータのディスプレイ上で読影する際に、注目する臓器ごとに自動で最適な輝度値に調節し、読影しやすくするといった読影作業の支援にも大変有用である。

図 9 の結果より、読影医の従来の読影診断のプロセスを変えることなく、その記録から異常な集積のある肝臓（RegionID：R0042002012400, R00420002900）について、その集積（AccumulationID:A00220020124000901, A00220020129000100）を検索し、集約演算を行うことで定量的な診断の検証を行うことが可能であることを示している。本スキーマにおいて診断の情報と画像とが結び付いているので病状など条件を指定した検索を行うことができ、さらにその条件を満たす複数の診断記録に対し横断的に検索することができる。このことにより、2 章での述べた PET 画像による検査の利点をさらに活かすことが可能となる。たとえば同一患者の再発の診断や、治療効果の確認では、患者の ID を条件に、検査時間順に結果を表示すれば、本スキーマにより画像中での集積の位置が明確なので、経時的差分画像により、集積の大きさ・範囲・集積の具合の変化を定量的に知ることができ、これまでより効率的に比較することが可能になる。

5.3 実験 2:クエリーの実行時間の検証

実験 2 では検索対象の患者数が増えた際にも、提案している PET 画像スキーマによる検索が有効であることを、実際に検索を行うことで検証する。

本実験で用いる実験データとして患者の PET 画像と診断データが 10, 50, 100 人分ずつ蓄積された 3 種類のデータベース（DB-1, DB-2, DB-3）を用意した。DB-1 では実験 1 で使用した表 2 のデータを利用した。DB-2, DB-3 のデータは表 2 のデータの患者 ID, 検査 ID, 領域抽出したデータの ID を新しく振り直すことで、模擬の診断データを作成して、データベースに蓄積した。以上の模擬

診断データを残り 9 人のデータに対しても作成し使用した。本実験では、データ規模が増加したときの検索を検証することが目的なので、模擬データに対して検索を行った結果により有効性を示すことが可能であると考えた。

[実験方法]

(1) 領域 ID を指定した検索

この検索では、ある特定の画像から抽出された領域に対して検索を行う。

Query1-1: RegionID: 'R004200201290001' を指定して、その領域の SUV 値の最大値を求めよ。（同様に平均値についても Query1-2 で検索を行った。）

(2) 患者 ID を指定しての検索

この検索ではある特定の患者の臓器に注目して検索を行う（今回は患者 B の肝臓）。

Query2-1: PatientID: '0001' の肝臓に対応する領域の SUV 値の最大値を求めよ。（同様に平均値についても Query2-2 で検索を行った。）

(3) DB に蓄積されている患者全体に対しての検索

この検索ではデータベースに蓄積されている全患者が対象になる。

Query3-1: 肝臓に対応する領域の SUV 値の患者ごとの最大値を求めよ（同様に平均値についても Query3-2 で検索を行った。）

(4) 条件により対象を絞り込んで検索

この検索ではデータベース内の情報を条件により絞り込んで検索を行う。表 2 より DB-1, 2, 3 において

正常な人: 80%

肝臓に異常集積のある人: 20%

の割合で含まれている。

Query4-1: 正常な人の肝臓に対応する領域の患者ごとの SUV 値の最大値を求めよ。（同様に平均値についても Query4-2 で検索を行った。）

Query4-3: 肝臓に異常集積がある患者の肝臓の SUV 値の患者ごとの最大値を求めよ。（同様に平均値についても Query4-4 で検索を行った。）

[実験結果]

以下の実験結果は DBserver 上で検索を実行してから、結果が返ってくるまでの時間を検索処理時間として計測したものである。結果を表 3 に示す。

表 3. Query の結果

対象データベース	DB-1 (10件) [sec]	DB-2 (60件) [sec]	DB-3 (100件) [sec]
Query1-1 (MAX)	3	8	18
Query1-2 (AVG)	3	8	18
Query2-1 (MAX)	15	20	24
Query2-2 (AVG)	15	20	24
Query3-1 (MAX)	19	68	141
Query3-2 (AVG)	19	70	143
Query4-1 (MAX)	17	58	129
Query4-2 (AVG)	18	52	123
Query4-3 (MAX)	7	17	35
Query4-4 (AVG)	7	17	35

[実験考察]

各 Query の実行時間から本データベースにおける検索が実現可能であることを示した。また各患者の平均や最大値の結果からさらに、患者全体に渡っての平均、最大値を求めることも可能である。これにより3章で述べた「臨床の知識」を検索により利用可能であることを示した。

ただし、実際に使用することを考慮すると高々100件のデータで最大134秒という時間は実用的ではない。この原因として1人あたり画像が300枚というデータ量が直接影響していると考えられる。画像データの取り扱いについては4.1節でも述べたとおり今後改良すべき課題の一つである。

また実験に用いたSQLを見るとJOINが多用されている。よってSQLチューニングによる検索の高速化の余地があると考えられる。将来的に想定されるデータ量は、例えば、2004年夏に稼働予定の新横浜PETセンターでは1日200人、年間7万2000人にPET検査を実施する計画なので、それだけのデータ量に対応できるように今後改良が必要である。

5.4 検索結果の出力方法

実験1では、検索の有効性を確認するために結果のみ確認できれば十分と考え、検索の結果は表形式で表示した。今後、PET画像データベースシステムの機能拡張として、結果をDICOM形式のファイルとして出力することも予定している。例えば、異常な集積を含む肺領域を検索した結果、抽出した肺領域の画素データのみが書き込まれたDICOMファイルを生成することが可能である。それを出力して3D Viewer[18]などを用いて表示すれば対象領域だけが浮かび上がったCG画像としてみることができ、医師が直感的により理解しやすい形での診断データの再利用が可能である。

6. まとめと今後の課題

本稿では、DBを利用した診断支援を目的とし、医師が読影時に抽出した臓器や集積に対応する領域を明示的に定義することで医学的知識と画像データを関連付けて蓄積、検索が可能で、「知識」-「読影時の概念的な存在」、「画像データ」の3層によって表現されるスキーマの設計を行った。また、DBMS上に実装し、実際に検索を行った。その結果の妥当性と有効性を評価することで提案スキーマの有効性と本システムの実用性を示した。

今後の課題としては、

読影診断後の確定診断により異常とされた集積が本当に悪性腫瘍だったかどうかまでを追跡してデータベースに蓄積するためのスキーマの拡張

SQLを知らない医師でも検索を行えるようにするための検索インタフェースの導入

などがあげられる。また、今後は画像処理による自動領域抽出、ルールベースによる自動診断システムなどとの連携を図っていく。

謝辞 本研究は、横浜市立大学・横浜国立大学共同研究の一部である。研究遂行にあたり横浜市立大学医学部放射線医学教室の井上登美夫教授よりPET画像データの提供と多くのアドバイスを頂いた。また、鈴木昌子医師に読影方法の調査において協力して頂いた。加えて、横浜国立大学環境情報研究院の有澤博教授をはじめ、共同研究の多くのスタッフの方々の協力を得た。ここに記して謝意を表す。

尚、本研究の一部は文部科学省科学研究費補助金(課題番号40313473)の支援による。

文 献

- [1] 社団法人日本アイソトープ協会 HP
<http://www.jrias.or.jp/jrias/>
- [2] Bjorn Johansson, "A Survey on: Content Based Search in Image Databases",
<http://www.isy.liu.se/cvl/Projects/VISIT-bjojo/>
- [3] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, , "Content-Based Image Retrieval at the End of the Early Years", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 12, DECEMBER 2000
- [4] W.W.Chu, I.T.Ieong, and R.K.Taira, "A Semantic Modeling Approach for Image Retrieval by Content", VLDB J, 3:445-477, 1994.
- [5] Daniel Keysers, Jorg Dahmen, Hermann Ney, Berthold B. Wein, Thomas M. Lehmann, "Statistical framework for model-based image retrieval in medical applications", Journal of Electronic Imaging 12(1), pp59-68, January 2003.
- [6] 富井尚志, "高度コミュニティ空間の構成方式", 情報処理学会研究報告, DBS-131(II), pp. 33-40, Aug, 2003.
- [7] 山口慶一郎, 鷺野谷利幸, 伊藤正敏, "正常画像と生理的変動", 画像診断, Vol. 23, No. 10, pp. 1129 - 1141, October 2003.
- [8] JIRA/DICOM 委員会 HP
<http://www.jfcr.or.jp/DICOM/index.html>
- [9] 清水哲男, 医科学のためのシステム論, 放射線科学, Vol. 45, No. 1, pp13-23, 2002.
- [10] 中野雄太, 長尾智晴, "3次元医用画像処理の自動生成システム 3D-ACTIT", 医学情報処理システム研究会 MIPS2003 論文集, 横浜国大, 2003
- [11] 橋本麗, 荒井淳, 有澤博, 田村直良, 鈴木昌子, 井上登美夫, "PET画像を用いたガン自動診断システムにおける医学知識のルール化", 信学技報, DE2003-112, pp7-12, October, 2003
- [12] JOHN F. SOWA, "Knowledge Representation: Logical, Philosophical, and Computational Foundations", Brooks/Cole, 2000
- [13] Gruber, Thomas R. "A translation approach to portable ontology specifications. In Knowledge Acquisition", vol. 5, pp199-220, 1993.
- [14] Unified Medical Language System(UMLS)
<http://www.nlm.nih.gov/research/umls/>
- [15] 医用画像管理ソフトウェア Aconomy
<http://www.risecorp.co.jp/acconomy.html>
- [16] Tianqiu Wang, Simone Santini, Amarnath Gupta, "An Interpolated Volume Model for Database", ER2003, LNCS 2813, pp335-348, 2003.
- [17] 北本朝展, 高木幹雄, "類似画像検索システム構築のフレームワークとしての階層モデル", 電子情報通信学会技術報告, Vol.PRMU2002-159, pp.25-32, 1997
- [18] Amira
<http://www.amiravis.com/>