

PET/CT 画像より抽出した腫瘍候補からの 偽陽性除去について

澤田 好秀^{*1}, 本谷 秀堅^{*1}, 呉 勁^{*2}, 武田 徹^{*3}, 渡邊 順久^{*4}

要旨

本稿では FDG-PET/CT 画像からの腫瘍抽出精度を向上させるための、真の腫瘍と偽陽性の識別器について報告する。我々は全身の FDG-PET/CT 画像より腫瘍を抽出して医師に提示するための診断支援システムを開発している。このシステムは、まず PET 画像を二値化して腫瘍候補を抽出する。この際、偽陽性が多く含まれることを許容し、真の腫瘍の取りこぼしを避ける。次に、これら偽陽性を除去するための識別器を適用する。この識別器は、腫瘍候補画像の集合より学習により構築する。本発表では、この識別器の汎化誤差と学習データの関係について調査した結果を報告する。

キーワード：FDG-PET/CT, 診断支援, 腫瘍検出, 偽陽性除去, 汎化誤差

1. はじめに

近年、様々なモダリティの中で、癌の診断などを目的として PET/CT 画像が広く利用されるようになってきた[1]。PET 画像は患者の体内に投与した薬剤の 3 次元分布を表したものであり、体内の機能情報を表している。特に薬剤として FDG を利用する FDG-PET 画像

は、体内の各位置における糖代謝情報を表す。FDG-PET 画像においては糖代謝の活発な臓器や腫瘍が高輝度に撮影され、腫瘍の抽出に有用である。一方、CT 画像は X 線の透過率の 3 次元分布を表したものであり、体内の形態情報の取得に有用である。PET/CT 装置はこれらの画像をほぼ同時に撮影するため、姿勢や臓器などの位置ずれの小さな融合画像を容易に得ることができる。このため、文献[2]のような画像の位置あわせを必要としない。このことは、体内の部位や位置ごとの糖代謝情報を得る上で大きな利点である。Fig.1 に FDG-PET/CT 画像を示す。

PET/CT 装置の普及に伴い、PET/CT 画像を対象とした計算機による診断支援(Computer Aided Diagnosis: CAD)システムの必要性が高まりつつある[3], [4], [5]。我々は全身の FDG-PET(以下 PET)画像と CT 画像より腫瘍候補を抽出し、それぞれの位置と大きさを医師に提示するシステムを開発してきた[6], [7]。先に

*1 名古屋工業大学

[〒466-8555 名古屋市昭和区御器所町]

e-mail: sawada@hilbert.elcom.nitech.ac.jp

e-mail: hontani@nitech.ac.jp

*2 千葉大学社会精神保健教育研究センター

[〒260-8670 千葉市中央区亥鼻 1-8-1]

*3 筑波大学臨床医学系

[〒305-8575 茨城県つくば市天王台 1-1-1]

*4 山形済生病院

[〒990-8545 山形県山形市沖町 79-1]

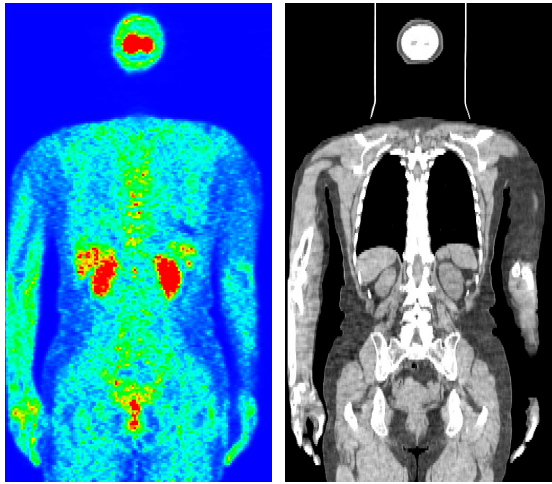


Fig. 1 An example of original PET/CT images.
Left: FDG-PET image, Right: CT image.

述べたとおり、腫瘍では糖代謝が活発であるため、PET 画像中で高輝度な領域として撮影される。但し、PET 画像中の高輝度な領域の全てが腫瘍であるわけではなく、腫瘍でない領域でも糖代謝が活発な領域は高輝度に撮影される。そこで我々のシステムは、まず PET 画像を二値化することにより腫瘍の候補領域を抽出し、次に腫瘍候補中の偽陽性を除去することにより悪性腫瘍のみを抽出する。この偽陽性除去のためにカーネル Fisher 判別分析 (Kernel Fisher Discriminant Analysis: KFDA)[8], [9]を利用して識別器を構築した。

KFDA は、通常の Fisher 判別分析にカーネル法を適用したものであり、非線形の特徴空間において学習データのクラス内・クラス間分散比が最大となるような識別空間を求めることができる。本稿では、学習データを腫瘍候補画像にそれぞれ位置と大きさの情報を付与して作成する。

一般に、対象の画像を元画像から切り出して学習データを作成する際、切り出した画像の大きさを正規化する必要がある。この正規化後の画像の大きさは、学習データの次元数を決定するパラメータであり識別器の性能を左右する。実際本稿でも報告するとおり、同一の学習データ集合を用い、同じ学習アルゴリズムを採用しても、画像を正規化する大き

さを変化させると、識別性能も変化する。正規化後の画像の大きさに依存して識別性能が変化する場合、その識別性能が最良となる大きさに画像を正規化すべきである。

対象がある確率分布に従って生起するとき、その確率分布に基づいて評価される識別器の性能は汎化性能と呼ばれる。しかし一般には対象が従う確率分布は未知であるため、汎化性能は評価できない。このため、識別器の性能は学習データ(の一部)に基づいて、例えば Leave-One-Out (LOO)法などにより評価せざるをえない。このとき、評価された識別性能と汎化性能との間には差がありうる。識別器を構築する際の画像は、本来、汎化性能を最良にする大きさに正規化すべきである。

本稿では、KFDA で構築した識別器の汎化性能を評価するために、KFDA で求めた識別空間の uniform stability[10]を用いた。具体的には、KFDA で求めた識別空間の uniform stability β_u を近似的に求め、その値を用いて汎化誤差を評価し、約 1000 個の腫瘍候補画像を用いて画像の各次元における汎化性能を比較した。

本稿の構成は以下の通りである。まず第 2 節では、PET/CT 画像からの腫瘍候補の自動抽出について述べる。次に第 3 節では、学習データの生成、第 4 節では識別器の構築方法を述べる。第 5 節では、汎化誤差の評価方法について述べる。そして第 6 節では評価実験を行い、第 7 節にて結論を述べる。

2. 腫瘍候補抽出

本節では、PET 画像を二値化し腫瘍候補を抽出する手法について説明する[6], [7]。

Fig.2 に示すように PET 画像の輝度値分布は患者や部位ごとに大きく異なっている。このため、まず CT 画像から体内を各部位に分割し、次に PET 画像より部位ごとに閾値を自動決定する。体内の分割は CT 画像を用いて行い、頭部、頸部、肺、胸腔、胸部、腹部、下腹部、脚部に分割した。また、確率アトラス[11]を用いて肝臓、腎臓、脾臓を抽出した。

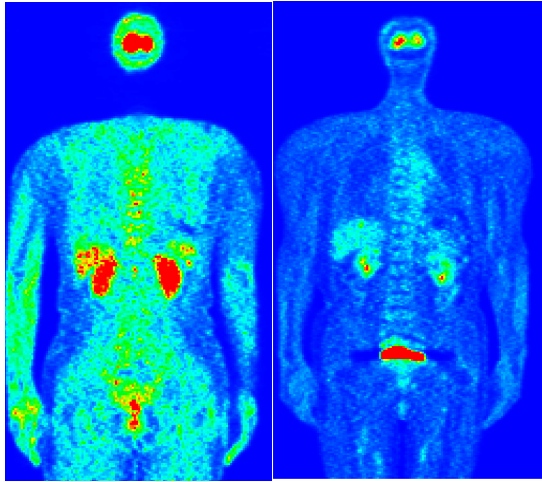


Fig. 2 An example of original PET images.

These two images are different by patients and parts.

次に各部位ごとに PET 値のヒストグラムを求める。そして、二値化した結果得られる腫瘍候補領域のボクセル数の期待値が 1 ボクセル程度となるように、ヒストグラムに基づき閾値 T_{PET} を決定する。この方法を実際に 99 個の腫瘍を含む 11 症例に適用したところ、すべての腫瘍を抽出することに成功した。一方、偽陽性は正常患者 32 症例に適用したところ、1 症例あたり平均 150 個抽出された。

3. 学習データの生成

腫瘍候補に含まれる多数の偽陽性を除去するために識別器を構築する。偽陽性の多くは腹部から抽出された。そこで本稿では腹部より抽出された腫瘍候補を悪性腫瘍とそれ以外に分離する識別器を構築する。

腫瘍候補は PET 画像の二値化により抽出した。腫瘍候補を腫瘍と偽陽性として識別するためには PET 画像の個々の画素値以外の情報が必要である。本稿では、PET 画像パターン (Fig.3), 腫瘍候補の位置, 大きさ情報のそれぞれを含む学習データを生成する。

1) 腫瘍候補画像の切り出しと正規化

前節でも述べたとおり、PET 画像は患者や部位ごとに輝度値分布が異なるため、PET 画

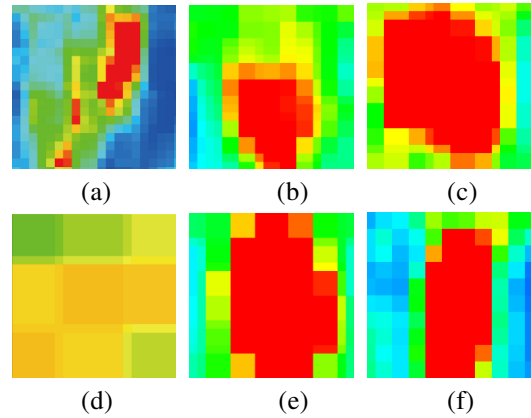


Fig. 3 Examples of the extracted tumor candidates in PET images. (a)~(c): true malignant tumors, (d)~(e): false positives.

像パターンを生成する前に、PET 値を正規化する必要がある。そこで、PET 値 $I(x,y,z)$ を次式に従い正規化する。

$$\tilde{I}(x, y, z) = \frac{I(x, y, z)}{T_{PET}}. \quad (1)$$

ここで、 T_{PET} は腫瘍候補抽出の際に部位ごとに定めた閾値である。この正規化により腫瘍候補内部の画素値が 1 以上になり、それ以外の領域が 1 未満となる。

次に変換された PET 値を用いて PET 画像パターンを生成する。PET 画像パターンは上記の変換を行った PET 画像から、重心を中心とした一辺 r_i の外接立方体を用いて切り出したものである。ここで、学習を行う際は、各学習データの次元をそろえる必要があるため、立方体の大きさを正規化する。正規化後の一辺の長さを L とおくと、PET 画像パターンは L^3 次元となる。

2) 腫瘍候補の位置と大きさ情報の表現

腫瘍か否かの識別には、腫瘍の大きさや体内における位置の情報も有用であると考えられる。しかし、腫瘍候補画像を元画像から切り出すと、位置情報が失われ、大きさの正規化を行うと、大きさ情報が失われる。このため、大きさ、位置情報は別途記録しておく必要がある。そこで本手法では、正規化前の立方体の

一辺の長さ r_i を腫瘍候補の大きさ情報として表す。また、体内における腫瘍候補の位置をシステムが元画像より自動的に抽出した肝臓、左右の腎臓、膀胱の4つの臓器からの相対位置で表す。肝臓、右の腎臓、左の腎臓、膀胱のそれぞれの重心から i 番目の腫瘍候補の重心までの距離をそれぞれ d_{1i} , d_{2i} , d_{3i} , d_{4i} で表すとき、腫瘍候補の位置を以下の4次元ベクトルで表す。

$$\mathbf{d}_i = \frac{[d_{1i}, d_{2i}, d_{3i}, d_{4i}]^T}{\sqrt{d_{1i}^2 + d_{2i}^2 + d_{3i}^2 + d_{4i}^2}}. \quad (2)$$

3) 学習データの生成

学習データは、各腫瘍候補画像から得られる上記のデータを並べたものとする。すなわち、(1)画素値を正規化したPET画像 $\mathbf{x}_{i, \text{PET}}$, (2)位置 \mathbf{d}_i , (3)大きさ r_i である。但し、これら4つのデータはモダリティが異なり、値のオーダーも異なる。そこで、 $\|\mathbf{x}_{i, \text{PET}}\|$, \mathbf{d}_i , r_i のそれぞれの分散の値を全腫瘍候補から計算し、 σ_{PET}^2 , σ_d^2 , σ_r^2 を求め、次式のように各モダリティの情報を正規化する。

$$\mathbf{x}_i = \left[\frac{\mathbf{x}_{i, \text{PET}}^T}{\sigma_{\text{PET}}}, \frac{\mathbf{d}_i^T}{\sigma_d}, \frac{r_i}{\sigma_r} \right]^T. \quad (3)$$

4. 識別器の構築

腫瘍候補を真の悪性腫瘍と偽陽性との識別する識別器を学習により構築する。学習にはカーネル Fisher 判別分析 (KFDA) を採用する [8], [9]。

悪性腫瘍のクラスを ω_1 で表し、偽陽性のクラスを ω_{-1} で表す。それぞれのクラスの要素数を m_1 , m_{-1} で表し、全学習データを $m = m_1 + m_{-1}$ で表す。また、学習データを \mathbf{x}_i ($i, j = 1, 2, \dots, m$) で表し、 \mathbf{x}_i のクラスが ω_1 (ω_{-1}) のとき、それぞれラベルを $y_i = 1$ (-1) と定める。これら m 個の学習データに基づき、非線形高次元特徴空間におけるクラス内・クラス間分散比が最大となるような識別空間を、KFDAを用い

て計算する。具体的には以下の分離度 J を最大にするような m 次元ベクトル \mathbf{a} を求める。

$$J(\mathbf{a}) = \frac{\mathbf{a}^T \mathbf{M} \mathbf{a}}{\mathbf{a}^T \mathbf{N} \mathbf{a}}. \quad (4)$$

但し、 \mathbf{M} と \mathbf{N} は双方 $m \times m$ の行列であり、前者はクラス間分散、後者はクラス内分散に関わり、それぞれ次式で計算される。

$$\mathbf{M} = \boldsymbol{\mu} \boldsymbol{\mu}^T, \mathbf{N} = \mathbf{K} \mathbf{K}^T - \sum_{k=1, -1} m_k \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T. \quad (5)$$

ここで、 \mathbf{K} は $m \times m$ の行列 $\mathbf{K} = (k(\mathbf{x}_i, \mathbf{x}_j))$ である。 $k(\mathbf{x}_i, \mathbf{x}_j)$ はカーネル関数であり、本稿では以下のガウスクーネルを使用する。

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right). \quad (6)$$

但し、 σ^2 はガウスクーネルの分散を表す。また、式(5)上の $\boldsymbol{\mu}$ は $\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ であり、それぞれのクラスの学習データの平均どうしの隔たりを表すベクトルである。 $\boldsymbol{\mu}_1$ と $\boldsymbol{\mu}_2$ はそれぞれ悪性腫瘍の学習データと偽陽性の学習データの平均に関連するベクトルであり、 $\boldsymbol{\mu}_k = [\mu_k^1, \mu_k^2, \dots, \mu_k^m]^T$ ($k = 1, -1$) と表すとき、各成分 μ_k^j は次式で定義される。

$$\mu_k^j = \frac{1}{m_k} \sum_{\mathbf{x}_i \in \omega_k} k(\mathbf{x}_i, \mathbf{x}_j). \quad (7)$$

式(4)を最大にする m 次元ベクトル \mathbf{a} は $m \times m$ の行列 $\mathbf{N}^{-1} \mathbf{M}$ の最大固有値に対応する固有ベクトルとして計算できる。但し、このままでは m 個の学習データから m 次元の分散を求める必要があるため、不良設定である。そこで文献[8]のように、 \mathbf{N} を

$$\mathbf{N}_\lambda = \mathbf{N} + \lambda \mathbf{I}, \quad (8)$$

へと変更する。 λ は正則化パラメータ、 \mathbf{I} は単位行列である。ここまで記した手法により計算されたベクトル \mathbf{a} の各成分を $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_m]^T$ で表す。また閾値を β で表す。このとき、識別器は新たに入力されたデータ \mathbf{x} が悪性腫瘍か否かを判定するために、

$$\sum_{i=1}^m k(\mathbf{x}, \mathbf{x}_i) \alpha_i + \beta, \quad (9)$$

の符号を求め、正なら腫瘍、負なら偽陽性と判定する。

ところで、式(4)を最大にする α は、以下の目的関数 E を最小にする α_M とスケールを除いて一致する[9].

$$\begin{aligned} E(\alpha_M, \beta_M) &= E_0(\alpha_M, \beta_M) + \frac{1}{2} \lambda \alpha_M^T \alpha_M, \\ E_0(\alpha_M, \beta_M) &= \frac{1}{2} (\mathbf{y} - \mathbf{K}^T \alpha_M - \beta_M \mathbf{u})^T (\mathbf{y} - \mathbf{K}^T \alpha_M - \beta_M \mathbf{u}). \end{aligned} \quad (10)$$

ここで m 次元ベクトル $\mathbf{y} = [y'_1, y'_2, \dots, y'_m]^T$ は、各学習データに新たに付したラベル y'_i を並べたものであり、 \mathbf{x}_i のクラスが ω_1 (ω_{-1}) のとき、 $y'_i = m/m_1$ ($-m/m_2$) である。また、 \mathbf{u} は m 次元ベクトル $\mathbf{u} = [1, 1, \dots, 1]^T$ である。式(10)を最小にする α_M , β_M は以下のように求めることができる。

$$\begin{bmatrix} \alpha_M \\ \beta_M \end{bmatrix} = \begin{bmatrix} \mathbf{K}\mathbf{K}^T + \lambda \mathbf{I} & \mathbf{K}\mathbf{u} \\ (\mathbf{K}\mathbf{u})^T & m \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{K}\mathbf{y} \\ \mathbf{u}^T \mathbf{y} \end{bmatrix}. \quad (11)$$

先に述べたとおり式(11)で得られる α_M は式(4)を最大にする α と方向が一致する。また、式(11)で得られる β_M は

$$\beta_M = -\alpha^T \frac{(m_1 \boldsymbol{\mu}_1 + m_2 \boldsymbol{\mu}_2)}{m}. \quad (12)$$

となる。すなわち、腫瘍と偽陽性の各クラスの重心間を $1/m_1:1/m_{-1}$ に内分する位置が閾値となる。医用画像においては、腫瘍のデータより偽陽性のデータの方が圧倒的に収集しやすい。このため、 $m_{-1} \gg m_1$ となる場合がほとんどであり、 $1/m_1:1/m_{-1}$ で内分すると、閾値は偽陽性からのマージンが小さくなり、悪性腫瘍からのマージンが大きくなる。このことは、偽陰性をより強く抑えるという方針と合致する。このため、 β_M を閾値としてそのまま採用しても良い。そこで本稿では式(10)を最小にすることにより識別器を構築する。

5. 汎化性能評価

識別対象がある確率分布に従うときに、同じ確率分布に従う特定の学習データ集合 S を用いて構築した識別器がその確率分布に対して示す誤りのことを汎化誤差という。しかし、一般的に確率分布は未知であるため、汎化誤差を求めることはできない。本節では、式(10)を最小化することにより構築された識別器の安定性と汎化誤差について述べる。

1) Uniform Stability と汎化誤差

識別器の性能の評価は、学習データと同じデータでテストをして経験損失を求めたり、Leave-One-Out (LOO)法のような交差法により誤差を求めたりする。前者で得られた誤り率を

$$R_{\text{emp}} = \frac{1}{m} \sum_{i=1}^m l(\alpha_S, y^i), \quad (13)$$

で表し、LOO法で得られた誤り率を

$$R_{\text{loo}} = \frac{1}{m} \sum_{i=1}^m l(\alpha_{S_{-i}}, y^i), \quad (14)$$

で表す。また、汎化誤差を R で表す。但し、

$$l(\alpha_S, y^i) = \frac{1}{2} \left(y^i - \sum_{i=1}^m k(\mathbf{x}, \mathbf{x}_i) \alpha_i - \beta \right)^2, \quad (15)$$

であり、正しいラベルと識別器が予測する値との差の二乗を表す。 m 個の学習データの集合を S で表すとき、 α_S は m 個のデータ全てを用いて得られた識別器 (式(10)を最小にする α , β) を表し、 $\alpha_{S_{-i}}$ は m 個のデータのうち i 番目のデータを取り除いて得られる識別器を表す。 R と R_{emp} や R_{loo} との差の上界は uniform stability β_u で求めることができ[10], 以降では、本稿の枠組みに即して uniform stability について説明する。

式(10)を最小にする識別器の uniform stability β_u は次式を満たす最小値である。

$$\begin{aligned} \forall S \in Z^m, \forall i \in \{1, \dots, m\}, \\ \|l(\alpha_S, \cdot) - l(\alpha_{S^{(i)}, \cdot})\|_\infty \leq \beta_u. \end{aligned} \quad (16)$$

uniform stability β_u は上式を見るとわかるとおり、学習データのひとつを取り除くことが識別器に与える影響の大きさの最大値である。

ところで、誤差を評価する関数 $l(\alpha_S, y)$ が次式を満たすとき、その識別器は σ -admissible と呼ばれる。

$$|l(\alpha, y') - l(\alpha', y')| \leq \sigma_\alpha |(\alpha^T \mathbf{x} + \beta) - (\alpha'^T \mathbf{x} + \beta')|. \quad (17)$$

この σ の値は m 次元ベクトル α により定まる量であり、この σ の値を求めることができれば、式(10)を最小化することで得られた識別器の uniform stability β_u は次式で与えられる。

$$\beta_u \leq \frac{\sigma_\alpha^2}{\lambda}. \quad (18)$$

uniform stability β_u を用いることにより、学習データに対しての誤り率 R_{emp} 、LOO 法で得られた誤り率 R_{loo} と汎化誤差 R との間に、確率 $1 - \delta$ ($\delta \in (0, 1)$) で次式が成立する。 β_u が大きいほど、誤り率と汎化誤差との差が高確率で大きくなることがわかる。

$$R \leq R_{\text{emp}} + 2\beta_u + \left(4m\beta_u + \sqrt{\frac{mB}{\lambda}}\right) \sqrt{\frac{\log(1/\delta)}{2m}}, \quad (19)$$

$$R \leq R_{\text{loo}} + \beta_u + \left(4m\beta_u + \sqrt{\frac{mB}{\lambda}}\right) \sqrt{\frac{\log(1/\delta)}{2m}}. \quad (20)$$

但し、 B はデータ \mathbf{x} を α に投影したときとらうる最大値で、具体的には

$$B = \arg \max_{\mathbf{x}} \left| \sum_{i=1}^m k(\mathbf{x}, \mathbf{x}_i) \alpha_i + \beta \right|, \quad (21)$$

である。以降では、 B を近似的に求める方法を示す。 $k(\mathbf{x}, \mathbf{x}_i)$ はガウスカネルであり、 \mathbf{x} が \mathbf{x}_i の近傍に位置するときに最大値 1 に近い値をとり、それ以外では 0 に近い値になる。本稿ではガウスカネルのスケール σ と比べて学習データ \mathbf{x}_i がそれほど密に分布していな

い状況を仮定する。このとき、 $\sum_i^m k(\mathbf{x}, \mathbf{x}_i) \alpha_i$

の値は、 \mathbf{x} が例えば \mathbf{x}_i の近傍にあるときに α_i に近い値となり、学習データのいずれからも遠い場所には 0 に近い値になる。このため、 \mathbf{x} を変化させたときの $\sum_i^m k(\mathbf{x}, \mathbf{x}_i) \alpha_i$ の最大値

は、ベクトル α の要素 m 個のうち最大値で近似できる。この最大値を α_{i^*} で表す。このとき式(21)は、

$$B \approx \alpha_{i^*} + \beta, \quad (22)$$

と近似できる。

2) Uniform Stability の近似評価

式(19)、(20)の右辺を求めるには、式(17)の σ_α の値を求める必要がある。前節の B を用いると、 $\sigma_\alpha = 2B$ であることが容易に導出でき、これを式(18)に代入すると次式を得ることができる。

$$\beta_u \leq \frac{4B^2}{\lambda}. \quad (23)$$

上記の β_u を式(19)、(20)に代入することにより、 R と R_{emp} 、 R_{loo} との差の上界を計算することができる。

3) パラメータの設定

識別器を構築する際、式(6)中のスケール σ の値と、式(8)中の λ の値を決定する必要がある。これらパラメータを決定する方法として、本稿では KFDA を用いた効率的な LOO 法である Efficient Leave-One-Out (ELOO)法を採用した[12]。ELOO 法とは、 R_{loo} を計算する際に必要な $l(\alpha_{S^{(i)}, y'})$ を以下のように与える方法のことである。

$$l(\alpha_{S^{(i)}, y'}) = \left(\frac{y'_i - \sum_j^m k(\mathbf{x}, \mathbf{x}_j) - \beta}{1 - H_{ii}} \right)^2. \quad (24)$$

但し、 $H = Z[Z^T Z + \lambda I]Z$ 、 $Z = [K \mathbf{u}]$ である。

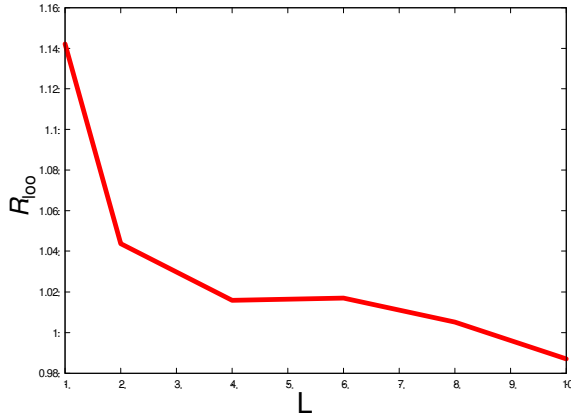


Fig 4. The change of R_{100} for the change of the dimension of the training data.

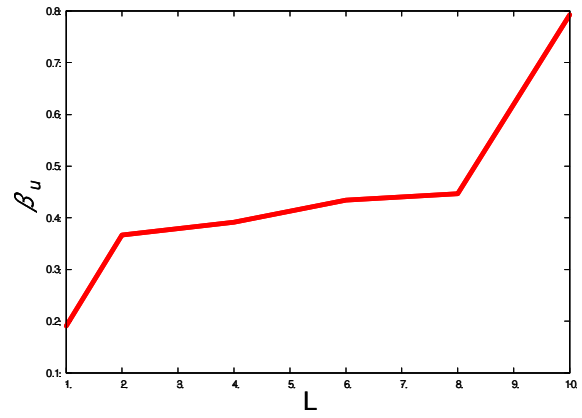


Fig 5. The change of uniform stability β_u for the change of the dimension of the training data.

Table 1. The evaluated performance of the classifiers constructed from different dimensional training data.

L	False negatives	False positives
1	5.12% (2/39)	25.8% (258/1000)
2	0% (0/39)	23.7% (237/1000)
4	0% (0/39)	28.6% (286/1000)
6	0% (0/39)	26.0% (260/1000)
8	0% (0/39)	25.8% (258/1000)
10	0% (0/39)	14.6% (146/1000)

通常、LOO法は学習データ分だけ繰り返すため、学習データが多くなればなるほど、多大なコストがかかる。しかし、ELOO法を用いると、学習データの数がどれだけ増加しても識別器を一度構築するのと同じコストでLOOを評価できる。これがELOO法の利点である。

第6節ではこの方法を用いて、 σ^2 の値を0.5から10.0まで0.5刻みで変化させ、一方 λ の値を 10^{-5} から10.0まで10倍ずつ変化させ、すべての組み合わせの中から最も式(14)の値を最小にする組み合わせを採用した。

6. 評価実験

第2節で述べた手法により抽出した腫瘍候補の集合に基づいて、KFDAにより識別器を構築する。本稿で実験に用いる腫瘍候補の集合は、悪性腫瘍39個と偽陽性1000個から構成されている。いずれもPET画像の二値化により抽出されたものであり、前者は癌患者8症例から抽出された腫瘍候補のうち共著者の放射線科医が悪性腫瘍と判断したものであり、後者は健康な患者244症例から抽出された腫瘍候補である。それぞれの腫瘍候補の画像より、第3節で述べた手続きより学習データを作成する。このとき、画像の大きさは一辺 L の立方体に正規化する。この L の値の変化に対する識別器の性能の変化の評価が、本実験の目的である。

学習に利用した腫瘍データの数に比べて偽陽性データの数が多く、データ数について不均衡である。これは病変データの入手がそうではないデータの入手と比べて困難だからであり、このような状況はCADシステムを実現する際に頻繁に起こることである。クラス間のデータ数の不均衡は慎重に考慮すべき問題ではあるが、本稿ではこの問題については扱わない。腫瘍39個と偽陽性1000個の学習データに対してKFDAを適用し、性能と汎化誤差の評価を行った。

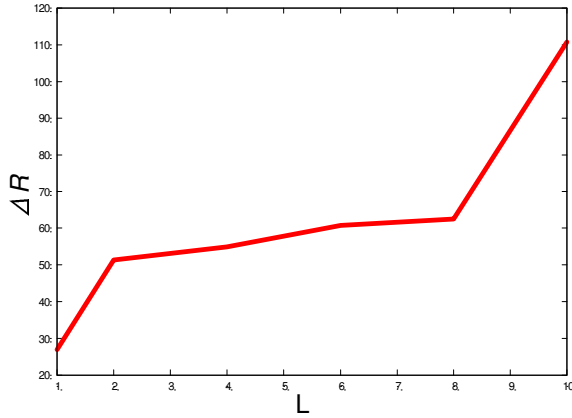


Fig 6. The change ΔR for the change of the dimension of the training data.

1) 安定性評価

まず、学習データの次元の変化に対する識別器の性能の変化を評価した。Fig.4 にその結果を示す。但し、 $L=1$ のときは、PET の平均輝度を用いている。また、学習に利用したデータ数は上述のとおり腫瘍 39 個、偽陽性 1000 個である。学習の評価には ELOO 法で得られた R_{100} の値を用いた。図を見ると分かる通り、同一の学習データに対して同一のアルゴリズムを用いても、画像を正規化する大きさ L を大きくする(データ空間の次元を高くすると)、 R_{100} は減少した。Fig.4 に示した結果に従って画像を正規化する大きさ L を定めるとき、 L の値は大きいほど良い(偽陽性の除去率が高くなる)と判断されることになる。また、学習データの次元の変化に対する識別率の変化を示した表を Table 1. に示す(学習データに対する識別率)。この表から、 $L=1$ のとき最も識別性能が悪く、 $L=10$ のときに最も識別性能が良いということがわかる。

次に、データ空間の次元数の変化に対する uniform stability β_u の変化を評価した。 L の変化に対する式(23)の右辺の値の変化を Fig.5 に示す。図をみるとわかる通り、次元数の増加とともに uniform stability の値が大きくなっている。すなわち、データ空間の次元数の増加とともに、KFDA で計算される識別器 α の安定性が低くなっており、学習データの若

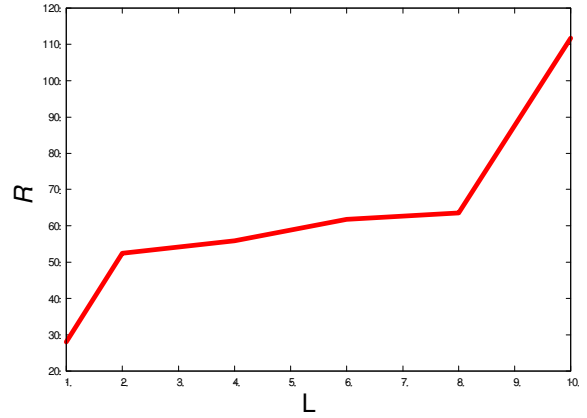


Fig 7. The change of the generalization error R for the change of the dimension of the training data.

干の変化に対してより大きく変化するようになってきていることを示す。

2) 汎化性能評価

データ空間の次元の変化に対する汎化誤差の変化を式(18)に基づいて評価した。結果を Fig.6, 7 に示す。但し、 $\delta=0.1$ とした。Fig.6 は式(18)の右辺の第 2 項以降

$$\Delta R = \beta_u + \left(4m\beta_u + \sqrt{\frac{mB}{\lambda}} \right) \sqrt{\frac{\log(1/\delta)}{2m}}, \quad (25)$$

を表す。 ΔR は式(14)で評価した R_{100} と汎化誤差 R の差の上限を示している。図より次元が上昇するにつれて ΔR は増加していくことがわかる。

また、Fig.7 は $R=R_{100} + \Delta R$ を表している。この図から、 $L=1$ のときが最も汎化性能が良いことがわかる。つまり、PET 画像を用いた識別では、PET の平均輝度などの低次元な情報のみを使用している可能性がある。また、Fig.4, 6, 7 から ΔR の値の方が R_{100} の値よりも大きいことがわかる。つまり、 ΔR の方が R_{100} よりも汎化誤差 R に与える影響が大きい。このことは、今回の手法が汎化誤差を悲観的に見積もっているということを示している。

7. おわりに

本稿では PET/CT 画像から腫瘍を抽出する際の偽陽性を除去するために用いる識別器について報告した。画像から抽出される腫瘍候補の画像の大きさは様々である。画像パターンを直接学習に用いるアピアランスベースト手法を採用する場合、識別器を構築する際や実際に識別器を適用する際に画像の大きさを正規化する必要がある。この正規化する大きさの決定には自由度があり、基本的に識別性能が最大となるように定めるべきである。但し、Leave-One-Out 法により評価される性能は、いずれも汎化性能との間に差がある。学習に用いるデータの集合が同一であるならば、多くの場合、画像を正規化する大きさが大きいほど高い性能を示す傾向がある。本稿では、識別器の uniform stability が、画像を正規化する大きさが大きいほど大きく不安定になることを実験的に確認し、著者らが構築した識別器に対して汎化性能の変化も評価した。また、各次元に対する取りこぼし、誤抽出の数も求めた。

また、今回の手法は実験結果からもわかるとおり、悲観的な汎化誤差を求めている。そのため、今後はより厳密な評価方法を使用する必要がある。

文 献

- [1] 西村恒彦, 佐治英郎, 飯田英博: クリニカル PET 一望千里, MEDICAL VIEW, 東京, 2004, 8-17
- [2] 福田宙也, 清水昭伸, 小畑秀文, 他: 臓器情報に基づく医用画像の位置合わせ法の開発. PRMU2007-4, 107: 17-22, 2007
- [3] 多次元医用画像の知的診断支援第一回シンポジウム論文集. 1, 2005
- [4] 鳥脇純一郎: わが国における CAD 研究の歴史と将来. インナービジョン 19: 5-9, 2004
- [5] Yongkang Nie, Qiang Li, Feng Li et al: Integrating PET and CT information to Improve Diagnostic Accuracy for Lung Nodules: A Semiautomatic Computer-Aided Method, The Journal of Nuclear Medicine 47:1075-1080, 2006
- [6] 新田修平, 本谷秀堅, 深見忠典, 他: 画素値の確率分布に基づく PET/CT 画像中の腫瘍領域の抽出. MEDICAL IMAGING TECHNOLOGY, 24: 409-418, 2006
- [7] Shuhei Nitta, Yoshihide Sawada, Hidekata Hontani, et al: Malignant tumor detection from PET/CT images for CAD system. International Journal of Computer Assisted Radiology and Surgery, 2:520, 2007
- [8] Sebastian Mika, Gunnar Ratsch, Jason Weston, et al: FISHER DISCRIMINANT ANALYSIS WITH KERNELS, IEEE press 41-48, 1999
- [9] Jinhua Xu, Xuegong Zhang, Yanada Li: Kernel MSE Algorithm: A Unified Framework for KFD, LS-SVM and KRR. Proceedings of the International Joint Conference on Neural Networks, 2: 1486-1491, 2001
- [10] Oliver Bousquet, Andre Elisself: Stability and Generalization. The Journal of Machine Learning Research, 2:499-526, 2002
- [11] 清水昭伸, 桜井博紀, 柳田友尚, 他: 人体の電子アトラスに基づく 3 次元腹部 CT 像からの複数臓器の抽出とその性能評価. 信学技報 MI2005-15, 105: 7-12, 2005
- [12] Gavin C. Cawley, Nicola L. C. Talbot: Efficient Leave-One-Out Cross Validation of Kernel Fisher Discriminant Classifiers. Pattern Recognition 36:2585-2592

Rejection of false positives included in tumor candidates detected from PET/CT images

Yoshihide SAWADA^{*1}, Hidekata HONTANI^{*1}, Jin WU^{*2}, Tohoru TAKEDA^{*3}, Yoriyisa WATANABE^{*4}

*1 Nagoya Institute of Technology

*2 Center for Forensic Mental Health, Chiba University

*3 Tsukuba University

*4 Yamagata Saisei Hospital

In this paper, we reported about a classifier separating the true malignant tumors and the false positives. We developed the Computer Aided Diagnosis (CAD) system to show the true malignant tumors using FDG-PET/CT images. This system, firstly, extracts tumor candidates by binarizing FDG-PET images. Secondly, we constructed the classifier that rejects those false positives by means of the kernel Fisher discriminant analysis. This classifier is constructed by means of tumor candidate's images. In this article, we reported the results about the relationship between the generalization and the learning data.

Key words: FDG-PET/CT, Computer Aided Diagnosis, tumor extraction, false positive rejection, generalization error