

CCD スキャナで取り込んだじん肺写真に対する ランダムツリーを用いた型分類

Categorization of Pneumoconiosis Radiographs Using Random Trees Obtained with a CCD Scanner

中村 宗広† 阿部 孝司† 南 昌秀‡ 井口 信和†
Munehiro Nakamura Koji Abe Masahide Minami Nobukazu Iguchi

1. まえがき

じん肺の診断結果は熟練された医師間でも異なることが多く、計算機によりじん肺を症状レベル別に分類し、医師の診断を支援する方法の確立が求められている。従来手法では、テクスチャ解析による方法[1][2]から個々の粒状影を抽出する方法[3][4]に発展したが、粒状影の抽出精度が完全ではなく、かつ、胸部 X 線写真を鮮明に画像化するスキャナが高価で臨床の現場では実用化に至っていない。本稿では、ILO と厚生労働省のそれぞれが提供しているじん肺標準写真と健常の写真を市販 CCD スキャナで画像化し、ランダムツリーにより症状レベル別に型分類することを提案する。

2. じん肺画像（けい肺）の型分類

じん肺の症状であるけい肺の診断基準は、粒状影の数とサイズによって 0 型（健常）と 1~4 型（異常）の 5 段階に分類され、型が上がるにつれ重度になる。4 型は目視で明確に判別できるため、本研究では 0~3 型までを対象とする。

図 1 に示すように、ドラムスキャナを用いた場合と比較して、CCD スキャナで取り込んだ胸部 X 線像は肋骨影や粒状影が不鮮明となる。この画像に対して、粒状影を抽出する方法[3][4]や画像品質の影響を大きく受けるテクスチャ解析による方法[1][2]の適用は困難であり、この画像の特徴に着目した別の手法が必要である。

CCD スキャナで取り込んだ画像を目視観察すると、粒状影は 0 型には無く、1 型では肋骨影内部（以下、肋骨）に少数、2 型では肋骨に多数と肋間に少数ある。3 型では、肋骨と肋間に多数の粒状影がある。また、粒状影は肺の内側ほど不鮮明であることがわかる。

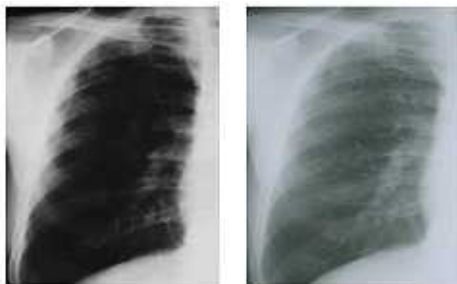


図 1 異なるスキャナで取り込んだ胸部 X 線像
(左: CCD スキャナ, 右: ドラムスキャナ)

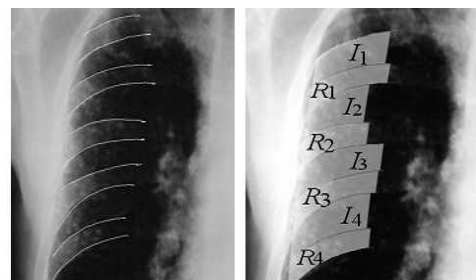
3. 提案手法

じん肺を症状レベル別に分類するために、以下の 1)~5) の手順でじん肺画像の異常度を定量的に求める。まず、前処理として、医師がタブレット PC 上で胸部 X 線像（原画像）に 10 本の肋骨縁を描く（図 2）。ただし、患者によって鎖骨影に重なることがある最も上の肋骨縁は使用しない。この画像において、画素値 255 を閾値とし 2 値化し、ラベリングを行い上位 9 個の面積を残し、肋骨縁のみの画像を作成する。

- 1) 原画像の濃度ヒストグラムを伸張化する。
- 2) 原画像に選択的局所平均法を 2 回適用する。
- 3) 原画像にマスクサイズ 9×9 の移動平均を適用後、原画像と肋骨縁のみの画像を重ね合わせる。
- 4) 各肋骨縁において最も外側の画素を用いてベジェ曲線を作成する。ベジェ曲線を肺の内側に 1 ピクセルずつ平行方向にずらす走査をベジェ曲線と 2 つの肋骨縁の交点のうち一方が途切れるまで続け、肋骨領域 R_m ($m: 1 \sim 4$) と肋間領域 I_m を設定する（図 2）。 R_m において、肺の最も外側のベジェ曲線上の midpoint を $c[m, 1]$ 、 $c[m, 1]$ から上方向と下方向の肋骨縁までベジェ曲線に沿って 1 ピクセルずつ走査したときの走査数を $drmR[m, 1]$ 、 $c[m, 1]$ から上方向と下方向の走査における j 番目の濃度をそれぞれ $u[m, 1, j]$ および $l[m, 1, j]$ とする。 $c[m, 1]$ からベジェ曲線を肺の最も内側のベジェ曲線上の midpoint まで水平方向に 1 ピクセルずつずらしたときの走査数を $lengthR_m$ 、これらの総和 L_R とし、 $c[m, 1]$ および $u[m, k, j]$ と $l[m, k, j]$ の k をベジェ曲線上の k 番目の走査点として、各肋骨 R_m の異常度 $Abn(R_m)$ と肋骨全体における異常度 Abn_R をそれぞれ以下の式で定義する。

$$Abn(R_m) = \left[\frac{1}{LengthR_m} \sum_{k=1}^{lengthR_m} \left\{ \frac{1}{drmR[m, k]} \sum_{j=1}^{drmR[m, k]} |u[m, k, j] - l[m, k, j]|^2 \right\} \right]^{1/2}$$

$$Abn_R = \left[\frac{1}{L_R} \sum_{m=1}^4 \sum_{k=1}^{lengthR_m} \left\{ \frac{1}{drmR[m, k]} \sum_{j=1}^{drmR[m, k]} |u[m, k, j] - l[m, k, j]|^2 \right\} \right]^{1/2}$$



肋骨縁の描画画像 走査領域
図 2 肋骨と肋間に分割した走査領域の設定

†近畿大学大学院、Graduate School of Kinki University

‡ 東京大学大学院、Grad. School of the University of Tokyo

同様に、任意の肋間 I_m の異常度 $Abn(I_m)$ と肋間全体における異常度 Abn_I をそれぞれ以下の式で定義する。

$$Abn(I_m) = \left[\frac{1}{Length I_m} \sum_{k=1}^{length I_m} \left\{ \frac{1}{drrml[m,k]} \sum_{j=1}^{drrml[m,k]} |u[m,k,j] - l[m,k,j]|^2 \right\} \right]^{1/2}$$

$$Abn_I = \left[\frac{1}{L_I} \sum_{m=1}^4 \sum_{k=1}^{length I_m} \left\{ \frac{1}{drrml[m,k]} \sum_{j=1}^{drrml[m,k]} |u[m,k,j] - l[m,k,j]|^2 \right\} \right]^{1/2}$$

5) Abn_R と Abn_I の 2 つの特徴量を用いて、ランダムツリー[5]により肺画像を 0 型と 1,2,3 型 (1-3)、および、0,1 型 (0-1) と 2,3 型 (2-3) の 2 通りに型分類する。ランダムツリーは予想精度が高く過学習しないという利点がある。本研究では、肋骨と肋間のそれぞれにおける総合的な異常度を評価するため、 $Abn(R_m)$ と $Abn(I_m)$ は型分類に用いる特徴量に含めない。

4. 実験結果

まず、CCD スキャナ (EPSON GT-X750) で取り込んだ 67 枚の胸部 X 線画像 (右肺, 8bit 白黒濃淡, 300dpi) に対して Abn_R と Abn_I を算出した結果を表 1 に示す。表 1 において mean は平均, STD は標準偏差を示す。画像の内訳はILOのじん肺標準写真 13 枚 (0 型 2 枚, 1 型 3 枚, 2 型 3 枚, 3 型 5 枚), 厚生労働省のじん肺標準写真 9 枚 (0 型 3 枚, 1 型 4 枚, 2 型 1 枚, 3 型 1 枚), 医師から提供された胸部 X 線写真 45 枚 (0 型) である。表 1 から、型が上がるにつれて異常度は上昇していることがわかる。

次に、提案手法によりじん肺写真の型分類を行った。このとき、実験データを有効に活用するために、学習データを 2 枚選択し、残りの 65 枚をテストデータとし平均認識率を求める動作を交差法[6]によって繰り返した。ランダムツリーの主なパラメータである決定木の数と決定木のサイズは、どちらも 50 とした。0 型と 1,2,3 型 (1-3) 型の分類結果を表 2 に示し、0,1 型 (0-1) と 2,3 型 (2-3) の分類結果を表 3 に示す。また、比較のためにニューラルネットワーク (NN) と線形判別分析 (LDA) を用いた分類結果も示した。分類結果は、表 2 表 3 とともに F 値で示した。0-1 型と 2-3 型の分類において、3 種類の分類手法の中でランダムツリー (RT) による分類結果が最も良好であった。

表 4 に 0 型であると誤判定された 1 型の画像 1 枚 (Err1) と 2-3 型であると判断された 1 型の画像 1 枚 (Err2) の $Abn(R_m)$ と $Abn(I_m)$ を示す。表 1 に示す 1 型の画像における異常度の平均と比べて、Err1 の異常度が低くなっているのがわかる。図 3(a) に Err1 の画像を示す。この画像において肋骨影はほとんど目視できない。図 3(b) は、図 3(a) において、肋骨縁が描画されている箇所を拡大したものであるが、このようにじん肺写真のスキャンにおいて肋骨影や粒状影が抽出されず肋骨縁のみが手動で描画される場合、実際の型よりも異常度が低く算出された。また、 Abn_R において、 $|u[m, k, j] - l[m, k, j]|^2$ で求まる全値

表 1 異常度の算出結果

category level	Abn_I	Abn_R
0 (45 images)	8.5 (mean) 4.4 (STD)	7.1 (mean) 2.5 (STD)
1 (7 images)	16.4 (mean) 4.2 (STD)	14.7 (mean) 5.1 (STD)
2 (4 images)	20.6 (mean) 7.7 (STD)	20.1 (mean) 3.1 (STD)
3 (6 images)	25.4 (mean) 6.2 (STD)	21.7 (mean) 3.6 (STD)

表 2 0 型と 1-3 型の分類における F 値

型分類	0	1-3
RT	91.9%	77.8%
NN	92.9%	80.0%
LDA	92.9%	80.0%

表 3 0-1 型と 2-3 型の分類における F 値

型分類	0-1	2-3
RT	96.4%	83.3%
NN	95.0%	74.9%
LDA	95.6%	73.7%

表 4 型分類に失敗した画像の異常度 $Abn(x)$

$Abn(x)$								
x	R_1	R_1	R_1	R_1	I_1	I_1	I_1	I_1
Err1	11.1	9.8	7.2	8.3	6.2	7.1	7.4	7.6
Err2	11.3	10.5	13.5	30.5	17.6	25.8	19.0	19.3

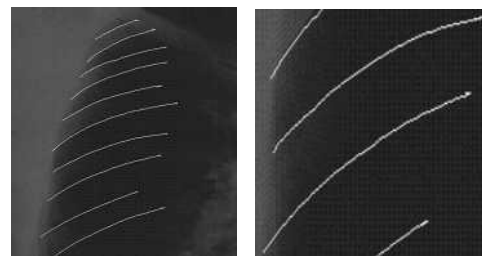


図 3 Err0 (表 4) の画像

の標準偏差について、Err2 の画像における標準偏差と 3 型の全画像における標準偏差の平均はそれぞれ 414.6 と 213.6 であった。このように、描画された肋骨縁と実際の肋骨縁とのずれが大きい場合、異常度の算出結果が実際の型よりも高くなると示唆される。

5. むすび

本稿では、ILO と厚生省のそれぞれのじん肺標準写真と健常の写真画像化し、ランダムツリーにより症状別に型分類する方法を提案した。異常度の信頼性は描画される肋骨縁の精度に依存すると考えられるため、今後は肋骨縁の描画マニュアルの作成、もしくは簡易かつ正確に肋骨縁を描画するためのインタフェースを作成し、型分類の精度を高めることを検討する。

参考文献

- [1] R. P. Kruger and W. B. Thompson: Computer Diagnosis of Pneumoconiosis, Trans. on Systems, Man and Cybernetics, Vol. SMC-4, No.1, pp.40-49 (1974).
- [2] 小畑秀文, 渡辺敏博: 2 次元自己回帰モデルに基づくテクスチャ識別, 信学論 A, Vol.71, No.2, pp.512-518 (1988).
- [3] 陳旋, 長谷川純一, 鳥脇純一郎: じん肺 X 線写真の定量診断のための粒状影の識別, 信学論 D-II, Vol.72, No.6, pp.944-953 (1989).
- [4] 小畑秀文, 久保淳, 岡田守弘: 等濃度線処理による X 線写真中のじん肺陰影抽出と自動診断への応用, 信学論, vol.J76-D-II, No.2, pp.261-267 (1993).
- [5] Leo, Breiman: Random forests, Machine Learning, Vol.45, pp.5-32 (2001).
- [6] F. Mosteller: A k-Sample Slippage Test for an Extreme Population, The Annals of Mathematical Statistics, Vol.19, No.1, pp.58-65 (1948).