

深層学習とマルチチャンネル化を用いた  
嚙下時 X 線動画からの頸椎椎間板領域の  
形態解析に関する研究

2021年 3月

藤中 彩乃

深層学習とマルチチャンネル化を用いた  
嚙下時 X 線動画からの頸椎椎間板領域の  
形態解析に関する研究

藤中 彩乃

システム情報工学研究科

筑波大学

2021年 3月

## 概要

嚙下は生命維持のために不可欠な動作であり、頸椎椎間板、頸椎椎体、舌骨などあらゆる頸部構造物の協調により成立する。本研究では、嚙下時 X 線動画 (videofluorography, VF) を用いて嚙下障害と頸椎の形動態の関連性を調べることを目的とした。はじめに、頸椎固定術術前患者群 (患者群) 19 症例と頸椎疾患のない実験参加者群 (正常群) 39 症例の VF から医師が手動抽出した椎間板領域の形態解析を行い、頸椎疾患の有無によって椎間板の形態が異なることを示した。嚙下障害と頸椎の形動態の関連性を調べるためには、嚙下障害の有無が明確な大規模な VF データセットを用いる必要があると考え、大規模なデータセットに対応するために、深層学習とモルフォロジー演算による後処理を用いて椎間板領域を自動抽出する手法を提案した。さらに、後処理の種類やパラメータが経験に依存しないよう、深層学習とマルチチャンネル化 (multi-channelization, MC) による前処理を用いて椎間板領域を自動抽出する方法を提案した。VF の各フレームは 8 ビットのグレースケール画像である。各フレーム画像に  $M$  種類の画像フィルタを別個に適用し、 $M$  種類の特徴画像を生成する。 $N$  種類を選択し、マルチチャンネル化画像 (MC 画像) と呼ばれる多チャンネル画像の各チャンネルに保存する。MC 画像に深層学習を適用し、椎間板領域を抽出する。焼きなまし法を用いて椎間板抽出精度が最大となる  $N$  種類の特徴画像の組合せを求める。実際の VF に適用した結果を示す。今後、嚙下障害の有無が明確な大規模な VF データセットに提案手法を適用し、自動抽出した嚙下障害患者群と正常嚙下群の頸椎領域に形態解析を行う手法を開発し、嚙下障害と頸椎の形動態の関連性を評価することによって、本研究の目的を達成できると考える。

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	研究背景	1
1.2	研究目的	2
1.3	本論文の構成	2
<b>第2章</b>	<b>関連研究</b>	<b>3</b>
2.1	人体を対象とした画像処理研究	3
2.2	頸椎疾患を対象とした研究	5
<b>第3章</b>	<b>医師が手動抽出した椎間板領域の形態解析</b>	<b>6</b>
3.1	患者群と正常群における椎間板の形態解析	6
3.2	椎間板の形態からの患者群と正常群の予測	8
3.3	実験	8
3.3.1	実験条件	8
3.3.2	実験結果	8
3.4	考察	9
<b>第4章</b>	<b>深層学習を用いた椎間板領域の抽出</b>	<b>11</b>
4.1	モルフォロジー演算による後処理を用いた方法	11
4.1.1	頸部マスクの抽出	11
4.1.2	深層学習とモルフォロジー演算を用いた椎間板領域の抽出	12
4.1.3	実験	18
4.1.3.1	パッチサイズを比較した実験	18
4.1.3.2	画素値正規化補正と回転補正の有無を比較した実験	21
4.1.3.3	椎間板領域全体のデータを用いてパッチサイズを比較した実験	23
4.1.3.4	椎間板領域全体のデータを用いて画素値正規化補正と回転補正の有無を比較した実験	23
4.1.3.5	深層学習の尤度を比較した実験	27
4.1.3.6	データ拡張を行った実験	29
4.1.3.7	VFデータセットの増加に伴い椎間板領域のデータと頸部マスクを改良した実験	31
4.1.3.8	58症例のVFを用いてパッチサイズを比較した実験	37



4.1.3.9	モルフォロジー演算と二値化, 小領域削除を用いた後処理の実験 . . . . .	37
4.1.3.10	58 症例の VF にホールドアウト法を適用してパッチサイズと画素値正規化補正の有無を比較した実験 . . . . .	42
4.1.3.11	モルフォロジー演算と二値化, 小領域削除, 深層学習の尤度を用いた後処理の実験 . . . . .	44
4.1.3.12	モルフォロジー演算と二値化, 深層学習の尤度を用いた後処理の実験 . . . . .	47
4.2	マルチチャンネル化による前処理を用いた方法 . . . . .	49
4.2.1	深層学習とマルチチャンネル化, 特徴画像の選択による椎間板領域の抽出 . . . . .	49
4.2.2	画像前処理を用いた従来研究 . . . . .	50
4.2.3	実験 . . . . .	53
4.2.3.1	特徴画像の作成 . . . . .	53
4.2.3.2	ガウシアンフィルタを基準に特徴画像の組合せを変更した実験 . . . . .	72
4.2.3.3	21 種類の特徴画像に焼きなまし法を適用した実験 . . . . .	74
4.2.3.4	学習, テスト症例と焼きなまし条件を変更して 21 種類の特徴画像に焼きなまし法を適用した実験 . . . . .	77
4.2.3.5	焼きなまし条件を変更して 45 種類の特徴画像に焼きなまし法を適用した実験 . . . . .	81
4.2.3.6	3-fold 3-phase cross validation を用いて LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet を比較した実験 . . . . .	82
4.2.3.7	焼きなまし条件を変更して U-Net, FPN, LinkNet, PSPNet を比較した実験 . . . . .	85
4.2.3.8	アウトライアを除外した実験 . . . . .	91
4.2.3.9	21 種類の特徴画像に焼きなまし法と PSPNet を適用した実験 . . . . .	92
4.2.3.10	深層学習の層数を変更した実験 . . . . .	93
4.2.3.11	チャンネル数を変更した実験 . . . . .	94
4.2.3.12	焼きなましの開始位置を変更した実験 . . . . .	95
4.2.4	考察 . . . . .	96
<b>第 5 章</b>	<b>結論</b>	<b>121</b>
	謝辞	122
	発表論文	123
	採用	125
	獲得研究費	125

受賞	125
参考文献	126

# 目次

2.1	人体を対象とした画像処理研究	3
3.1	患者症例のVFフレーム画像	7
3.2	患者症例の椎間板領域画像	7
3.3	正常症例のVFフレーム画像	7
3.4	正常症例の椎間板領域画像	7
3.5	椎間板の特徴量	8
4.1	VFフレーム画像	13
4.2	判別分析法を適用後の人体領域と背景領域	13
4.3	平滑化後の人体領域と背景領域	13
4.4	孤立領域削除後の人体領域と背景領域	13
4.5	人体領域と背景領域の境界線画像	14
4.6	背景の最大2領域の重心を結ぶ直線	14
4.7	直線の上下で二分した境界線画像	14
4.8	人体領域の距離画像のカラーマップ	14
4.9	背景領域の距離画像のカラーマップ	15
4.10	画像全体の距離画像のカラーマップ	15
4.11	首中心線と領域分割された境界線画像	15
4.12	首中心線からの距離画像のカラーマップ	15
4.13	領域分割した境界線画像	16
4.14	頭部と肩部の放物線の重畳表示	16
4.15	頸部領域	16
4.16	判別分析法を適用後の頸部領域	16
4.17	頸部マスク画像	17
4.18	学習症例のVFフレーム画像	19
4.19	椎間板領域の教師画像	19
4.20	テスト症例のVFフレーム画像	19
4.21	椎間板領域のGround truth	19
4.22	4.1.3.1節で $S_p = 15$ とした結果画像 ( $F$ 値 = 0.173)	20
4.23	4.1.3.1節で $S_p = 17$ とした結果画像 ( $F$ 値 = 0.181)	20
4.24	4.1.3.1節で $S_p = 19$ とした結果画像 ( $F$ 値 = 0.188)	20

4.25	4.1.3.1 節で $S_p = 21$ とした結果画像 ( $F$ 値 = <b>0.208</b> ) . . . . .	20
4.26	4.1.3.1 節で $S_p = 23$ とした結果画像 ( $F$ 値 = 0.201) . . . . .	21
4.27	4.1.3.2 節で $S_p = 15$ , 画素値正規化補正あり, 回転補正なしとした結果画像 ( $F$ 値 = <b>0.197</b> ) . . . . .	22
4.28	4.1.3.2 節で $S_p = 15$ , 画素値正規化補正なし, 回転補正ありとした結果画像 ( $F$ 値 = 0.149) . . . . .	22
4.29	4.1.3.2 節で $S_p = 15$ , 画素値正規化補正なし, 回転補正なしとした結果画像 ( $F$ 値 = 0.124) . . . . .	22
4.30	椎間板領域全体の教師画像 . . . . .	24
4.31	椎間板領域全体の Ground truth . . . . .	24
4.32	4.1.3.3 節で $S_p = 15$ とした結果画像 ( $F$ 値 = 0.390) . . . . .	24
4.33	4.1.3.3 節で $S_p = 17$ とした結果画像 ( $F$ 値 = 0.410) . . . . .	25
4.34	4.1.3.3 節で $S_p = 19$ とした結果画像 ( $F$ 値 = <b>0.424</b> ) . . . . .	25
4.35	4.1.3.3 節で $S_p = 21$ とした結果画像 ( $F$ 値 = 0.418) . . . . .	25
4.36	4.1.3.3 節で $S_p = 23$ とした結果画像 ( $F$ 値 = 0.395) . . . . .	25
4.37	4.1.3.3 節で $S_p = 25$ とした結果画像 ( $F$ 値 = 0.363) . . . . .	26
4.38	4.1.3.3 節で $S_p = 27$ とした結果画像 ( $F$ 値 = 0.374) . . . . .	26
4.39	4.1.3.3 節で $S_p = 29$ とした結果画像 ( $F$ 値 = 0.342) . . . . .	26
4.40	4.1.3.3 節で $S_p = 31$ とした結果画像 ( $F$ 値 = 0.332) . . . . .	26
4.41	4.1.3.3 節で $S_p = 33$ としたときの椎間板抽出結果画像 ( $F$ 値 = 0.318) . . . . .	28
4.42	4.1.3.3 節で $S_p = 35$ としたときの椎間板抽出結果画像 ( $F$ 値 = 0.301) . . . . .	28
4.43	4.1.3.4 節で画素値正規化補正あり, 回転補正なしとした結果画像 ( $F$ 値 = <b>0.406</b> ) . . . . .	28
4.44	4.1.3.4 節で画素値正規化補正なし, 回転補正ありとした結果画像 ( $F$ 値 = 0.363) . . . . .	28
4.45	4.1.3.4 節で画素値正規化補正なし, 回転補正なしとした結果画像 ( $F$ 値 = 0.403) . . . . .	29
4.46	データ拡張を適用して得られた結果画像 ( $F$ 値 = 0.503) . . . . .	31
4.47	図 4.147 を二値化して得られた人体領域と背景領域 . . . . .	32
4.48	平滑化後の人体領域と背景領域 . . . . .	32
4.49	孤立領域削除後の人体領域と背景領域 . . . . .	32
4.50	頸部マスク A . . . . .	32
4.51	図 4.147 を二値化して得られた人体領域と背景領域 . . . . .	33
4.52	画像を $-5$ 度回転して得られた XY 座標系 . . . . .	33
4.53	人体領域の画素数でしきい値を定めて得られた <i>imgKeep</i> . . . . .	34
4.54	頸部マスク B . . . . .	34
4.55	頸部マスク C . . . . .	34
4.56	患者症例の VF フレーム画像 . . . . .	35
4.57	患者症例の椎間板領域画像 . . . . .	35
4.58	正常症例の VF フレーム画像 . . . . .	35
4.59	正常症例の椎間板領域画像 . . . . .	35

4.60	頸部マスクの作成が不可能な症例の VF フレーム画像	36
4.61	判別分析法を適用後の人体領域と背景領域	36
4.62	平滑化後の人体領域と背景領域	36
4.63	図 4.20 と同一の患者症例の Ground truth	38
4.64	4.1.3.8 節で $S_p = 15$ とした結果画像 ( $F$ 値 = 0.387)	38
4.65	4.1.3.8 節で $S_p = 17$ とした結果画像 ( $F$ 値 = 0.420)	38
4.66	4.1.3.8 節で $S_p = 19$ とした結果画像 ( $F$ 値 = 0.440)	38
4.67	4.1.3.8 節で $S_p = 21$ とした結果画像 ( $F$ 値 = 0.454)	39
4.68	4.1.3.8 節で $S_p = 23$ とした結果画像 ( $F$ 値 = <b>0.464</b> )	39
4.69	4.1.3.8 節で $S_p = 25$ とした結果画像 ( $F$ 値 = 0.461)	39
4.70	4.1.3.9 節で図 4.66 に closing フィルタを適用して得られた画像 ( $F$ 値 = 0.557)	40
4.71	4.1.3.9 節で VF フレーム画像に Top-hat 変換フィルタと二値化を適用して得られた画像	40
4.72	4.1.3.9 節で図 4.70 と図 4.71 の論理積を計算して得られた画像 ( $F$ 値 = 0.695)	41
4.73	4.1.3.9 節で体積が 1000 ボクセル以下の小領域を削除して得られた画像 ( $F$ 値 = 0.714)	41
4.74	4.1.3.9 節で erosion フィルタを適用して得られた画像 ( $F$ 値 = 0.655)	41
4.75	4.1.3.9 節で面積が 30 ピクセル以下の小領域を削除して得られた画像 ( $F$ 値 = 0.671)	41
4.76	4.1.3.9 節で dilation フィルタを適用して得られた画像 ( $F$ 値 = 0.795)	42
4.77	4.1.3.9 節の最終結果画像 (半径 2 ピクセルの誤差を許容した $F$ 値 = <b>0.909</b> )	42
4.78	患者テスト症例の VF フレーム画像	45
4.79	患者テスト症例の Ground truth	45
4.80	4.1.3.10 節で $S_p = 19$ とした画素値正規化補正後に深層学習を適用して得られた椎間板候補領域 ( $F$ 値 = 0.568)	45
4.81	4.1.3.11 節で作成した深層学習の尤度画像のカラーマップ	45
4.82	4.1.3.11 節で図 4.81 を二値化して得られた画像 ( $F$ 値 = 0.567)	46
4.83	4.1.3.11 節で小領域を削除して得られた画像 ( $F$ 値 = 0.596)	46
4.84	4.1.3.11 節で VF フレーム画像に Top-hat 変換フィルタを適用して得られた画像	46
4.85	4.1.3.11 節で図 4.84 を二値化して得られた画像 ( $F$ 値 = 0.642)	46
4.86	4.1.3.11 節で closing フィルタを適用して得られた画像 ( $F$ 値 = 0.610)	47
4.87	4.1.3.11 節の最終結果画像 (半径 2 ピクセルの誤差を許容した $F$ 値 = <b>0.765</b> )	47
4.88	4.1.3.12 節で尤度画像に erosion フィルタを適用後のカラーマップ	48
4.89	4.1.3.12 節で二値化して得られた画像 ( $F$ 値 = 0.627)	48
4.90	4.1.3.12 節の最終結果画像 (半径 2 ピクセルの誤差を許容した $F$ 値 = <b>0.850</b> )	48
4.91	深層学習と MC, 焼きなまし法に基づく特徴画像の選択による椎間板の領域抽出	51
4.92	ヒストグラム平坦化フィルタ	55
4.93	正常症例の VF フレーム画像	55

4.94 特徴画像 HGE . . . . .	55
4.95 ネガ・ポジ反転フィルタ . . . . .	56
4.96 特徴画像 NGP . . . . .	56
4.97 ソラリゼーションフィルタ . . . . .	57
4.98 特徴画像 SOL . . . . .	57
4.99 特徴画像 MEA1 . . . . .	57
4.100特徴画像 MEA2 . . . . .	57
4.101ガウシアンフィルタのカーネル . . . . .	58
4.102特徴画像 GAU1 . . . . .	58
4.103特徴画像 GAU2 . . . . .	58
4.104特徴画像 MED1 . . . . .	59
4.105特徴画像 MED2 . . . . .	59
4.106K 最近傍平均化フィルタの処理の具体例 . . . . .	60
4.107特徴画像 KNN1 . . . . .	60
4.108特徴画像 KNN2 . . . . .	60
4.109特徴画像 BLT1 . . . . .	61
4.110特徴画像 BLT2 . . . . .	61
4.111ラプラシアンフィルタのカーネル . . . . .	61
4.112特徴画像 LPL1 . . . . .	61
4.113特徴画像 LPL2 . . . . .	61
4.114特徴画像 ETC1 . . . . .	62
4.115特徴画像 ETC2 . . . . .	62
4.116特徴画像 STC1 . . . . .	63
4.117特徴画像 STC2 . . . . .	63
4.118特徴画像 NML1 . . . . .	63
4.119特徴画像 NML2 . . . . .	63
4.120特徴画像 NML3 . . . . .	64
4.121Local binary pattern フィルタの処理の具体例 . . . . .	64
4.122Dilated フィルタの処理の具体例 . . . . .	65
4.123特徴画像 LBP1 . . . . .	65
4.124特徴画像 LBP2 . . . . .	65
4.125特徴画像 LBP3 . . . . .	65
4.126特徴画像 POS1 . . . . .	66
4.127特徴画像 POS2 . . . . .	66
4.128特徴画像 POS3 . . . . .	66
4.129エンボスフィルタのカーネル . . . . .	67
4.130特徴画像 EMB1 . . . . .	67
4.131特徴画像 EMB2 . . . . .	67

4.132特徴画像 EMB3	67
4.133特徴画像 TOP1	68
4.134特徴画像 TOP2	68
4.135特徴画像 TOP3	69
4.136特徴画像 TOP4	69
4.137特徴画像 TOP5	69
4.138水平, 垂直方向の Sobel フィルタのカーネル	70
4.139特徴画像 SBLX1	70
4.140特徴画像 SBLX2	70
4.141特徴画像 SBLX3	71
4.142特徴画像 SBLY1	71
4.143特徴画像 SBLY2	71
4.144特徴画像 SBLY3	71
4.145特徴画像 SBLM	72
4.146特徴画像 SBLD	72
4.147正常症例の VF フレーム画像	75
4.148正常症例の Ground truth	75
4.149(GAU1, GAU1, GAU1) を適用して得られた結果画像 (F 値=0.664)	75
4.150(GAU1, TOP5, LBP2) を適用して得られた結果画像 (F 値=0.733)	75
4.151(GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.650)	78
4.152(SBLM, SOL, LPL1) を適用して得られた結果画像 (F 値=0.701)	78
4.153正常症例の VF フレーム画像	80
4.154正常症例の Ground truth	80
4.155(GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.521)	80
4.156(SBLM, TOP5, SBLM) を適用して得られた結果画像 (F 値=0.596)	80
4.157(GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.533)	83
4.158(STC1, TOP1, POS2) を適用して得られた結果画像 (F 値=0.559)	83
4.159(GRY, GRY, GRY) に LeNet を適用して得られた結果画像 (F 値=0.634)	86
4.160(GRY, GRY, GRY) に AlexNet を適用して得られた結果画像 (F 値=0.214)	86
4.161(GRY, GRY, GRY) に U-Net を適用して得られた結果画像 (F 値=0.680)	86
4.162(GRY, GRY, GRY) に FPN を適用して得られた結果画像 (F 値=0.791)	86
4.163(GRY, GRY, GRY) に LinkNet を適用して得られた結果画像 (F 値=0.800)	87
4.164(GRY, GRY, GRY) に PSPNet を適用して得られた結果画像 (F 値=0.765)	87
4.165(SBLX2, SBLY1, POS3) に LeNet を適用して得られた結果画像 (F 値=0.657)	87
4.166(SBLY1, SBLX2, SBLY3) に AlexNet を適用して得られた結果画像 (F 値=0.666)	87
4.167(HGE, HGE, HGE) に U-Net を適用して得られた結果画像 (F 値=0.792)	88
4.168(EMB3, EMB3, GAU1) に FPN を適用して得られた結果画像 (F 値=0.782)	88
4.169(EMB3, ETC1, BLT1) に LinkNet を適用して得られた結果画像 (F 値=0.782)	88

4.170(ETC2, MEA2, MED2) に PSPNet を適用して得られた結果画像 (F 値=0.766) . . .	88
4.171(EMB1, SBLD, SBLY1) に U-Net を適用して得られた結果画像 (F 値=0.808) . . .	90
4.172(ETC2, NML1, MED2) に FPN を適用して得られた結果画像 (F 値=0.783) . . .	90
4.173(BLT1, BLT1, BLT1) に LinkNet を適用して得られた結果画像 (F 値=0.787) . . .	90
4.174(NML3, NML3, NML3) に PSPNet を適用して得られた結果画像 (F 値=0.769) . . .	90
4.175除外した患者症例 1 . . . . .	99
4.176除外した患者症例 2 . . . . .	99
4.177正常症例の VF フレーム画像 . . . . .	99
4.178U-Net を適用して得られた EMB2 の近似画像 . . . . .	99
4.179EMB2 の Ground truth . . . . .	100
4.180EMB2 の差分画像 . . . . .	100
4.181U-Net を適用して得られた TOP2 の近似画像 . . . . .	100
4.182TOP2 の Ground truth . . . . .	100
4.183TOP2 の差分画像 . . . . .	101
4.184U-Net を適用して得られた LBP1 の近似画像 . . . . .	101
4.185LBP1 の Ground truth . . . . .	101
4.186LBP1 の差分画像 . . . . .	101
4.187U-Net を適用して得られた SBLD の近似画像 . . . . .	102
4.188SBLD の Ground truth . . . . .	102
4.189SBLD の差分画像 . . . . .	102



# 表目次

3.1	患者群と正常群における $lrat\_d$ , $ang\_d$ , $lam\_d$ , $len2\_d$ , $len1\_d$ , $brat\_d$ , $bod\_d$ の有意差検定 . . . . .	9
4.1	混同行列 . . . . .	17
4.2	4.1.3.1 節で $S_p$ を 15 から 23 まで 2 刻みで変更して得られた平均椎間板抽出精度	21
4.3	4.1.3.2 節で画素値正規化補正と回転補正の有無を変更して得られた平均椎間板抽出精度 . . . . .	23
4.4	4.1.3.3 節で $S_p$ を 15 から 35 まで 2 刻みで変更して得られた平均椎間板抽出精度	27
4.5	4.1.3.4 節で画素値正規化補正と回転補正の有無を変更して得られた平均椎間板抽出精度 . . . . .	27
4.6	深層学習の尤度を変更して得られた平均椎間板抽出精度 . . . . .	30
4.7	データ拡張の有無を変更して得られた平均椎間板抽出精度 . . . . .	30
4.8	4.1.3.8 節で $S_p$ を 15 から 25 まで 2 刻みで変更して得られた平均椎間板抽出精度	37
4.9	4.1.3.9 節で $S_p$ を 15 から 25 まで 2 刻みに変更して得られた各後処理段階の平均 F 値 . . . . .	43
4.10	4.1.3.10 節で $S_p$ と画素値正規化補正の有無を変更して得られた平均椎間板抽出精度 . . . . .	43
4.11	4.1.3.11 節で $S_p$ を 15 から 25 まで 2 刻みに変更して画素値正規化補正を適用して得られた各後処理段階の平均 F 値 . . . . .	44
4.12	4.1.3.12 節で $S_p$ を 15 から 25 まで 2 刻みに変更して画素値正規化補正を適用して得られた各後処理段階の平均 F 値 . . . . .	47
4.13	特徴画像の種類と略語の一覧 . . . . .	53
4.14	4.2.3.2 節で得られた 11 種類の特徴画像の順位 . . . . .	73
4.15	GAU1 を含む全通りの特徴画像の組合せ . . . . .	73
4.16	4.2.3.3 節で得られた 21 種類の特徴画像の順位 . . . . .	76
4.17	4.2.3.3 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	77
4.18	4.2.3.4 節で得られた 21 種類の特徴画像の順位 . . . . .	78
4.19	4.2.3.4 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	79
4.20	4.2.3.5 節で得られた上位 10 種類の同一の特徴画像 . . . . .	81

4.21	4.2.3.5 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	82
4.22	深層学習のパラメータ . . . . .	84
4.23	3-fold 3-phase cross validation を用いた実験の流れ . . . . .	84
4.24	4.2.3.6 節で LeNet を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	89
4.25	4.2.3.6 節で AlexNet を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	91
4.26	4.2.3.6 節で U-Net を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	92
4.27	4.2.3.6 節で FPN を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	93
4.28	4.2.3.6 節で LinkNet を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	94
4.29	4.2.3.6 節で PSPNet を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	95
4.30	4.2.3.6 節で焼きなまし法と LeNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	96
4.31	4.2.3.6 節で焼きなまし法と AlexNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	97
4.32	4.2.3.6 節で焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	98
4.33	4.2.3.6 節で焼きなまし法と FPN を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	103
4.34	4.2.3.6 節で焼きなまし法と LinkNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	104
4.35	4.2.3.6 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	105
4.36	4.2.3.6 節で 6 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せを入力して得られた平均 F 値 . . . . .	105
4.37	4.2.3.7 節で焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	106
4.38	4.2.3.7 節で焼きなまし法と FPN を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	107
4.39	4.2.3.7 節で焼きなまし法と LinkNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	108
4.40	4.2.3.7 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	109
4.41	4.2.3.7 節で 6 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せを入力して得られた平均 F 値 . . . . .	109
4.42	アウトライアを除外し, 焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	110
4.43	4.2.3.7 節と 4.2.3.8 節の (GRY, GRY, GRY) と最適な特徴画像の組合せの比較 . . . . .	110
4.44	4.2.3.7 節と 4.2.3.8 節の症例ベースの比較 . . . . .	111

4.45	4.2.3.9 節で PSPNet を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	111
4.46	4.2.3.9 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	112
4.47	4.2.3.7 節と 4.2.3.9 節の (GRY, GRY, GRY) と最適な特徴画像の組合せの比較 . . . . .	112
4.48	2 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	112
4.49	3 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	113
4.50	4 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像 . . . . .	113
4.51	焼きなまし法と 2 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	114
4.52	焼きなまし法と 3 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	114
4.53	焼きなまし法と 4 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	115
4.54	2 層, 3 層, 4 層, 5 層の U-Net を用いて得られた (GRY, GRY, GRY) と最適な特徴画像の組合せの比較 . . . . .	115
4.55	1 チャンネルの実験で得られた上位 10 種類の同一の特徴画像 . . . . .	115
4.56	2 チャンネルの実験で得られた上位 10 種類の同一の特徴画像 . . . . .	116
4.57	4 チャンネルの実験で得られた上位 10 種類の同一の特徴画像 . . . . .	116
4.58	2 チャンネルの実験で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	117
4.59	4 チャンネルの実験で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	117
4.60	チャンネル数を 1, 2, 3, 4 と変更して得られた (GRY, GRY, GRY) と最適な特徴画像の組合せの比較 . . . . .	117
4.61	4.2.3.12 節で得られた上位 10 種類の同一の特徴画像 . . . . .	118
4.62	焼きなまし法を (1 位, 1 位, 1 位) から開始して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	118
4.63	焼きなまし法を (14 位, 14 位, 14 位) から開始して得られた上位 10 種類の最適な特徴画像の組合せ . . . . .	119
4.64	EMB2, TOP2, LBP1, SBLD を近似して得られた各症例の RMSE . . . . .	119
4.65	(1 位, 1 位, 1 位) とその周囲 26 近傍の特徴画像の組合せ . . . . .	120

# 第1章 序論

## 1.1 研究背景

嚥下とは、狭義には口腔にある食べ物を胃まで飲み込むことをいう。狭義の嚥下に食物の認知、食物の口腔への取り込み、咀嚼を加えた概念を摂食嚥下という。摂食嚥下は、食物を認知して口腔へ取り込む認知期、咀嚼をして食塊を形成する口腔準備期、舌の運動により食塊を口峽に送り込む口腔期、嚥下反射により食塊が上咽頭から下咽頭へ送り込まれる咽頭期、食道の蠕動運動により食塊が頸部食道から腹部食道へ送り込まれる食道期に分けられる。特に、咽頭期の嚥下反射では、軟口蓋が挙上して鼻腔への逆流を防ぎ、咽頭が収縮して食塊が下咽頭へ移動し、喉頭が挙上して食道入口部が開大し、喉頭蓋が反転して声帯が閉鎖することで気管への誤嚥を防止する。したがって、嚥下は、舌、軟口蓋、咽頭、喉頭、喉頭蓋、喉頭蓋谷、梨状窩、声帯、舌骨、甲状軟骨、食道、頸椎などあらゆる頸部構造物の協調運動によって成立する。それぞれの頸部構造物の形態や運動が障害されることで嚥下障害が生じる。嚥下障害の原因疾患を考える際には、嚥下のいずれの期が障害されるかという観点が重要となる。口腔期が障害される疾患の例として舌癌、パーキンソン病、筋委縮性側索硬化症が、咽頭期が障害される疾患の例として前縦靭帯骨化症、多発性硬化症が、食道期が障害される疾患の例として食道炎、変形性脊椎症、強皮症が挙げられる。また、嚥下障害の症状として、口腔内汚染、誤嚥、咽喉頭の異常感覚、湿性嘔声を呈することがあり [1][2][3]、嚥下障害が進行すると、脱水や低栄養、誤嚥性肺炎を来し、生活の質が低下する [4][5]。誤嚥性肺炎は令和元年の死因第6位となっており [6]、嚥下障害は患者個人のみならず社会全体の問題となっている。したがって、嚥下障害を早期に発見し、診断、治療を行うことが重要である。嚥下障害の診断に最も頻繁に用いられているモダリティの一つが嚥下時 X 線動画 (videofluorography, VF) である。VF により、口腔期、咽頭期、食道期の嚥下運動を視覚的に捉え、機能の異常を指摘することができる。VF の平均透視時間は 5 分であり、入射表面線量は 30 mGy である。これは胸部 1 相 CT の 13 mGy、胸部正面 X 線画像の 0.3 mGy [10] と比較して大きい値であるが、VF のフレームあたりの線量は少なく、時間分解能も高い。VF の観察項目には、喉頭蓋谷残留、喉頭閉鎖、誤嚥などが含まれており、各項目の異常度を点数化することで嚥下障害の発見が可能となる [7][8][9]。

## 1.2 研究目的

従来の嚙下に関する研究は、舌骨、声帯など嚙下に伴って変形、移動する頸部構造物に着目したものが多かった [28][29][32][35]。本論文では、嚙下障害と頸椎の形動態の関連性を調べることを目的とする。これを実現するために、嚙下障害を確認するモダリティを VF に限定し、頸椎を頸椎椎間板に限定し、形動態を形態に限定したうえで、VF を用いて嚙下障害と頸椎椎間板の形態の関連性を調べるという小目的を設定する。椎間板は主に軟骨組織からなる円盤状の構造物であり、椎体と椎体の間に存在する。本論文では、VF を用いて嚙下障害と頸椎椎間板の形態の関連性を調べるために、まず、VF から椎間板の形態解析を行う際の指標値の作成と評価を行う。次に、それらの指標値を用いて大規模な VF データセットから嚙下障害と椎間板の形態の関連性を調べることを目的とし、深層学習を用いて VF から椎間板領域を自動抽出する手法を提案する。

## 1.3 本論文の構成

本論文の構成を述べる。第 2 章では関連研究、第 3 章では医師が手動抽出した椎間板領域の形態解析、第 4 章では深層学習を用いた椎間板領域の抽出、第 5 章では結語を述べる。

## 第2章 関連研究

### 2.1 人体を対象とした画像処理研究

本論文ではVFを用いて嚥下障害と頸椎の形動態との関連性解明を目指すため、関連研究として人体を対象とした画像処理研究を紹介する。人体は、神経系、呼吸器系、循環器系、消化器系、筋軟部組織、骨などあらゆる器官に分けられる。骨には、脊椎、頭蓋骨、肋骨、肩甲骨、腸骨、四肢の骨などが含まれる。さらに、脊椎は、頸椎、胸椎、腰椎、仙椎に分けられる。本論文では、頸椎に含まれる椎間板を対象とする。図2.1に人体を対象とした画像処理研究の区分を示す。

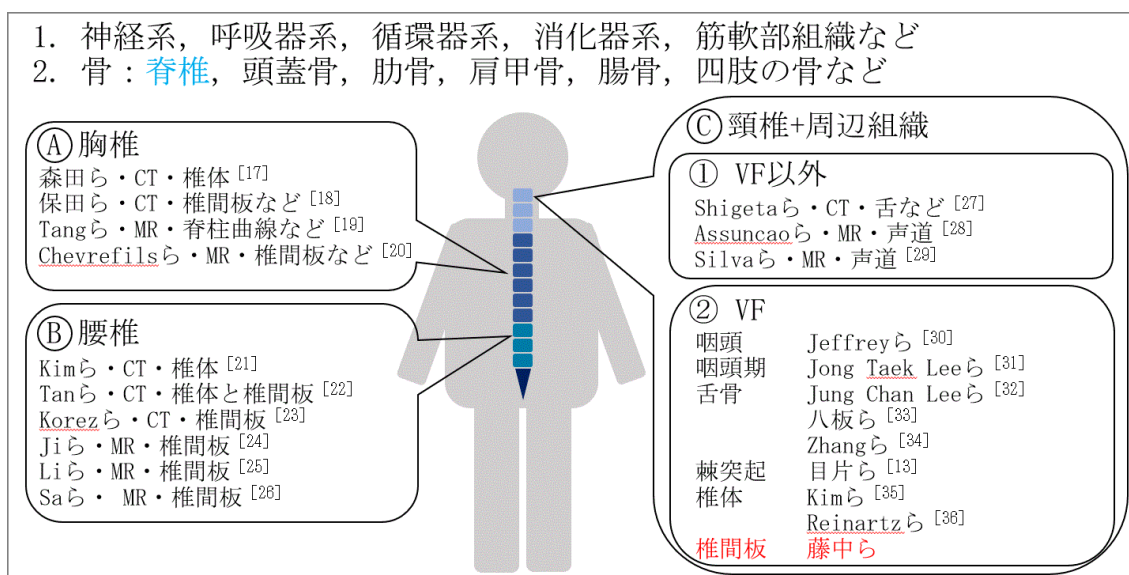


図 2.1: 人体を対象とした画像処理研究

胸椎を対象とした画像処理研究として、以下の例が挙げられる。森田ら [17] は、骨粗鬆症診断支援システムを用いて胸部 CT 画像から胸椎、腰椎の椎体を抽出し、皮質骨と海綿骨の骨密度を測定した。保田ら [18] は、モルフォロジー演算と平面モデル、曲面モデルを用いて腹部 CT 画像から脊柱、肋骨、胸椎椎間板、椎体を段階的に自動認識する手法を提案した。Tangら [19] は、勾配法, graph-based method, active shape model (ASM) を用いて MR 画像から頸椎、胸椎、腰椎の椎間板、椎体、脊柱曲線を自動抽出する手法を提案した。Chevrefilsら [20] は、Watershed 変換とモルフォロジー演算を用いて、MR 画像から胸椎と腰椎の椎間板と脊柱

管を自動抽出する手法を提案した。

腰椎を対象とした画像処理研究として、以下の例が挙げられる。Yiebin Kim ら [21] は、画素値と位置情報を利用した 3 次元 fence の作成とラベリングを用いて、CT 画像から腰椎椎体を自動抽出する手法を提案した。Tan ら [22] は、geodesic active contour を用いて腰部 CT 画像から半自動的に腰椎椎体と椎間板の厚みを計測する手法を提案した。Korez ら [23] は、パラメトリックモデリングを用いて CT 画像から腰椎椎間板を自動抽出する手法を提案した。Ji ら [24] は、パッチベースの convolutional neural network (CNN) を用いて MR 画像から腰椎椎間板を自動抽出する手法を提案した。Li ら [25] は、fully convolutional network を用いて腰部マルチモデル MR 画像から腰椎椎間板を自動抽出する手法を提案した。Sa ら [26] は、faster-RCNN を用いて X 線画像から腰椎椎間板を自動検出する手法を提案した。

頸椎と周辺組織を対象とした画像処理研究のうち、VF 以外のモダリティを用いた研究として以下の例が挙げられる。Shigeta ら [27] は、睡眠時無呼吸症候群の男性患者 40 例の CT 画像を撮影し、半自動セグメンテーションツールを用いて 3 次元再構成と舌領域の抽出を行うことで、舌と下顎の容積比率が気道に影響を与えることを示した。Assuncao ら [28] は、レベルセット法を用いてリアルタイム MR 画像から声道を自動抽出する手法を提案した。ただし、声道とは下顎、舌、硬口蓋、軟口蓋、喉頭蓋、声帯の総称を指す。Silva ら [29] は、active appearance model (AAM) を用いて MR 画像から声道を自動抽出する手法を提案した。

頸椎と周辺組織を対象とした画像処理研究のうち、VF を用いた研究として以下の例が挙げられる。Jeffrey ら [30] は、健康な実験参加者の上部中咽頭後壁、下部中咽頭後壁、咽頭食道セグメント、頸部食道の計 4 か所に放射線不透過性吸引マーカーを取り付けて VF を撮影し、嚥下中のマーカーの変位を手動で測定することで、嚥下中に咽頭が短縮することを示した。Jong Taek Lee ら [31] は、3D CNN を用いて VF から咽頭期を自動検出する手法を提案した。Seong Jae Lee ら [33] は、U-Net[80] を用いて VF から誤嚥を自動検出する手法を提案した。Jung Chan Lee ら [32] は、VF から舌骨運動に関連する 26 個の特徴量を手動抽出し、SVM を用いて嚥下障害を識別する手法を提案した。八板ら [34] は、support vector machine (SVM) と時空間最大体積法を用いて VF から舌骨を自動追跡する手法を提案した。Zhang ら [35] は、SDD500-ResNet や YOLOv2 をはじめとする 6 種類の深層学習モデルを用いて VF から舌骨を自動検出する手法を提案した。目片ら [13] は、テンプレートマッチングを用いて VF から椎体と棘突起を半自動追跡する手法を提案した。Youngkook Kim ら [36] は、脳卒中後の患者の VF を用いて頸椎を手動計測することで、患者群を前頸部骨棘の有無で二分し、前頸部骨棘の発生が誤嚥や経管栄養、持続する嚥下障害のリスクとなることを示した。Reinartz ら [37] は、正規化勾配場モデルを用いて VF から頸椎椎体を自動追跡する手法を提案した。しかしながら、著者の知る限り、嚥下障害と頸椎の形動態の関連性を調べることを目的とし、VF から椎間板の形態解析を行うための指標値の作成と評価を行った研究は他にない。

## 2.2 頸椎疾患を対象とした研究

本論文では頸椎の形動態に着目して嚙下障害の解明を目指すため、頸椎疾患を対象とした関連研究を紹介する。中瀬ら [11] と中寺ら [12] は、椎間板が加齢に伴い変性、変形することを示した。加齢とともに、椎間板の髄核と繊維輪における軟骨細胞が変性し、MR 画像内で肉眼で確認できる椎間板の変形も増加することが示された。目片ら [13] は、頸椎固定術術前患者と頸椎疾患の既往がない実験参加者 (以下、正常群) から構成される VF データセットにおいて、上述の通り、テンプレートマッチングを用いて椎体と棘突起を半自動追跡し、頸椎前弯角の動態解析を行うことで、頸椎の形状は嚙下障害と関連することを示した。頸椎固定術は、頸椎間の関節の不安定性を解消するための手術であり、頸椎間の関節を螺子で固定する。頸椎固定術の術式は、椎間板ヘルニアを対象とする前方固定術と、頸部脊柱管狭窄症や変形性頸椎症を対象とする後方固定術に分けられる。後方固定術の一種として、環軸椎後方固定を行う Magerl 法 [14] が挙げられる。藤縄ら [15] は、誤嚥性肺炎を起こした頸髄損傷患者 8 例を検討した。頸椎前方固定術を施行した既往歴短期群 5 例と比較し、頸椎前方固定術を施行しなかった既往歴長期群 3 例の嚙下障害の予後が不良であったことを報告した。Izeki ら [16] は、頸椎後方固定術を施行後の 33 例を検討し、手術中に決定した occipito-C2angle (O-C2A) が術後の呼吸困難や嚙下障害に影響を与えることを示唆した。このように、頸椎固定術と嚙下障害の関連性が広く報告されているが、頸椎間の位置関係や嚙下中の頸椎の形態変化が嚙下障害と関連している可能性があり、嚙下時の頸椎の形動態解析が重要となる。



## 第3章 医師が手動抽出した椎間板領域の形態解析

嚙下障害と頸椎の形態の関連性を調べることを目的とし、本章では、VFを用いて頸椎椎間板の形態解析を行う手法を提案する。提案手法を用いて、頸椎疾患と椎間板の形態の関連性を調べる。

椎間板の形態は頸椎疾患により容易に変化する。例えば、頸椎椎間板ヘルニアは椎間板の突出を来とし、関節リウマチは関節裂隙の狭小化を来たしうる。逆に、椎間板の形態の異常からそれらの疾患を鑑別に挙げることができる。本章では頸椎固定術前患者(以下、患者群)と頸椎疾患のない実験参加者(以下、正常群)で椎間板の大きさや厚さ、形状に差異があるかどうかを調査する。本章では、医師(著者)が医学的な観点から作成した椎間板領域のデータを用いる。

### 3.1 患者群と正常群における椎間板の形態解析

VFは矢状方向に撮影されたものであるため、本来は円盤状の椎間板が楕円形に見える。患者群と正常群の各症例のVF上で、舌骨最大挙上フレームを1フレームのみ選定し、舌骨最大挙上フレームで目視可能な頸椎椎間板を対象とした。椎間板の高位に応じて、第2頸椎椎体と第3頸椎椎体に挟まれた椎間板をC2、第3頸椎椎体と第4頸椎椎体に挟まれた椎間板をC3、第4頸椎椎体と第5頸椎椎体に挟まれた椎間板をC4、第5頸椎椎体と第6頸椎椎体に挟まれた椎間板をC5、第6頸椎椎体と第7頸椎椎体に挟まれた椎間板をC6、第7頸椎椎体と第1胸椎椎体に挟まれた椎間板をC7と表記する。舌骨最大挙上フレームで目視不能な頸椎椎間板については後述する指標値を求めなかった。

本章で作成した椎間板の形態に関連する指標値について述べる。まず、各症例のそれぞれの椎間板に対して主成分分析を適用し、第1主成分と第2主成分のベクトルを算出した。図3.3、図3.4、図3.1、図3.2に患者症例と正常症例のVFフレーム画像と椎間板領域画像をそれぞれ示す。なお、椎間板領域上の第1、第2主成分を赤線で示した。第2主成分の固有値を第1主成分の固有値で割った値を $lam$ 、椎間板上の第1、第2主軸方向の線分の長さをそれぞれ $len1$ 、 $len2$ とした。 $len1$ が椎間板の直径、 $len2$ が椎間板の厚さに相当する。ただし、実験参加者の体格に依存しないよう、全ての椎間板における $len1$ と $len2$ を同一症例のC4における $len2$ の値で割って正規化した。それぞれの椎間板における $len1$ を $len2$ で割った値を $lrat$ とした。 $C_k$ 重心と $C_{k+1}$ 重心の距離を $bod_k$ とした。 $bod_k$ を $C_k$ の $len2$ で割った値を $brat_k$ とした。 $bod$ は椎体の厚さに、 $brat$ は椎体と椎間板の厚さの比率に相当する。 $C_k$ の第1主成分と

$C_{k+1}$  の第 1 主成分のなす角を  $ang_k$  とした。ただし、値の正負は無視し、椎間板同士のなす角が  $90^\circ$  を超えることはないため  $0^\circ < ang < 90^\circ$  とした。それぞれの症例の  $C2$  から  $C7$  の  $lam$ ,  $len1$ ,  $len2$ ,  $lrat$  の分散を  $lam\_d$ ,  $len1\_d$ ,  $len2\_d$ ,  $lrat\_d$  とし、それぞれの症例の  $C2$  から  $C6$  の  $bod$ ,  $brat$ ,  $ang$  の分散を  $bod\_d$ ,  $brat\_d$ ,  $ang\_d$  とした。ただし、目視不能な椎間板を除いて分散を求めた。椎間板の形態に関連する指標値として、 $lam\_d$ ,  $len1\_d$ ,  $len2\_d$ ,  $lrat\_d$ ,  $bod\_d$ ,  $brat\_d$ ,  $ang\_d$  の合計 7 つを用いた。椎間板間の分散を指標値とした理由は、症例ごとに目視可能な頸椎椎間板の本数が異なり、椎間板の高位に依存する解析が困難であったが、椎間板の不整度合を確認する指標として分散が有用であると考えたからである。図 3.5 に  $len1$ ,  $len2$ ,  $lrat$ ,  $lam$ ,  $bod$ ,  $brat$ ,  $ang$  を図示する。

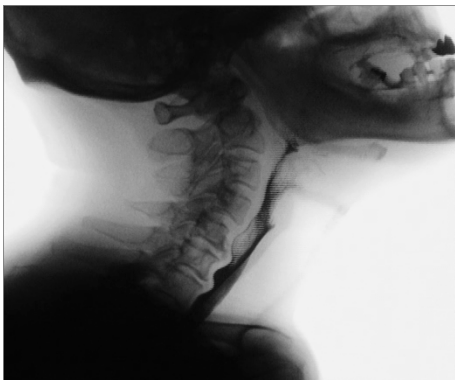


図 3.1: 患者症例の VF フレーム画像

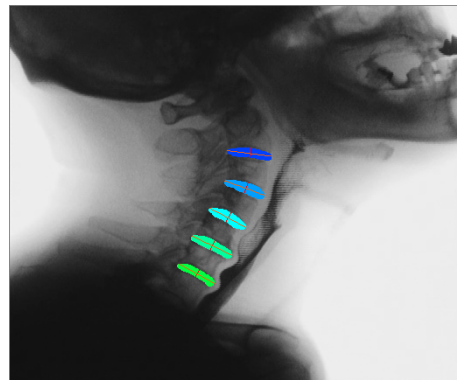


図 3.2: 患者症例の椎間板領域画像



図 3.3: 正常症例の VF フレーム画像

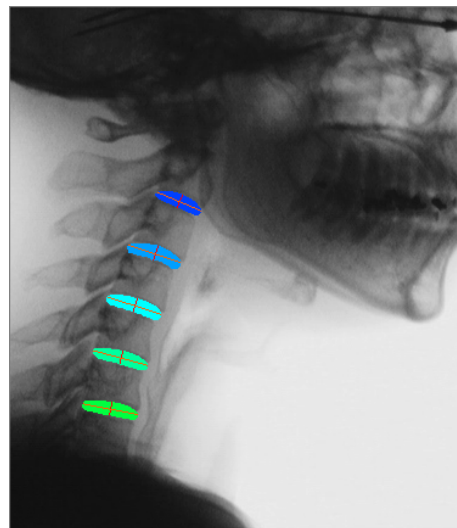


図 3.4: 正常症例の椎間板領域画像

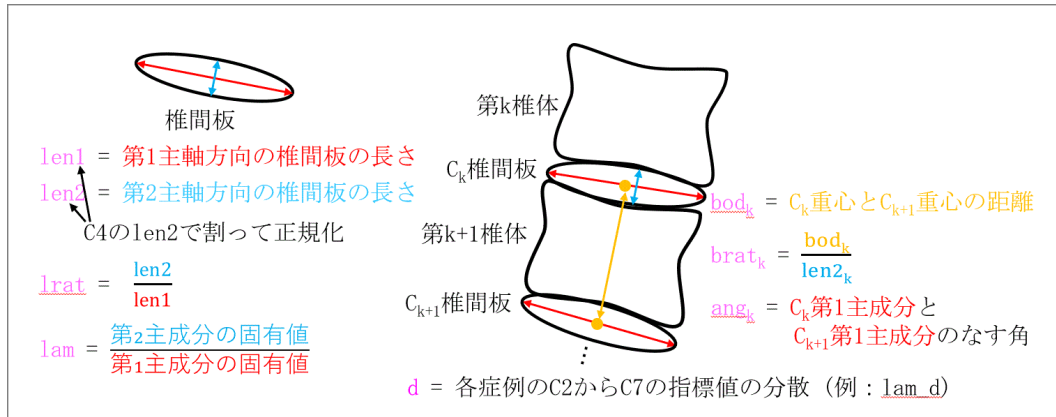


図 3.5: 椎間板の特徴量

## 3.2 椎間板の形態からの患者群と正常群の予測

3.1 節では患者群と正常群の椎間板の特徴量の有意差検定を行ったが、本章では特徴量のみを使用して SVM により患者群か正常群かを予測した。3.1 章で算出した  $lrat_d$ ,  $ang_d$ ,  $lam_d$ ,  $len2_d$ ,  $len1_d$ ,  $brat_d$ ,  $bod_d$  の合計 7 つの特徴量の内、任意の個数を選択して、それぞれの平均が 0、分散が 1 になるように標準化を行ったうえで SVM に入力した。

## 3.3 実験

### 3.3.1 実験条件

3.1 節では、患者群 19 症例と正常群 39 症例を合わせた 58 症例の VF の舌骨最大挙上フレームの目視可能な椎間板領域を著者が手で抽出し、椎間板領域のデータとした。舌骨最大挙上フレーム上の目視可能な椎間板の個数の最大値は、C2 から C7 までの 6 つであった。最小値は、C2 から C4 までの 3 つであった。C1 の形状は C2 以下とは異なるため、本研究では除外した。なお、正解領域データを作成する際、椎間板の高位ごとに異なる画素値を用い、画像から椎間板の高位を特定できるようにした。

3.2 節では、Leave one patient out 法を用いて、58 症例の VF の内 57 症例を学習症例に、残りの 1 症例をテスト症例に割り当てた。

### 3.3.2 実験結果

3.1 節で求めた患者群と正常群における  $lrat_d$ ,  $ang_d$ ,  $lam_d$ ,  $len2_d$ ,  $len1_d$ ,  $brat_d$ ,  $bod_d$  それぞれの平均値と、分散が等しくないと仮定した 2 標本による t-検定で得られた両側 p 値を表 3.1 に示す。

表 3.1: 患者群と正常群における  $lrat\_d$ ,  $ang\_d$ ,  $lam\_d$ ,  $len2\_d$ ,  $len1\_d$ ,  $brat\_d$ ,  $bod\_d$  の有意差検定

特徴量	特徴量の意味	患者群平均値	正常群平均値	両側 p 値
$lrat\_d$	第 1 / 第 2 主軸方向の長さの分散	$1.54 \times 10^{-3}$	$5.85 \times 10^{-4}$	0.0198
$ang\_d$	椎間板同士のなす角の分散	7.93	4.87	0.0331
$lam\_d$	第 1 / 第 2 主成分の固有値の分散	$2.99 \times 10^{-4}$	$1.19 \times 10^{-4}$	0.0406
$len2\_d$	第 2 主軸方向の長さの分散	$3.50 \times 10^{-2}$	$1.10 \times 10^{-2}$	0.0482
$len1\_d$	第 1 主軸方向の長さの分散	$1.45 \times 10^{-1}$	$5.45 \times 10^{-2}$	0.0567
$brat\_d$	椎体 / 椎間板の厚さの分散	$1.29 \times 10^{-3}$	$5.79 \times 10^{-4}$	0.0599
$bod\_d$	椎体の厚さの分散	$9.86 \times 10^{-2}$	$3.30 \times 10^{-2}$	0.0839

3.2 節で、3.1 節で算出した p 値が小さい 3 つの特徴量を選択した際、F 値ベースでの分類精度が 0.438 となった。  $ang\_d$ ,  $lam\_d$ ,  $brat\_d$  の 3 つを選択した際、F 値が 0.722 となった。

### 3.4 考察

3.1 節では C2 から C7 の分散に焦点をあてて有意差検定を行った。表 3.1 の 7 つの特徴量の内、第 1, 第 2 主軸方向の長さの比率の分散を表す  $lrat\_d$  の p 値が最小となった。このことから患者群の椎間板の形状の不整度合が正常群より大きいことが示唆される。患者群は頸椎固定術を受ける予定の患者であり、頸椎に何らかの疾患を抱えているため、頸椎椎体の扁平化や骨棘形成、椎間板の突出が見られる可能性があり、形態の不整度合が大きいという医学知識と合致する。3.2 節では、SVM を用いて患者群と正常群を F 値 = 0.722 という精度で識別できた。このことから、患者群と正常群の椎間板の形態に差異があることが示された。本章では、VF から椎間板の形態解析を行うための指標値を作成し、頸椎疾患の患者群と正常群の椎間板の形態の差異を示したが、今後、本章で提案した指標値を用いて、嚙下障害と椎間板の形態の関連性を調べることも可能となる。嚙下障害の原因は静的障害(器質的障害)と動的障害(機能的障害)に大きく分けられ、静的障害には前縦靭帯骨化症、頸椎症の術後、変形性脊椎症などの頸椎疾患が含まれていることから [1], 頸椎の形態学的異常は嚙下障害の原因となり得る。したがって、頸椎疾患の患者は嚙下障害を合併している可能性があり、場合によっては、頸椎疾患の有無による椎間板の形態の差異と、嚙下障害の有無による椎間板の形態の差異が関連している可能性もある。すなわち、本章の提案手法だけでなく実験結果も応用できる可能性がある。なお、別の応用先として、頸椎症、頸椎椎間板ヘルニア、頸椎椎体骨折、後縦靭帯骨化症など頸椎疾患ごとの椎間板の形態解析や、椎間板の形態からの頸椎疾患名の予測が挙げられる。

なお、本論文で用いた VF データセットは匿名化されており、氏名、年齢、性別、臨床症状、VF 以外の検査結果、臨床診断名を知らない状態で解析を行った。したがって、本章では頸椎疾患ごとの椎間板の形態解析を行うことはできなかった。また、いずれの実験参加者が嚙下

障害を合併しているかわからず，嚙下障害と椎間板の形態の関連性を直接調べることができなかった．ただし，本論文の提案手法を嚙下障害の有無が明確なVFデータセットに適用することで，研究の目的を達成できると予想される．その際に，大規模なデータセットを用いることで，嚙下障害と頸椎の形動態の関連性を詳細に解析することができる．

## 第4章 深層学習を用いた椎間板領域の抽出

3章では、VFを用いて頸椎椎間板の形態解析を行う手法を提案し、また、提案手法を用いて、頸椎疾患の有無による椎間板の形態の差異を調べた。今後の課題として、大規模なVFデータセットに提案手法を適用し、嚙下障害と椎間板の形態の関連性を調べることを挙げた。しかしながら、提案手法では手動による領域抽出が必要で、少数のVFデータセットに対しては有用であるが、解析できる症例数に限りがある。したがって、大規模なVFデータセットに対応するために、自動的に高精度に椎間板領域を抽出することが必要となる。本章では、現在最も抽出精度が高いと考えられる深層学習を基本とする椎間板抽出手法を述べる。4.1節では、深層学習による抽出結果にモルフォロジー演算等による後処理を施すことによって精度を向上させる手法を提案する。

### 4.1 モルフォロジー演算による後処理を用いた方法

#### 4.1.1 頸部マスクの抽出

VFの各フレーム画像は図4.1に示すように人体領域と背景領域から構成される。人体領域には頭部領域、頸部領域、肩領域が含まれる。齋藤らの手法[38]を使用して、以下の手順で椎間板領域を含む頸部マスクを抽出する。本論文で用いたVFデータセットでは、人体領域の画素値が低く、背景領域の画素値が高い傾向にあるため、判別分析法[43]により人体領域と背景領域を二分する。判別分析法とは別名大津の二値化とも呼ばれ、クラス間分散とクラス内分散の比である分離度が最大となるしきい値を自動で求める手法である。生成された画像にモルフォロジー演算[44]のopeningフィルタを10回適用し、平滑化を行う。Openingフィルタとは、モルフォロジー演算のerosionフィルタとdilationフィルタを組合わせたものである。Erosionフィルタを用いて小領域削除を行った後、小さくなってしまった物体の面積をdilationフィルタを用いて復活させることにより、物体の大きさを変えずにノイズを削減できる。Openingフィルタの構造要素を半径3ピクセルの円形画像とする。一部の症例では、平滑化後の画像に画像端や眼周辺の孤立領域が含まれているため、全ての症例で面積が2番目以下の人体領域と面積が3番目以下の背景領域を削除する。孤立領域削除後の画像における、人体領域と背景領域の境界線を8近傍で算出する。孤立領域削除後の背景領域に8近傍ラベリングを施し、面積が大きい2領域を抽出する。抽出した2領域それぞれの重心を求め、重心同士を直線で結ぶ。直線の上下で人体領域と背景領域の境界線を二分する。二分する前の境界線画像を用いて、人体領域のそれぞれの画素における境界線からの距離を求め、距離値を画素値とする画像を作成する。同じく二分する前の境界線画像を用いて、背景領域のそれぞれの

画素における境界線からの距離画像を作成する。背景領域の距離画像の画素値に  $-1$  をかけ、人体領域の距離画像の画素値と合計することで、画像全体の距離画像を作成する。画像全体の距離画像の上端と下端から順に 1 点ずつ選択し、それらの点同士を結ぶ線分を求めたうえで、線分上の画素値の平均値を計算する。上端と下端の全ての点において画素値の平均値を計算し、平均値が最大となるような線分を首中心線とする。画像内のそれぞれの画素における首中心線からの距離画像を作成する。首中心線から境界線への最短距離を  $dist_{min}$  とし、首中心線から境界線上の各画素への距離を  $dist_A$  としたうえで、 $dist_{min} \leq dist_A \leq dist_{min} + 50$  を満たす画素のみを削除する。残りの境界線を ㉑下顎骨、㉒後頭骨、㉓前胸部、㉔上背部の 4 つの領域に区分する。㉑下顎骨と㉒後頭骨から合計 3 点を選択し、境界線に random sample consensus method (RANSAC)[40] を適用することで頭部に放物線を当てはめる。ただし、㉑下顎骨と㉒後頭骨からそれぞれ 1 点以上を選択するという条件を課し、RANSAC を 2000 回適用した中で最良の放物線を選択する。同様に、㉓前胸部と㉔上背部に RANSAC を適用することで肩部に放物線を当てはめる。原画像の内、2 本の放物線に上下方向に挟まれた領域のみを選択し、頸部領域の画像を作成する。頸部領域の画像に判別分析法を適用する。生成された画像にモルフォロジー演算の closing フィルタを 4 回、opening フィルタを 6 回、closing フィルタを 2 回順に適用することで、平滑化を行い、頸部マスクとする。Closing フィルタとは、モルフォロジー演算の dilation フィルタと erosion フィルタを組合わせたものである。Closing フィルタと opening フィルタの構造要素を半径 3 ピクセルの円形画像とする。

図 4.2 に判別分析法を適用後の人体領域と背景領域を、図 4.3 に平滑化後の人体領域と背景領域を、図 4.4 に孤立領域削除後の人体領域と背景領域を、図 4.5 に境界線画像を、図 4.6 に背景の最大 2 領域の重心を直線で結んだ画像を、図 4.7 に直線の上下で二分した境界線画像を、図 4.8 に人体領域の距離画像のカラーマップを、図 4.9 に背景領域の距離画像のカラーマップを、図 4.10 に画像全体の距離画像のカラーマップを、図 4.11 に首中心線と領域分割された境界線画像を、図 4.12 に首中心線からの距離画像のカラーマップを、図 4.13 に領域分割した境界線画像を、図 4.14 に頭部と肩部の放物線を原画像に重畳表示した画像を、図 4.15 に頸部領域の画像を、図 4.16 に判別分析法を適用後の頸部領域を、図 4.17 に頸部マスク画像を示す。なお、図 4.5、図 4.6、図 4.7、図 4.11、図 4.12、図 4.13、図 4.14 では、視認性を重視し、線や点を実際より太く加工している。

#### 4.1.2 深層学習とモルフォロジー演算を用いた椎間板領域の抽出

作成した頸部マスクに以下の手順で深層学習を適用し、椎間板候補領域を抽出する。学習に先立ち、各症例の VF から舌骨最大挙上フレームを 1 フレーム選定し、舌骨最大挙上フレーム上の目視可能な椎間板領域を手動で抽出し、教師データあるいは Ground truth とする。学習段階では、学習症例の教師データの全ての椎間板画素を中心とする  $S_p \times S_p$  ピクセルのパッチを作成し、正例とする。頸部マスク内の非椎間板画素を中心とするパッチを正例と同数作成し、負例とする。ただし、椎間板画素と比較して非椎間板画素の数が十分多いことから、椎間板画素の近傍  $S_p$  ピクセルを除外して負例のパッチを作成する。テスト段階では、頸部マスク内の全ての画素を中心とするパッチを生成されたモデルに入力し、尤度  $prob$  で正例に分類



図 4.1: VF フレーム画像



図 4.2: 判別分析法を適用後の人体領域と背景領域



図 4.3: 平滑化後の人体領域と背景領域



図 4.4: 孤立領域削除後の人体領域と背景領域



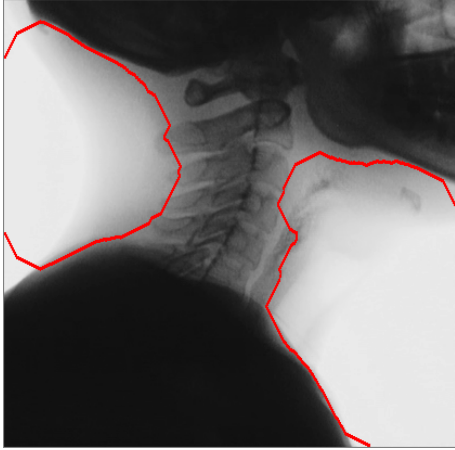


図 4.5: 人体領域と背景領域の境界線画像

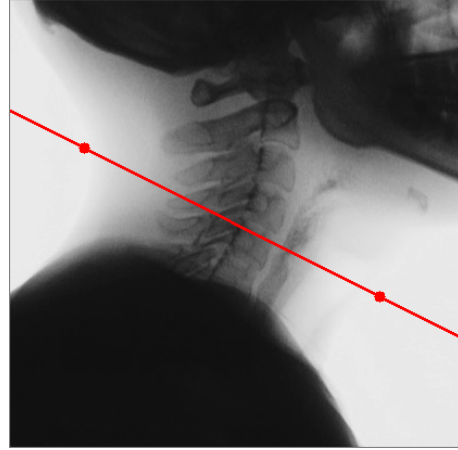


図 4.6: 背景の最大 2 領域の重心を結ぶ直線

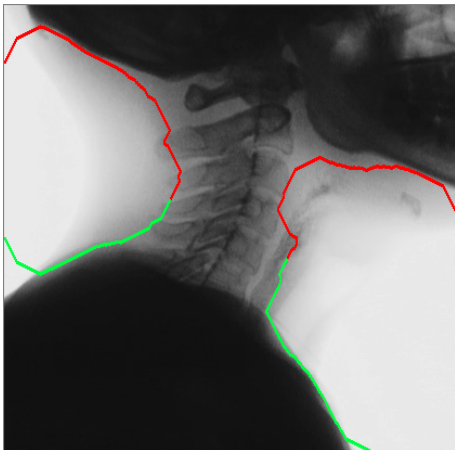


図 4.7: 直線の上下で二分した境界線画像

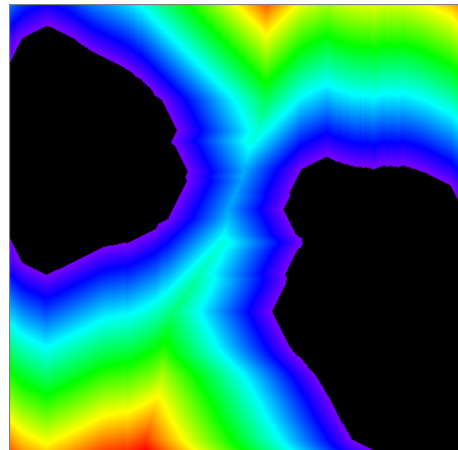


図 4.8: 人体領域の距離画像のカラーマップ

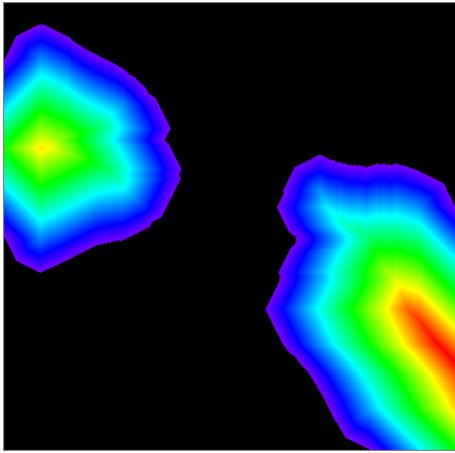


図 4.9: 背景領域の距離画像のカラーマップ

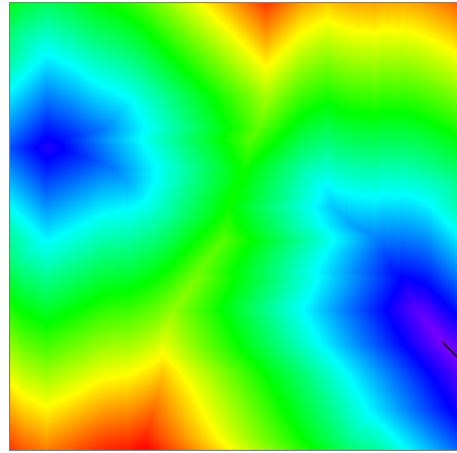


図 4.10: 画像全体の距離画像のカラーマップ

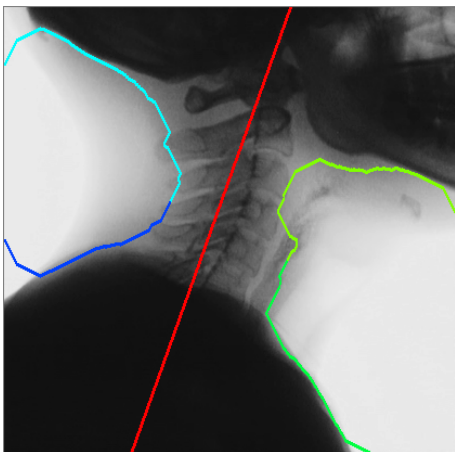


図 4.11: 首中心線と領域分割された境界線画像

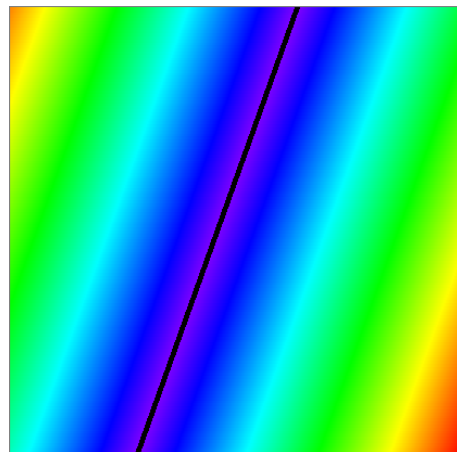


図 4.12: 首中心線からの距離画像のカラーマップ

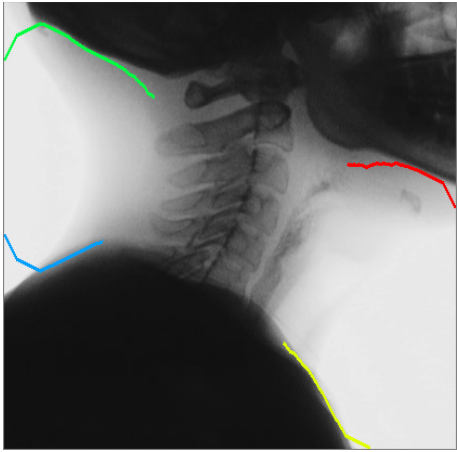


図 4.13: 領域分割した境界線画像

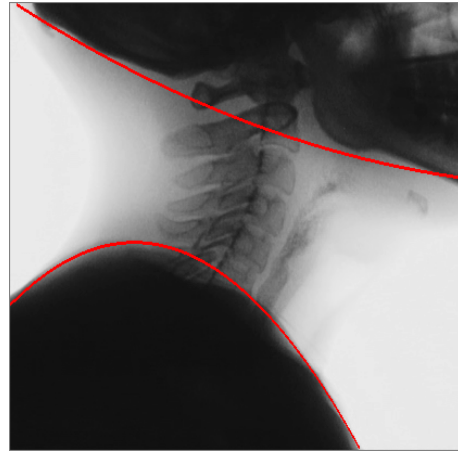


図 4.14: 頭部と肩部の放物線の重畳表示

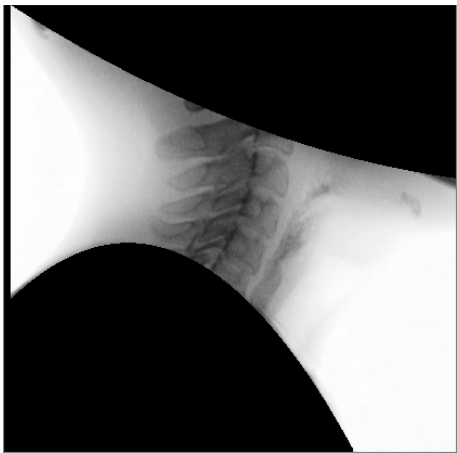


図 4.15: 頸部領域

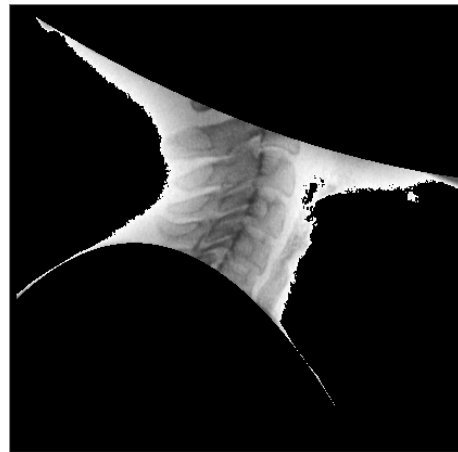


図 4.16: 判別分析法を適用後の頸部領域

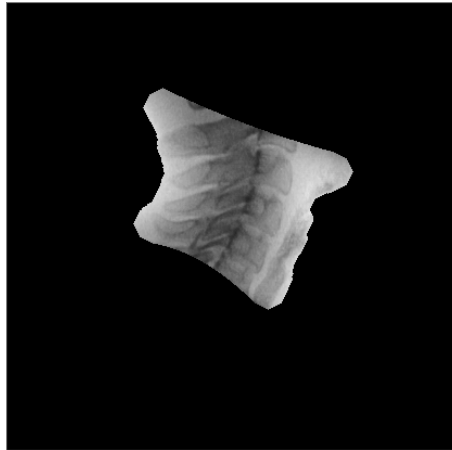


図 4.17: 頸部マスク画像

されたパッチの中心画素を椎間板候補画素とする．パッチベースの深層学習では実際の椎間板より太い椎間板候補領域が抽出される傾向にあり，椎体や棘突起を誤抽出した偽陽性も多い．したがって，4.1.3.9 節，4.1.3.11 節，4.1.3.12 節では，椎間板の形状が円盤状であるという特性を利用して，モルフォロジー演算を用いた後処理で偽陽性の削減を行う．

各テスト症例の Groun truth と抽出した椎間板候補領域を比較し，画素単位の precision, recall, accuracy, F 値 (F-measure) を求め，テスト症例間の平均値を計算する．なお，表 4.1 に示す通り，真のクラスと識別クラスの関係性から，画像内の画素は真陽性 (true positives, TP), 偽陰性 (false negatives, FN), 偽陽性 (false positives, FP), 真陰性 (true negatives, TN) の 4 つに分けられる． $precision = \frac{TP}{TP+FP}$ ,  $recall = \frac{TP}{TP+FN}$ ,  $accuracy = \frac{TP+TN}{TP+FP+FN+TN}$ ,  $F \text{ 値} = \frac{2 \times precision \times recall}{precision+recall}$  と定義される．偽陽性が多くなれば precision が下がり，偽陰性が多くなれば recall が下がるため，precision と recall はトレードオフの関係にある．F 値は precision と recall の調和平均を表す．

表 4.1: 混同行列

		識別クラス	
		正例	負例
真のクラス	正例	TP	FN
	負例	FP	TN

### 4.1.3 実験

#### 4.1.3.1 パッチサイズを比較した実験

##### 4.1.3.1.1 実験条件

患者群 11 症例の VF を用い、Leave one patient out 法により 10 症例を学習に、残りの 1 症例をテストに割り当てて実験を行った。椎間板領域のデータとして、舌骨最大挙上フレーム上の目視可能な椎間板を選定し、椎間板領域の垂直方向のおおよその中心画素を手動抽出したものをを用いた。すなわち、それぞれの椎間板を一本の線で表したものを椎間板領域のデータとした。正例、負例のパッチを深層学習に入力する前に、画素値正規化補正と回転補正を施した。ただし、画素値正規化補正では、それぞれのパッチの最大画素値が 255、最小画素値が 0 となるように線形に正規化補正を行った。回転補正では、パッチを  $-30$  度から  $30$  度まで  $5$  度刻みで回転した中で、椎間板が最も水平方向に近いパッチを選択した。補正後のパッチを深層学習に入力し、モデルを作成した。深層学習の尤度  $prob$  を  $0.5$  とした。パッチサイズ  $S_p$  を  $15$  から  $23$  まで  $2$  刻みで変更し、椎間板抽出精度を比較した。深層学習モデルとして LeNet [41] を用いた。LeNet は、畳み込みニューラルネットワーク (CNN) の一種であり、主に画像のクラス分類に用いられる。ネオコグニトロンを原型に、特徴抽出を行う畳み込み層と、位置ずれを許容するプーリング層を交互に配置している。学習時に重みや畳み込み層のカーネルを更新する際に、誤差逆伝搬法を用いている。本節では、症例数が 11 症例と少なかつたため、パッチベースでより多くの学習画像枚数を確保したいと考え、パッチベースの深層学習を使用した。その中でも、小さいパッチサイズで高精度の分類が可能な LeNet を選択した。学習サンプルの内評価に用いる割合を  $25\%$  とし、テストに用いる割合を  $0\%$  とした。エポック数を  $30$  回、学習率を  $0.01$ 、バッチサイズを  $32$ 、畳み込み層の重みの初期化フィルタを Xavier、損失関数の最小化手法を確率的勾配降下法とした。本節の実験に使用した計算機は、OS Linux 64 ビット Ubuntu 16.04、メモリ  $15.6$  GB、プロセッサ Intel® Core™ i7-8700 CPU @  $3.20$  GHz  $\times 12$ 、グラフィック GeForce GTX 1080/PCIe/SSE2 であり、使用したフレームワークは NVIDIA® DIGITS-6.1.1[42] であった。[78][46][47][48][49][50] を参考にプログラムを作成した。

##### 4.1.3.1.2 実験結果

図 4.18 に学習症例の VF フレーム画像を、図 4.19 に同一症例の教師画像を示す。 $S_p$  を変更して得られた平均椎間板抽出精度を表 4.2 に示す。 $S_p$  を  $23$  としたとき、precision, accuracy, F 値が最大となった。 $S_p$  を  $19$  としたとき、recall が最大となった。代表的な症例の VF フレーム画像を図 4.20 に、Ground truth を図 4.21 に、 $S_p$  を 5 通りに変更して得られた結果画像と同一症例の F 値を図 4.22, 図 4.23, 図 4.24, 図 4.25, 図 4.26 に示す。ただし、結果画像の真陽性、偽陰性、偽陽性の画素を順に緑色、青色、赤色で表す。



図 4.18: 学習症例の VF フレーム画像

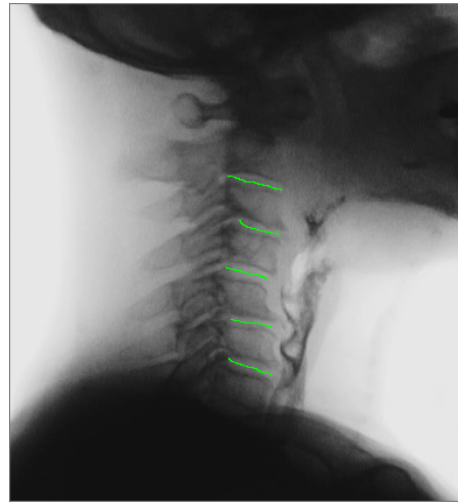


図 4.19: 椎間板領域の教師画像



図 4.20: テスト症例の VF フレーム画像

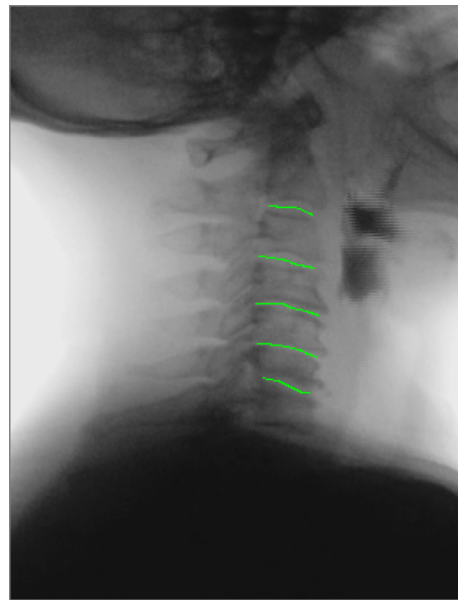


図 4.21: 椎間板領域の Ground truth

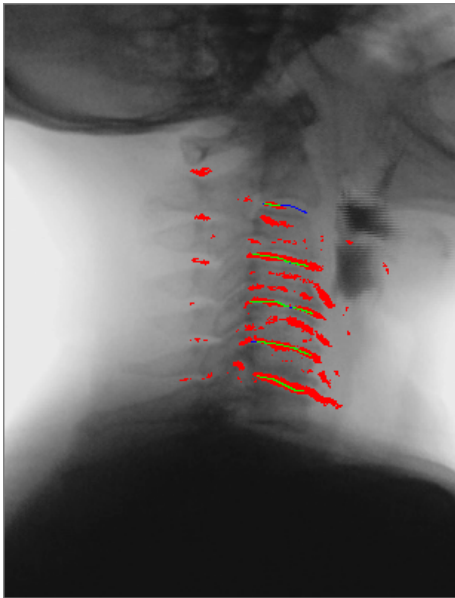


図 4.22: 4.1.3.1 節で  $S_p = 15$  とした結果画像 ( $F$  値 = 0.173)

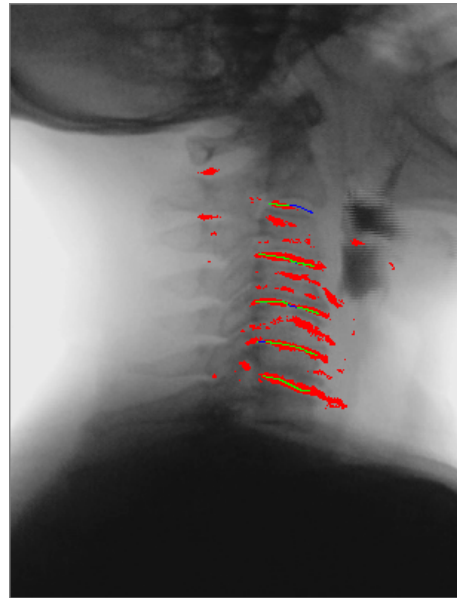


図 4.23: 4.1.3.1 節で  $S_p = 17$  とした結果画像 ( $F$  値 = 0.181)

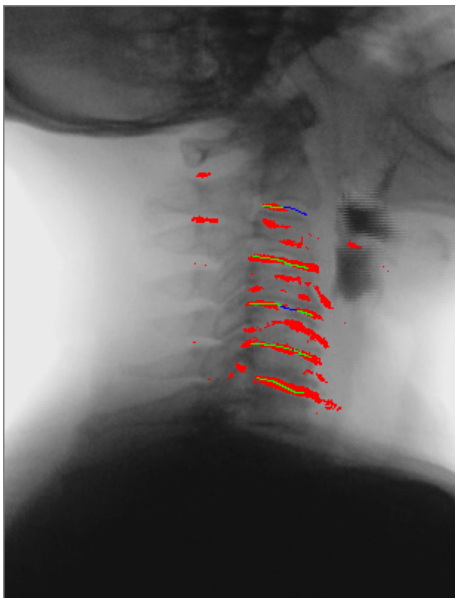


図 4.24: 4.1.3.1 節で  $S_p = 19$  とした結果画像 ( $F$  値 = 0.188)

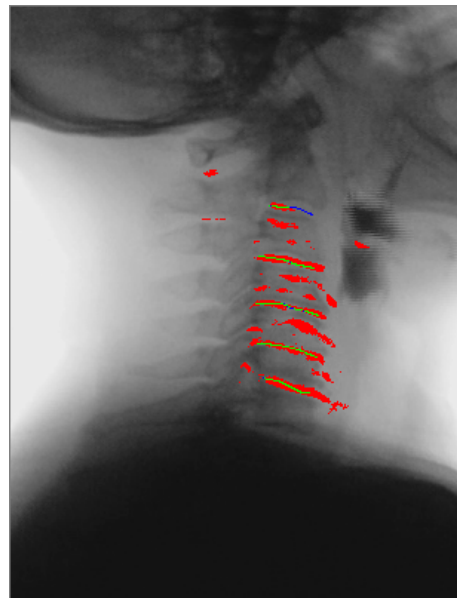


図 4.25: 4.1.3.1 節で  $S_p = 21$  とした結果画像 ( $F$  値 = 0.208)

表 4.2: 4.1.3.1 節で  $S_p$  を 15 から 23 まで 2 刻みで変更して得られた平均椎間板抽出精度

$S_p$	precision	recall	accuracy	F 値
15	0.0782	0.829	0.989	0.143
17	0.0807	0.827	0.989	0.147
19	0.0865	<b>0.832</b>	0.990	0.156
21	0.0840	0.824	0.990	0.152
23	<b>0.0874</b>	0.828	<b>0.990</b>	<b>0.158</b>

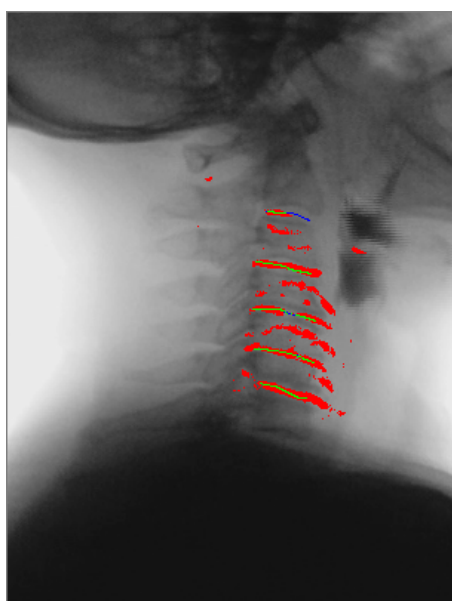


図 4.26: 4.1.3.1 節で  $S_p = 23$  とした結果画像 ( $F$  値 = 0.201)

#### 4.1.3.2 画素値正規化補正と回転補正の有無を比較した実験

##### 4.1.3.2.1 実験条件

$S_p$  を 15 に固定し，画素値正規化補正と回転補正の有無を変更して椎間板抽出精度を比較する 4 通りの実験を行った。ただし，4.1.3.1 節と同一の学習，テスト症例，椎間板領域データ，深層学習モデル，計算機，フレームワークを用いた。

##### 4.1.3.2.2 実験結果

4 通りの実験で得られた平均椎間板抽出精度を表 4.3 に示す。ただし，画素値正規化補正と回転補正を両方行った実験結果として表 4.2 の結果を再掲した。画素値正規化補正を行い回転補正を行わなかった場合に平均椎間板抽出精度が最大となった。図 4.20 にそれぞれの条件を適用して得られた結果画像と同一症例の  $F$  値を図 4.27，図 4.28，図 4.29 に示す。



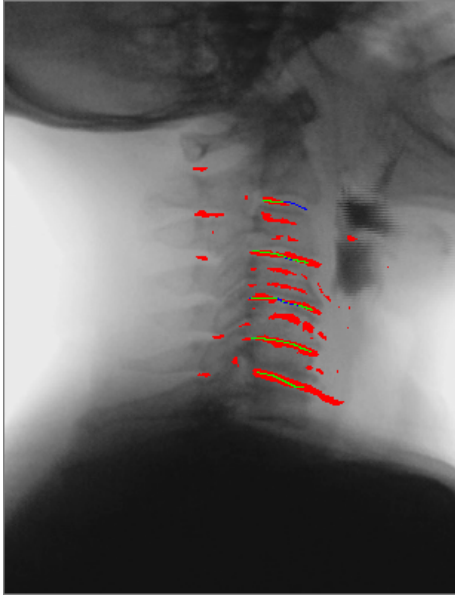


図 4.27: 4.1.3.2 節で  $S_p = 15$ , 画素値正規化補正あり, 回転補正なしとした結果画像 ( $F$  値 = **0.197**)

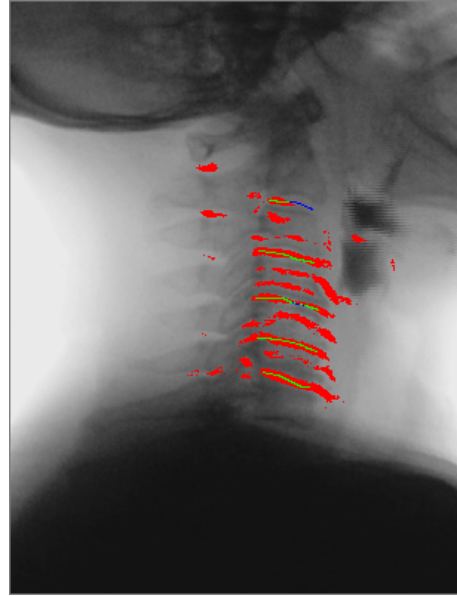


図 4.28: 4.1.3.2 節で  $S_p = 15$ , 画素値正規化補正なし, 回転補正ありとした結果画像 ( $F$  値 = 0.149)

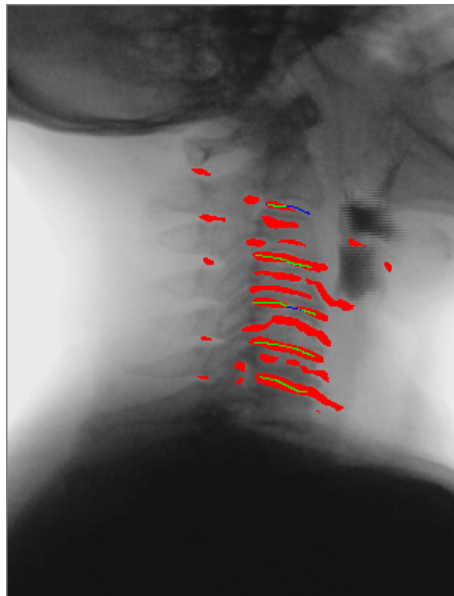


図 4.29: 4.1.3.2 節で  $S_p = 15$ , 画素値正規化補正なし, 回転補正なしとした結果画像 ( $F$  値 = 0.124)

表 4.3: 4.1.3.2 節で画素値正規化補正と回転補正の有無を変更して得られた平均椎間板抽出精度

$S_p$	画素値正規化補正	回転補正	precision	recall	accuracy	F 値
15	あり	あり	0.0782	0.829	0.989	0.143
15	あり	なし	<b>0.0975</b>	<b>0.830</b>	<b>0.991</b>	<b>0.174</b>
15	なし	あり	0.0749	0.802	0.988	0.136
15	なし	なし	0.0828	0.804	0.989	0.149

### 4.1.3.3 椎間板領域全体のデータを用いてパッチサイズを比較した実験

#### 4.1.3.3.1 実験条件

4.1.3.1 節と 4.1.3.2 節では、椎間板の垂直方向の中心画素のみを教師データとして用いていたため、学習データに用いるパッチの枚数が少なく、テストで椎体や棘突起を偽陽性として誤抽出してしまった。また、中心画素のみを Ground truth と見なして画素単位の抽出精度を求めていたため、真陽性の画素が少なく、F 値が低く出てしまった。これらの問題を解決するために、本節では、舌骨最大挙上フレーム上の目視可能な椎間板領域全体を手動で抽出し、椎間板領域のデータとした。パッチサイズ  $S_p$  を 15 から 35 まで 2 刻みで変更し、各パッチに画素値正規化補正と回転補正を施したうえで深層学習に入力し、椎間板抽出精度を比較した。ただし、4.1.3.1 節と同一の学習、テスト症例、深層学習モデル、計算機、フレームワークを用いた。

#### 4.1.3.3.2 実験結果

図 4.78 と同一の症例における椎間板領域全体の教師画像を図 4.30 に示す。図 4.19 と比較して椎間板領域が太く抽出され、椎間板領域の画素数が増えている。 $S_p$  を変更して得られた平均椎間板抽出精度を表 4.4 に示す。 $S_p$  を 19 としたとき、precision, accuracy, F 値が最大となった。 $S_p$  を 29 としたとき、recall が最大となった。4.1.3.1 節と比較して Ground truth の椎間板領域の画素数が多いため、F 値が向上する傾向にあった。図 4.20 と同一症例の Ground truth を図 4.31 に、 $S_p$  を 11 通りに変更して得られた結果画像と同一症例の F 値を図 4.32, 図 4.33, 図 4.34, 図 4.35, 図 4.36, 図 4.37, 図 4.38, 図 4.39, 図 4.40, 図 4.41, 図 4.42 に示す。

### 4.1.3.4 椎間板領域全体のデータを用いて画素値正規化補正と回転補正の有無を比較した実験

#### 4.1.3.4.1 実験条件

4.1.3.3 節で  $S_p$  を 19 とした実験で precision, accuracy, F 値が最大となったため、本節では  $S_p$  を 19 に固定し、画素値正規化補正と回転補正の有無を変更して椎間板抽出精度を比較する 4 通りの実験を行った。ただし、4.1.3.1 節と同一の学習、テスト症例、深層学習モデル、計算機、フレームワークを用い、4.1.3.3 節と同一の椎間板領域のデータを用いた。

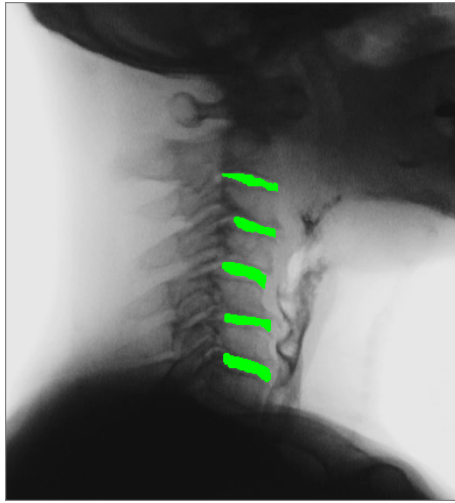


図 4.30: 椎間板領域全体の教師画像

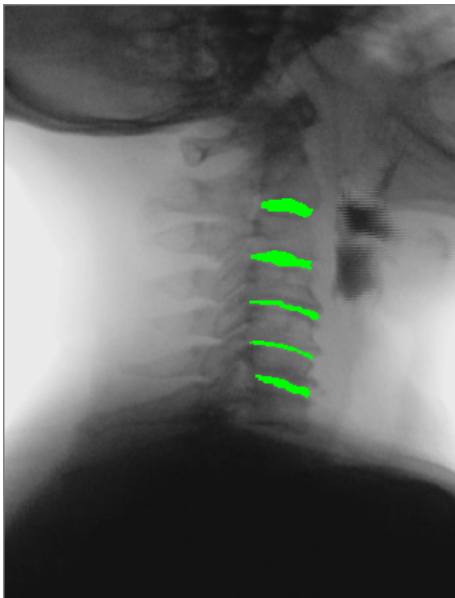


図 4.31: 椎間板領域全体の Ground truth

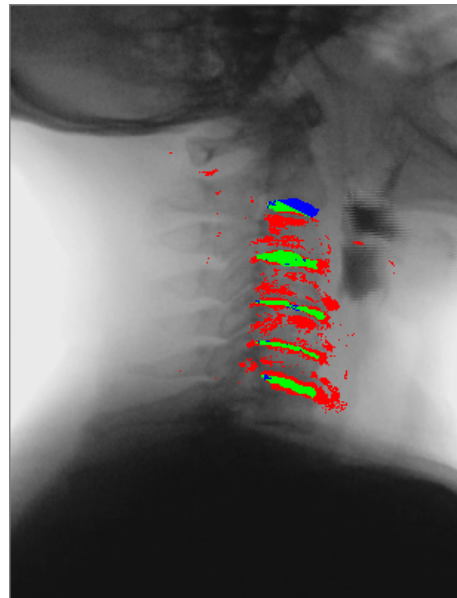


図 4.32: 4.1.3.3 節で  $S_p = 15$  とした結果画像 ( $F$  値 = 0.390)

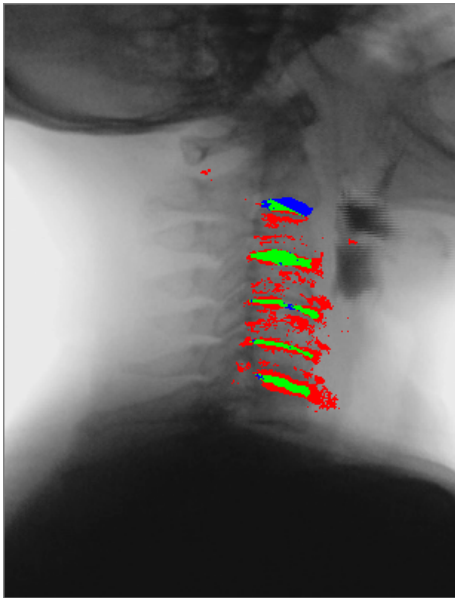


図 4.33: 4.1.3.3 節で  $S_p = 17$  とした結果画像 ( $F$  値 = 0.410)

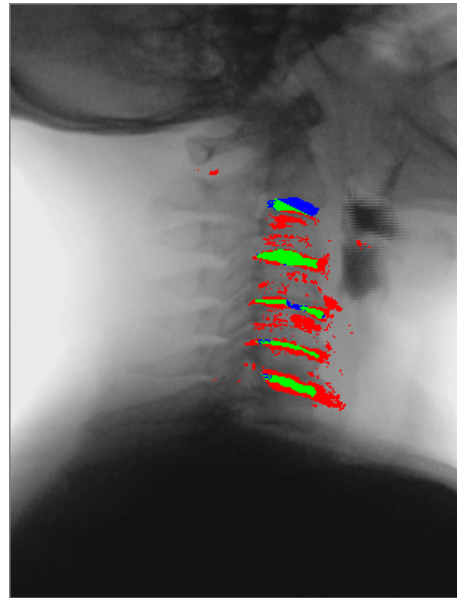


図 4.34: 4.1.3.3 節で  $S_p = 19$  とした結果画像 ( $F$  値 = **0.424**)

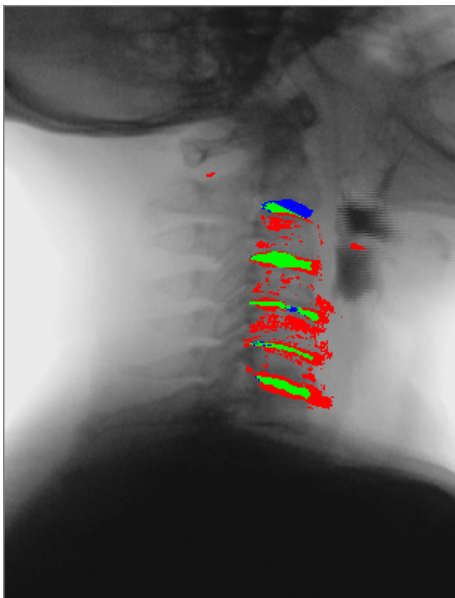


図 4.35: 4.1.3.3 節で  $S_p = 21$  とした結果画像 ( $F$  値 = 0.418)

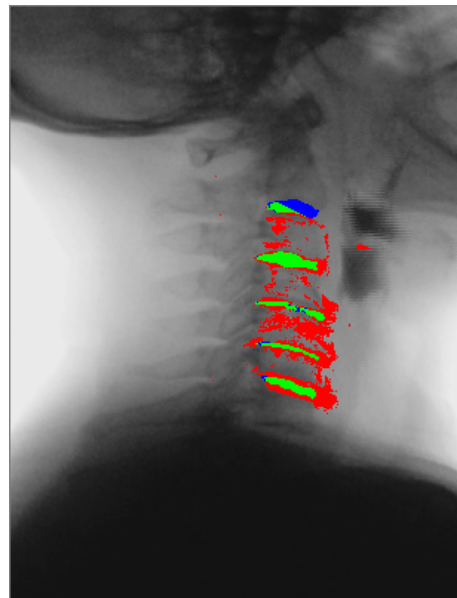


図 4.36: 4.1.3.3 節で  $S_p = 23$  とした結果画像 ( $F$  値 = 0.395)

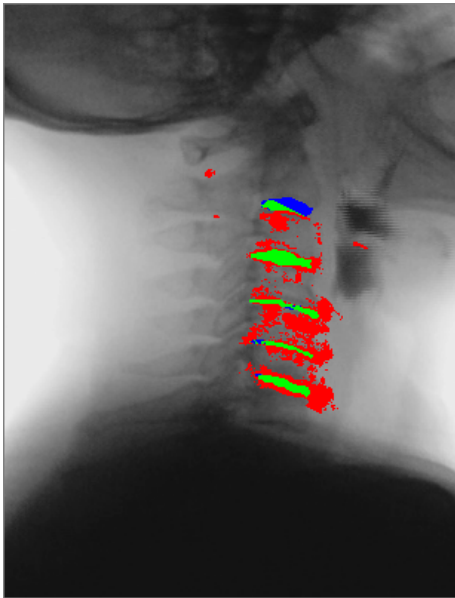


図 4.37: 4.1.3.3 節で  $S_p = 25$  とした結果画像 ( $F$  値 = 0.363)

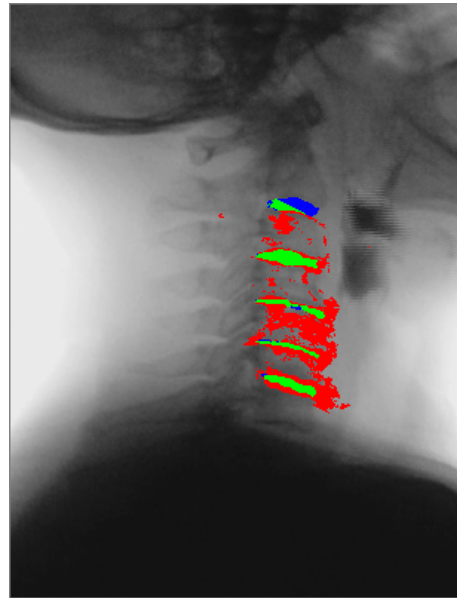


図 4.38: 4.1.3.3 節で  $S_p = 27$  とした結果画像 ( $F$  値 = 0.374)

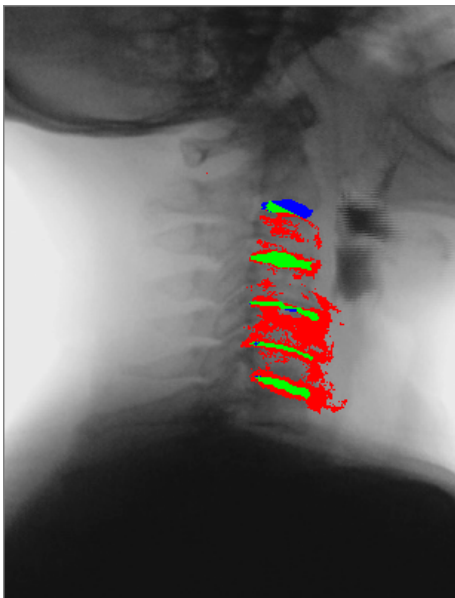


図 4.39: 4.1.3.3 節で  $S_p = 29$  とした結果画像 ( $F$  値 = 0.342)

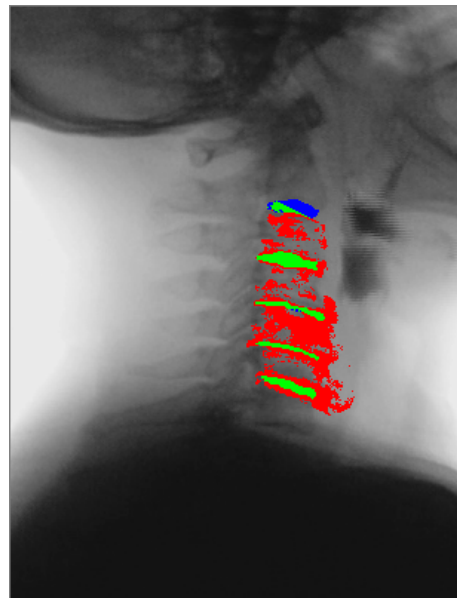


図 4.40: 4.1.3.3 節で  $S_p = 31$  とした結果画像 ( $F$  値 = 0.332)

表 4.4: 4.1.3.3 節で  $S_p$  を 15 から 35 まで 2 刻みで変更して得られた平均椎間板抽出精度

$S_p$	precision	recall	accuracy	F 値
15	0.283	0.814	0.984	0.418
17	0.301	0.814	0.985	0.437
19	<b>0.315</b>	0.828	<b>0.986</b>	<b>0.454</b>
21	0.314	0.832	0.986	0.453
23	0.305	0.834	0.986	0.444
25	0.281	0.840	0.984	0.419
27	0.268	0.840	0.983	0.405
29	0.244	<b>0.847</b>	0.981	0.377
31	0.230	0.843	0.980	0.360
33	0.209	0.844	0.977	0.334
35	0.198	0.839	0.976	0.319

#### 4.1.3.4.2 実験結果

4 通りの実験で得られた平均椎間板抽出精度を表 4.5 に示す。ただし、画素値正規化補正と回転補正を両方行った実験結果として表 4.4 の結果を再掲した。画素値正規化補正を行い回転補正を行わなかった場合に precision, accuracy, F 値が最大となり、画素値正規化補正と回転補正を両方行った場合に recall が最大となった。図 4.20 にそれぞれの条件を適用して得られた結果画像と同一症例の F 値を図 4.43, 図 4.44, 図 4.45 に示す。

表 4.5: 4.1.3.4 節で画素値正規化補正と回転補正の有無を変更して得られた平均椎間板抽出精度

$S_p$	画素値正規化補正	回転補正	precision	recall	accuracy	F 値
19	あり	あり	0.315	<b>0.828</b>	0.986	0.454
19	あり	なし	<b>0.344</b>	0.812	<b>0.988</b>	<b>0.481</b>
19	なし	あり	0.291	0.785	0.985	0.422
19	なし	なし	0.339	0.765	0.987	0.464

#### 4.1.3.5 深層学習の尤度を比較した実験

##### 4.1.3.5.1 実験条件

これまでの実験では、深層学習のテスト段階で、尤度 0.5 で正例に分類されたパッチの中心画素を椎間板候補画素としていた。本節では、尤度  $prob$  を 0.5, 0.85, 0.9, 0.95, 0.995, 0.999, 0.9999 の 7 通りに変更し、椎間板抽出精度を比較した。ただし、 $S_p$  を 19 に固定し、各パッチに画素値正規化補正と回転補正を施したうえで深層学習に入力した。4.1.3.1 節と同

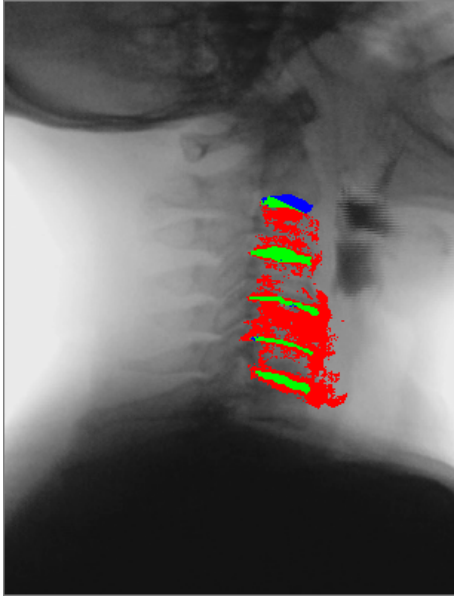


図 4.41: 4.1.3.3 節で  $S_p = 33$  としたときの椎間板抽出結果画像 ( $F$  値 = 0.318)

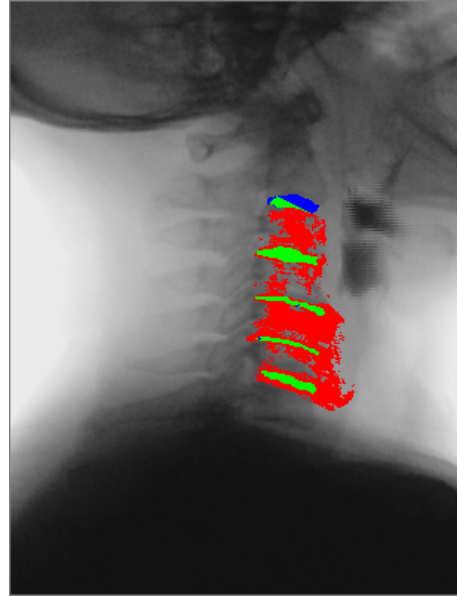


図 4.42: 4.1.3.3 節で  $S_p = 35$  としたときの椎間板抽出結果画像 ( $F$  値 = 0.301)

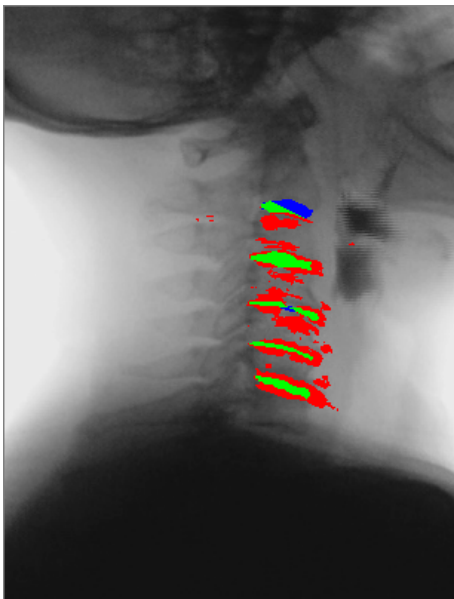


図 4.43: 4.1.3.4 節で画素値正規化補正あり，回転補正なしとした結果画像 ( $F$  値 = 0.406)

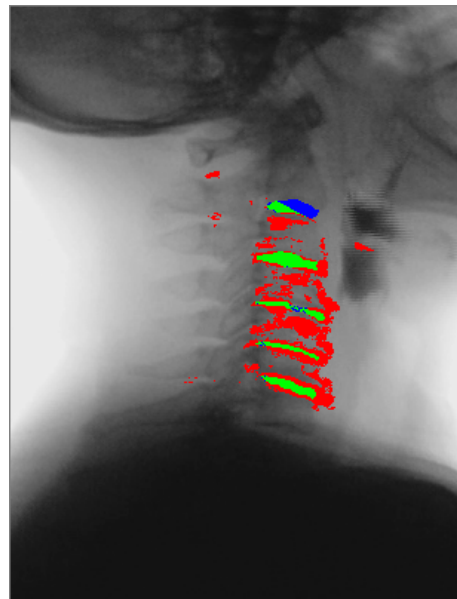


図 4.44: 4.1.3.4 節で画素値正規化補正なし，回転補正ありとした結果画像 ( $F$  値 = 0.363)

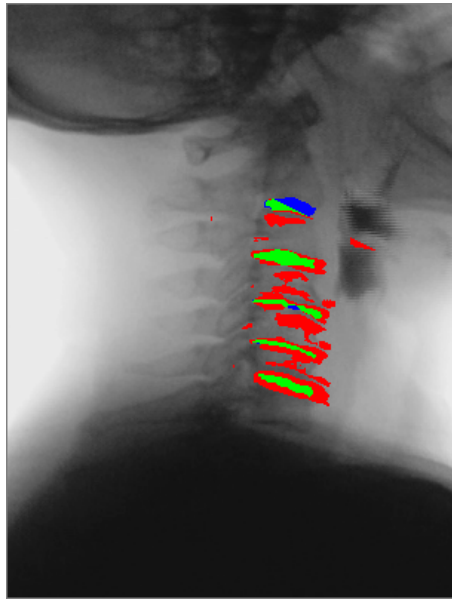


図 4.45: 4.1.3.4 節で画素値正規化補正なし，回転補正なしとした結果画像 ( $F$  値 = 0.403)

一の学習，テスト症例，深層学習モデル，計算機，フレームワークを用い，4.1.3.3 節と同一の椎間板領域のデータを用いた。

#### 4.1.3.5.2 実験結果

表 4.6 に深層学習の尤度を変更して得られた平均椎間板抽出精度を示す。ただし，尤度  $prob$  を 0.5 とした実験結果として表 4.4 の結果を再掲した。尤度を大きくした場合，precision と accuracy は増加し，recall は減少する傾向にあった。一因として，尤度が上がるにつれ負例に分類されるパッチが多くなるため，真陽性と偽陽性が減少し，偽陰性が増加することが考えられる。F 値は，尤度が 0.5 から 0.995 まで上がるにつれ増加し，0.995 から 0.9999 まで上がるにつれ減少した。尤度が上がるにつれ，precision は増加，recall は減少し，トレードオフの関係にあったが，precision と recall の調和平均である F 値は上に凸な関数となった。ただし，最適な尤度はデータセットや椎間板領域のデータ，深層学習のパラメータに依存するため，本来であれば，それらの条件が変わる度に最適な尤度を探索する必要がある。しかしながら，可能な限り実験条件を揃えるために，今後の実験では尤度  $prob$  を 0.5 に統一する。

#### 4.1.3.6 データ拡張を行った実験

##### 4.1.3.6.1 実験条件

これまでの実験では，高々10症例のVFを学習に用いていたため，パッチベースの実験を行ったとしても，十分な数の学習データを確保できなかった。本節では，11症例のVFに Leave one patient out 法を適用する実験を踏襲しつつ，データ拡張 (data augmentation) を行うことで



表 4.6: 深層学習の尤度を変更して得られた平均椎間板抽出精度

深層学習の尤度	precision	recall	accuracy	F 値
0.5	0.315	<b>0.828</b>	0.986	0.454
0.85	0.397	0.786	0.990	0.524
0.9	0.421	0.774	0.991	0.541
0.95	0.458	0.749	0.992	0.564
0.995	0.581	0.655	0.994	<b>0.610</b>
0.999	0.659	0.561	0.995	0.599
0.9999	<b>0.746</b>	0.408	<b>0.995</b>	0.519

学習に用いるパッチの枚数を水増しした。データ拡張とは、一枚の学習画像に幾何学的変換、色空間変更、カーネルフィルタを加えて複数の画像を生成し、それらを全て深層学習に入力する手法であり、少数のデータセットを用いた場合に精度が向上する [39]。本節では、データ拡張として、正例と負例の各パッチを水平、垂直方向それぞれに 0.6 倍、0.8 倍、1.0 倍、1.2 倍、1.4 倍の大きさに拡大、縮小した。拡大、縮小後の各パッチを  $-10$  度、 $-5$  度、 $0$  度、 $5$  度、 $10$  度回転させた。すなわち、一枚のパッチを 125 枚に水増しした。実験で使用した VF では、椎間板画素と比較して非椎間板画素の数が十分多かったため、本来であればデータ拡張を行わなくとも多数の負例を確保できたが、正例との対称性を重視し、負例のパッチに対してもデータ拡張を行った。計算機環境の制約から、10 症例の学習症例から生成したパッチ全てにデータ拡張を適用した後、正例、負例からそれぞれ 100 万枚のパッチをランダムに選択し、深層学習に入力した。ただし、 $S_p$  を 19、 $prob$  を 0.5 とし、各パッチに画素値正規化補正を施し回転補正を施さずに深層学習に入力した。4.1.3.1 節と同一の学習、テスト症例、深層学習モデル、計算機、フレームワークを用い、4.1.3.3 節と同一の椎間板領域のデータを用いた。

#### 4.1.3.6.2 実験結果

表 4.7 にデータ拡張の有無を変更して得られた平均椎間板抽出精度を示す。ただし、データ拡張を行わなかった実験結果として表 4.5 の結果を再掲した。データ拡張により、precision, accuracy, F 値が向上することがわかった。図 4.20 にデータ拡張を適用して得られた結果画像と同一症例の F 値を図 4.46 に示す。図 4.43 と比較し、椎体領域の偽陽性の減少を目視で確認できる。

表 4.7: データ拡張の有無を変更して得られた平均椎間板抽出精度

データ拡張	precision	recall	accuracy	F 値
あり	<b>0.409</b>	0.802	<b>0.991</b>	<b>0.538</b>
なし	0.344	<b>0.812</b>	0.988	0.481

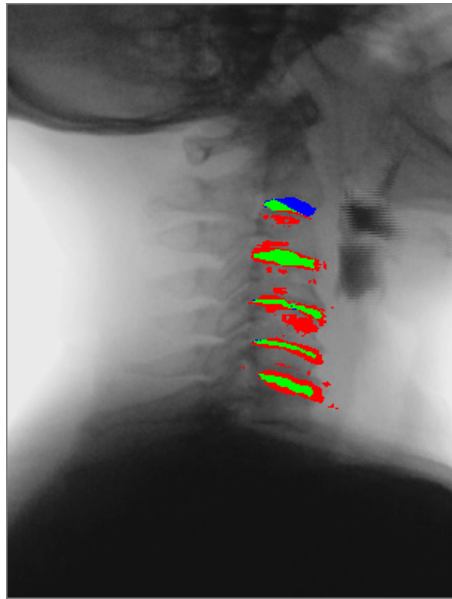


図 4.46: データ拡張を適用して得られた結果画像 ( $F$  値 = 0.503)

#### 4.1.3.7 VF データセットの増加に伴い椎間板領域のデータと頸部マスクを改良した実験

##### 4.1.3.7.1 実験条件

これまでの実験では、患者群 11 症例の VF を用いて実験を行った。本節では、前節までの 11 症例を含む患者群 19 症例と正常群 39 症例を合わせた 58 症例の VF データセットを用いて実験を行う。VF データセットの増加に伴い、前節までの 11 症例を含む全症例の椎間板領域のデータを以下の手順で作成した。各症例の VF には複数回の嚙下が含まれており、最大数秒程度、フレーム数にして最大数千フレーム程度の長さであった。各 VF から一回の嚙下を選択し、嚙下とその前後の連続する 71 フレームをクロップしたうえで、VF 画像端のスリットを削除した。続いて、71 フレームから舌骨最大挙上フレームを選定し、舌骨最大挙上フレーム上の目視可能な椎間板領域全体を著者が手動で抽出した。

VF データセットの増加に伴い、4.1.1 節に示した手順では頸部マスクを作成できない症例が生じたため、新たな形状の頸部マスクを考案することにした。新たな頸部マスクとして A, B, C の 3 種類を考案し、比較検討を行った。

頸部マスク A の作成手順を示す。4.1.1 節では判別分析法を用いて人体領域と背景領域を二分したが、頸部マスク A では、フレーム毎に画素値の最大値  $v1_{max}$  と最小値  $v1_{min}$  を求め、画素値  $\frac{v1_{max}+2v1_{min}}{3}$  をしきい値とする二値化を行った。図 4.147 を二値化して得られた画像を図 4.47 に、平滑化後の人体領域と背景領域を図 4.48 に示す。また、4.1.1 節では、眼周辺の孤立領域を削除する際に、面積が 2 番目以下の人体領域と面積が 3 番目以下の背景領域を削除したが、頸部マスク A では、全ての背景領域の 2 つの組合せに対して、重心同士を結ぶ直線の水平軸との角度を求め、角度が  $-60$  度から  $-30$  度の範囲内に収まる組合せのみを採

用した。図 4.49 に孤立領域削除後の人体領域と背景領域を示す。4.1.1 節では、頭部と肩部の放物線を抽出した後、二本の放物線に垂直方向に挟まれた頸部領域の画像に判別分析法を適用したが、頸部マスク A では、フレーム毎に画素値の最大値  $v_{2_{max}}$  と最小値  $v_{2_{min}}$  を求め、画素値  $\frac{v_{2_{max}}+v_{2_{min}}}{2}$  をしきい値とする二値化を行った。4.1.1 節では、二値化後の頸部領域に closing フィルタを 4 回、opening フィルタを 6 回、closing フィルタを 2 回順に適用したが、頸部マスク A では、closing フィルタを 4 回、opening フィルタを 4 回、dilation フィルタを 8 回順に適用し、複数の領域が存在する場合には各フレームで面積が最大の領域を選択し、頸部マスクとした。図 4.50 に頸部マスク A を示す。



図 4.47: 図 4.147 を二値化して得られた人体領域と背景領域



図 4.48: 平滑化後の人体領域と背景領域



図 4.49: 孤立領域削除後の人体領域と背景領域



図 4.50: 頸部マスク A

頸部マスク B の作成手順を示す。4.1.1 節では，判別分析法を用いて人体領域と背景領域を二分したが，頸部マスク B では，フレーム毎に画素値の最大値  $v1_{max}$  と最小値  $v1_{min}$  を求め，画素値  $\frac{2v1_{max}+v1_{min}}{3}$  をしきい値とする二値化を行った。図 4.147 を二値化して得られた人体領域と背景領域を図 4.51 に示す。二値化後の画像を水平軸に対して  $-5$  度傾けることで，新たな X, Y 座標を得た。図 4.52 に XY 座標系を示す。それぞれの Y 座標に対して線分  $Y = a$  を作成し，画像内の線分上の画素数に対する人体領域の画素数の割合を計算した。割合が 60% 以下の場合のみ，線分上の人体領域の画素を新規画像  $imgKeep$  に保存した。a を動かし， $imgKeep$  を完成させた。図 4.53 に  $imgKeep$  を示す。画像内の X, Y 軸方向の大きさをそれぞれ  $Xsize$ ,  $Ysize$  と定義し， $imgKeep$  内の人体領域の重心を中心とし，二辺の長さが  $0.6 \times Ysize$ ,  $0.4 \times Xsize$  となる長方形を生成した。長方形を右方向に  $0.25 \times Xsize$ ，上方向に  $0.125 \times Ysize$  移動し，頸部マスクとした。図 4.54 に頸部マスク B を示す。



図 4.51: 図 4.147 を二値化して得られた人体領域と背景領域

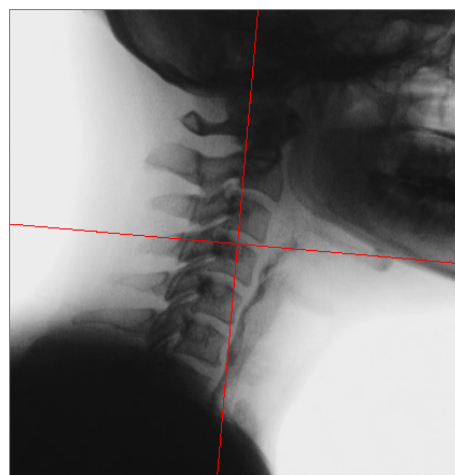


図 4.52: 画像を  $-5$  度回転して得られた XY 座標系

頸部マスク C の作成手順を示す。4.1.1 節に示した手順で頭部と肩部の放物線を抽出した後，二本の放物線に垂直方向に挟まれた首中心線分を抽出した。首中心線分から距離 100 ピクセル以内の画素を頸部マスクとした。図 4.55 に頸部マスク C を示す。

#### 4.1.3.7.2 実験結果

58 症例の舌骨最大挙上フレームの最小値は 32 フレーム目，最大値は 38 フレーム目であった。図 4.175 に患者症例の VF フレーム画像を，図 4.57 に同一患者症例の椎間板領域画像を，図 4.147 に正常症例の VF フレーム画像を，図 4.148 に同一正常症例の椎間板領域画像を示す。

4.1.1 節の手順で頸部マスクを作成できなかった症例画像を図 4.60 に示す。本症例では，判別分析法を用いて人体領域と背景領域を二分し，opening フィルタを用いて平滑化を行った際に，背景領域が一領域しか残らず，頸部マスクを作成できなかった。図 4.61 に判別分析法を適用後の人体領域と背景領域を，図 4.62 に平滑化後の人体領域と背景領域を示す。頸部マ



図 4.53: 人体領域の画素数でしきい値を定めて得られた *imgKeep*



図 4.54: 頸部マスク B



図 4.55: 頸部マスク C

ク A, B, C では, 58 症例全ての VF でマスクを作成することができた. ただし, 4.1.1 節の手順で作成した頸部マスクと頸部マスク A では, 造影剤が存在するフレームでは形状が大きく変形してしまい, 一部の椎間板が頸部マスク内に収まらなかった. 頸部マスク B では, 各症例の全てのフレームで面積が等しくなるような制約を加えたため, 造影剤に左右されない安定したマスク領域の抽出が可能となったが, 一部の椎間板が頸部マスク内に収まらないという問題は解決せず, また, 実験的に決定したパラメータが多く数学的に説明し難いという問題が新たに生じた. 一方, 頸部マスク C では, 造影剤に大きく左右されない角丸長方形のマスク領域の抽出が可能となり, 58 症例の合計 295 本の目視可能な椎間板が全てマスク内に収まり, 恣意的なパラメータの個数が減少した. 今後の実験では頸部マスク C を採用する.



図 4.56: 患者症例の VF フレーム画像

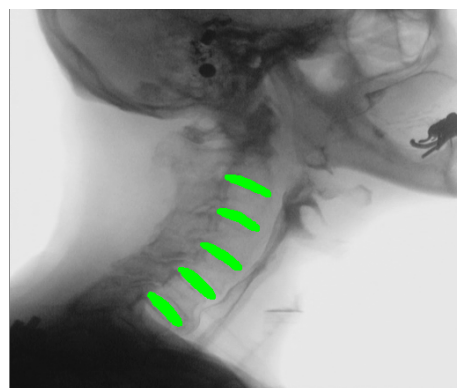


図 4.57: 患者症例の椎間板領域画像



図 4.58: 正常症例の VF フレーム画像

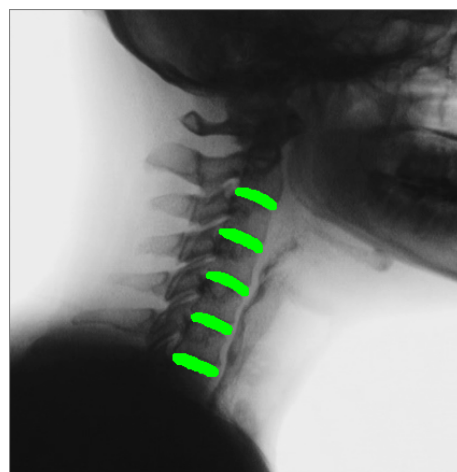


図 4.59: 正常症例の椎間板領域画像



図 4.60: 頸部マスクの作成が不可能な症例の VF フレーム画像

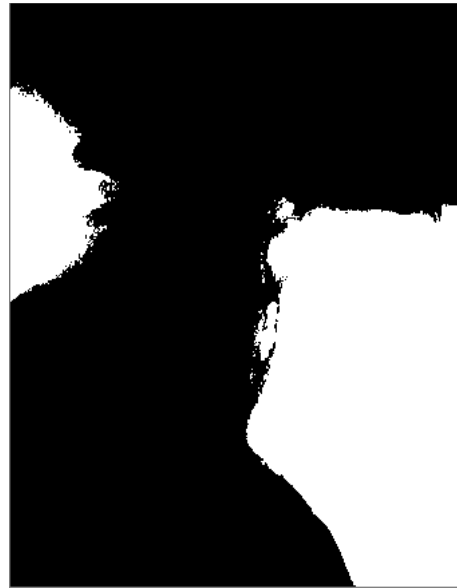


図 4.61: 判別分析法を適用後の人体領域と背景領域



図 4.62: 平滑化後の人体領域と背景領域

#### 4.1.3.8 58 症例の VF を用いてパッチサイズを比較した実験

##### 4.1.3.8.1 実験条件

58 症例の VF データセットに対し、4.1.3.7 節で作成した椎間板領域画像，頸部マスク C を適用し，深層学習を用いて椎間板領域を抽出した．Leave one patient out 法により，57 症例を学習に，残りの 1 症例をテストに割り当てた．パッチサイズ  $S_p$  を 15 から 25 まで 2 刻みで変更し，各パッチに画素値正規化補正を適用したうえで深層学習を入力し，椎間板抽出精度を比較した．ただし，4.1.3.1 節では，各パッチの最大画素値が 255 に，最小画素値が 0 になるように線形の画素値正規化補正を行ったのに対し，本節では，各パッチ内の画素値の平均値  $v_{mean}$  と分散  $v_{sigma}$  を求め，画素値  $v_{mean} + 3v_{sigma}$  が 255 に，画素値  $v_{mean} - 3v_{sigma}$  が 0 になるように線形の画素値正規化補正を行った． $prob$  を 0.5 とし，パッチの回転補正とデータ拡張は行わず，4.1.3.1 節と同一の深層学習モデル，計算機，フレームワークを用いて実験を行った．

##### 4.1.3.8.2 実験結果

$S_p$  を 15 から 25 まで 2 刻みで変更して得られた 58 症例の平均椎間板抽出精度を表 4.8 に示す． $S_p$  を 19 としたとき precision, accuracy, F 値が最大となり， $S_p$  を 23 としたとき recall が最大となった．図 4.20 と同一の患者症例の Ground truth を図 4.63 に， $S_p$  を 6 通りに変更して得られた結果画像と同一症例の F 値を図 4.64，図 4.65，図 4.66，図 4.67，図 4.68，図 4.69 に示す． $S_p$  を 15, 17, 19, 21, 23 としたとき，58 症例のパッチ作成，深層学習の学習，71 フレームのテストに費やした合計計算時間はそれぞれ約 22 時間，約 27 時間，約 28 時間，約 27 時間，約 26 時間となった．

表 4.8: 4.1.3.8 節で  $S_p$  を 15 から 25 まで 2 刻みで変更して得られた平均椎間板抽出精度

$S_p$	precision	recall	accuracy	F 値
15	0.336	0.923	0.973	0.488
17	0.364	0.926	0.976	0.519
19	<b>0.378</b>	0.928	<b>0.978</b>	<b>0.534</b>
21	0.374	0.934	0.977	0.531
23	0.371	<b>0.939</b>	0.977	0.528
25	0.356	0.939	0.975	0.513

#### 4.1.3.9 モルフォロジー演算と二値化，小領域削除を用いた後処理の実験

##### 4.1.3.9.1 実験条件

4.1.3.8 節で深層学習を適用して得られた椎間板候補領域に対して，半径 1 ピクセルの円形の closing フィルタを 3 回適用した．VF フレーム画像に半径 3 ピクセルの円形の median フィ



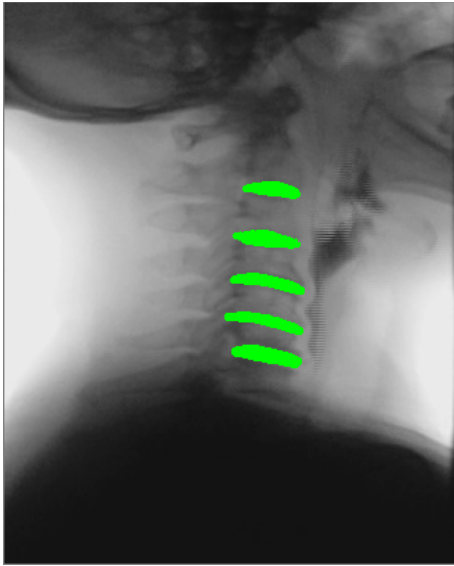


図 4.63: 図 4.20 と同一の患者症例の Ground truth

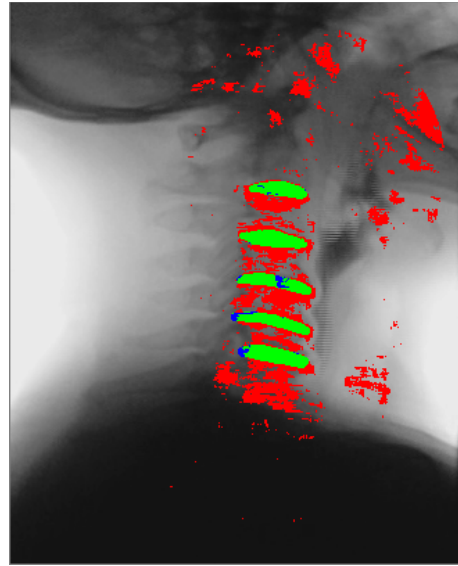


図 4.64: 4.1.3.8 節で  $S_p = 15$  とした結果画像 ( $F$  値 = 0.387)

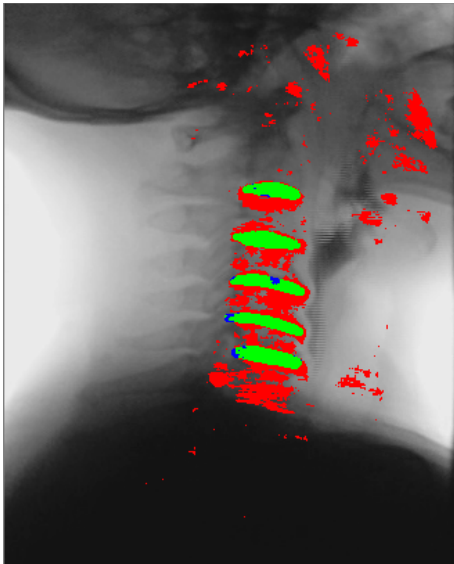


図 4.65: 4.1.3.8 節で  $S_p = 17$  とした結果画像 ( $F$  値 = 0.420)



図 4.66: 4.1.3.8 節で  $S_p = 19$  とした結果画像 ( $F$  値 = 0.440)

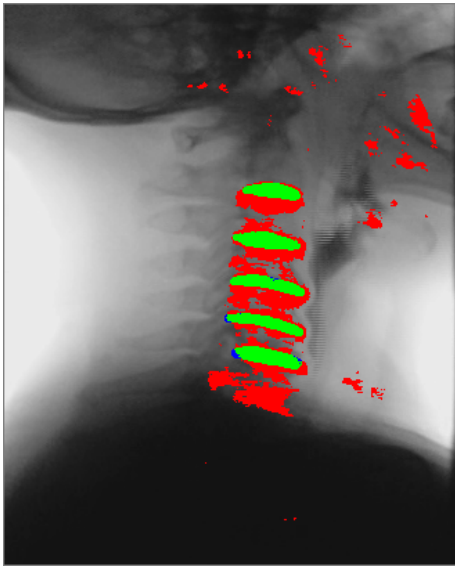


図 4.67: 4.1.3.8 節で  $S_p = 21$  とした結果画像 ( $F$  値 = 0.454)



図 4.68: 4.1.3.8 節で  $S_p = 23$  とした結果画像 ( $F$  値 = **0.464**)

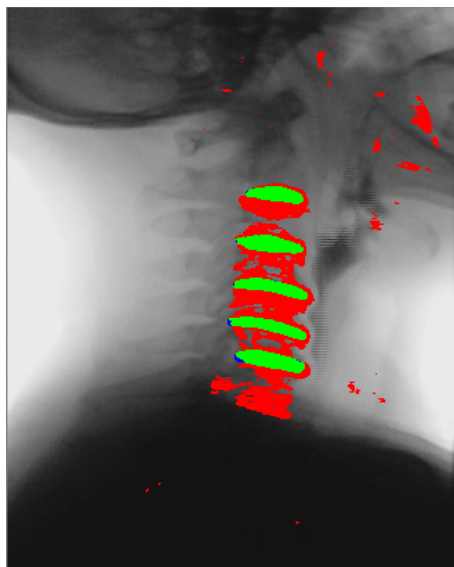


図 4.69: 4.1.3.8 節で  $S_p = 25$  とした結果画像 ( $F$  値 = 0.461)

ルタを適用後，半径 3 ピクセルの円形のもルフォロジー演算の Top-hat 変換フィルタを適用し，画素値のしきい値を 1 とする二値化を行った．二値化後の画像と closing フィルタを適用後の椎間板候補領域の両方の画素値が 255 の画素のみを抽出した．抽出後の画像から，体積が 1000 ボクセル以下の領域を削除した．小領域削除後の画像に半径 2 ピクセルの erosion フィルタを適用後，それぞれのフレームで面積が 30 ピクセル以下の領域を削除した．小領域削除後の画像に半径 3 ピクセルの円形の dilation フィルタを適用後，半径 3 ピクセルの円形の median フィルタを適用し，最終結果画像とした．最終結果画像の評価を行う際には，抽出精度が Ground truth の作り方に大きく依存しないよう，半径 2 ピクセルの誤差を許容して画素単位の F 値を求めた．

#### 4.1.3.9.2 実験結果

図 4.66 に closing フィルタを適用して得られた画像を図 4.70 に，VF フレーム画像に Top-hat 変換フィルタと二値化を適用して得られた画像を図 4.71 に，図 4.70 と図 4.71 の論理積を計算して得られた画像を図 4.72 に，体積が 1000 ボクセル以下の小領域を削除して得られた画像を図 4.73 に，erosion フィルタを適用して得られた画像を図 4.74 に，面積が 30 ピクセル以下の小領域を削除して得られた画像を図 4.75 に，dilation フィルタを適用して得られた画像を図 4.76 に，最終結果画像を図 4.77 に示す． $S_p$  を 15 から 25 まで 2 刻みに変更して得られた各後処理段階の平均 F 値を表 4.9 に示す．

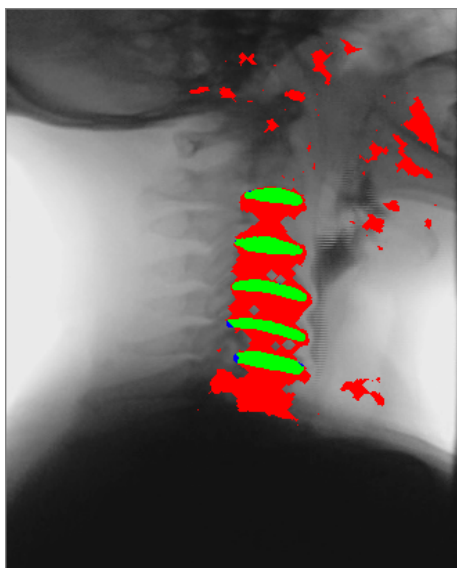


図 4.70: 4.1.3.9 節で図 4.66 に closing フィルタを適用して得られた画像 ( $F$  値 = 0.557)



図 4.71: 4.1.3.9 節で VF フレーム画像に Top-hat 変換フィルタと二値化を適用して得られた画像

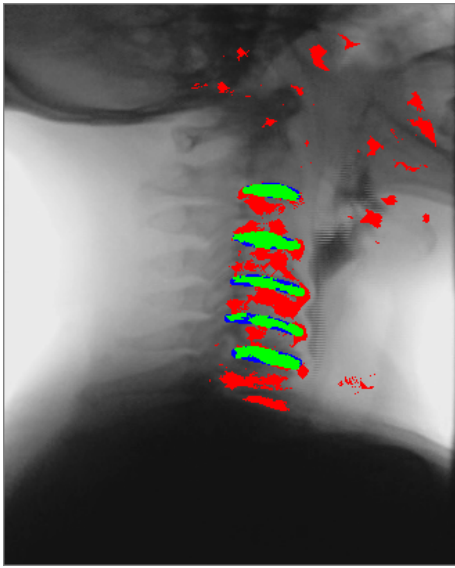


図 4.72: 4.1.3.9 節で図 4.70 と図 4.71 の論理積を計算して得られた画像 ( $F$  値 = 0.695)

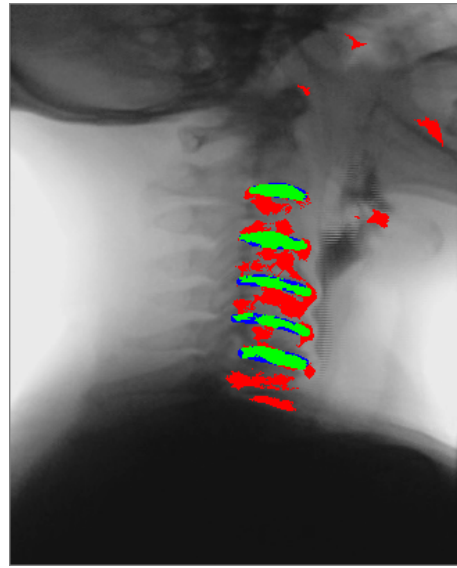


図 4.73: 4.1.3.9 節で体積が 1000 ボクセル以下の小領域を削除して得られた画像 ( $F$  値 = 0.714)

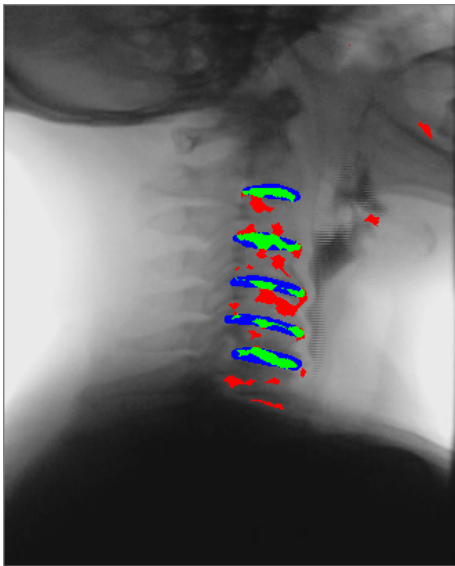


図 4.74: 4.1.3.9 節で erosion フィルタを適用して得られた画像 ( $F$  値 = 0.655)

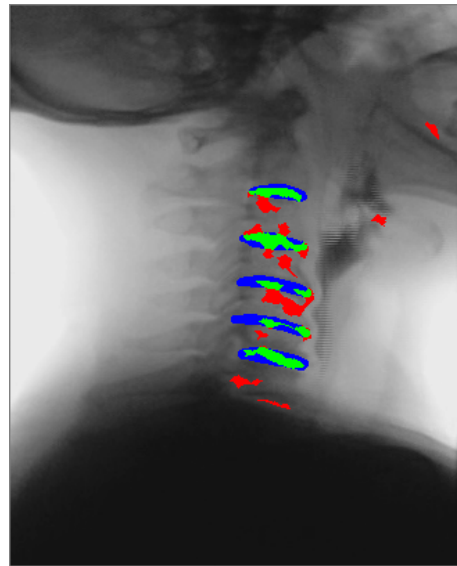


図 4.75: 4.1.3.9 節で面積が 30 ピクセル以下の小領域を削除して得られた画像 ( $F$  値 = 0.671)

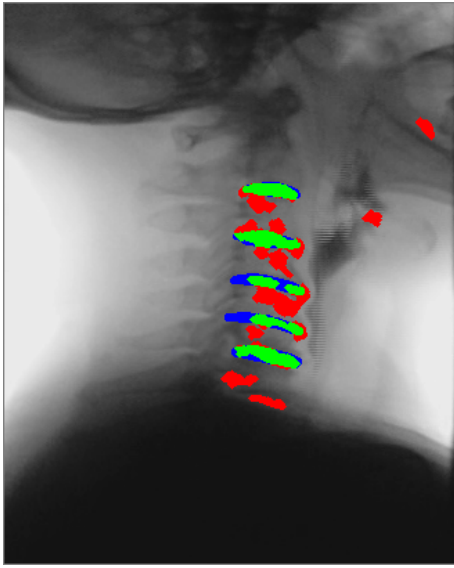


図 4.76: 4.1.3.9 節で dilation フィルタを適用して得られた画像 ( $F$  値 = 0.795)

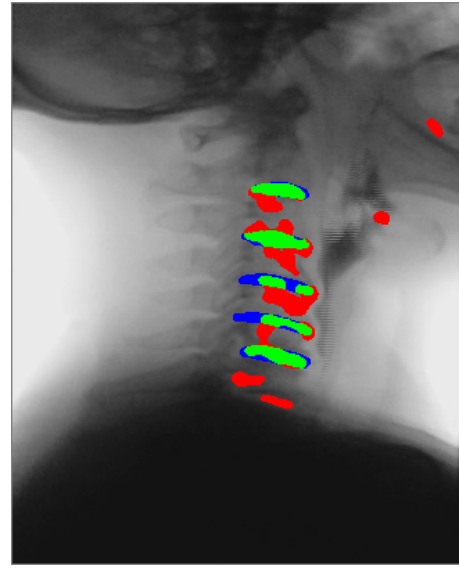


図 4.77: 4.1.3.9 節の最終結果画像 (半径 2 ピクセルの誤差を許容した  $F$  値 = 0.909)

#### 4.1.3.10 58 症例の VF にホールドアウト法を適用してパッチサイズと画素値正規化補正の有無を比較した実験

##### 4.1.3.10.1 実験条件

4.1.3.8 節では、Leave one patient out 法により、57 症例を学習に、残りの 1 症例をテストに割り当てて深層学習に入力する実験を 58 回行っていったため、計算時間が長くなってしまった。この問題を解決するために、本節では、ホールドアウト法により 6 割の VF を学習症例に、残りの 4 割の VF をテスト症例に割り当て、深層学習を一度だけ適用する実験を行った。学習、テスト症例間の患者群と正常群の割合が均等になるようにランダムに分割し、患者群 11 症例と正常群 24 症例を合わせた 35 症例を学習に、残りの患者群 8 症例と正常群 15 症例を合わせた 23 症例をテストに割り当てた。パッチサイズ  $S_p$  を 15 から 25 まで 2 刻みで変更し、それぞれの  $S_p$  に対して 4.1.3.8 節と同様の画素値正規化補正の有無を変更し、椎間板抽出精度を比較した。ただし、 $prob$  を 0.5 とし、パッチの回転補正とデータ拡張は行わず、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、4.1.3.1 節と同一の深層学習モデル、計算機、フレームワークを用いて実験を行った。

##### 4.1.3.10.2 実験結果

$S_p$  を 15 から 25 まで 2 刻みで変更し、それぞれの  $S_p$  に対して画素値正規化補正の有無を変更して得られた 23 症例の平均椎間板抽出精度を表 4.10 に示す。  $S_p$  を 19 とし、画素値正規化補正を行った実験で precision, accuracy, F 値が最大となり、  $S_p$  を 25 とし、画素値正規化

表 4.9: 4.1.3.9 節で  $S_p$  を 15 から 25 まで 2 刻みに変更して得られた各後処理段階の平均 F 値

$S_p$	15	17	19	21	23	25
深層学習を適用後の椎間板候補領域	0.488	0.519	0.534	0.531	0.528	0.513
closing フィルタ適用後	0.427	0.472	0.494	0.499	0.501	0.488
Top-hat 変換フィルタとの論理積	0.521	0.560	0.582	0.586	0.591	0.579
体積 1000 ボクセル未満の小領域削除後	0.569	0.598	0.610	0.607	0.606	0.591
erosion フィルタ適用後	0.572	0.584	0.590	0.587	0.586	0.576
面積 30 ピクセル未満の小領域削除後	0.581	0.590	0.593	0.592	0.590	0.580
dilation フィルタ適用後	0.644	0.661	0.664	0.660	0.657	0.641
最終結果 (誤差を許容しない)	0.657	0.673	0.676	0.672	0.669	0.652
最終結果 (半径 2 ピクセルの誤差を許容)	<b>0.791</b>	<b>0.810</b>	<b>0.813</b>	<b>0.811</b>	<b>0.809</b>	<b>0.792</b>

補正を行った実験で recall が最大となった。  $S_p$  を 15, 17, 19, 21, 23, 25 とし て画素値正規化補正を行った実験で、パッチ作成、深層学習の学習、71 フレームのテストに費やした合計計算時間がそれぞれ約 4 時間、約 7 時間、約 7 時間、約 7 時間、約 6 時間、約 7 時間となったのに対し、  $S_p$  を 15, 17, 19, 21, 23, 25 とし て画素値正規化補正を行わなかった実験では、合計計算時間がそれぞれ約 4 時間、約 6 時間、約 6 時間、約 6 時間、約 6 時間、約 5 時間となった。

表 4.10: 4.1.3.10 節で  $S_p$  と画素値正規化補正の有無を変更して得られた平均椎間板抽出精度

$S_p$	画素値正規化補正	precision	recall	accuracy	F 値
15	あり	0.344	0.923	0.974	0.498
17	あり	0.374	0.931	0.977	0.532
19	あり	<b>0.393</b>	0.935	<b>0.979</b>	<b>0.552</b>
21	あり	0.389	0.940	0.978	0.548
23	あり	0.377	0.944	0.977	0.537
25	あり	0.367	<b>0.953</b>	0.976	0.528
15	なし	0.301	0.854	0.971	0.443
17	なし	0.328	0.859	0.974	0.473
19	なし	0.337	0.894	0.974	0.488
21	なし	0.341	0.912	0.974	0.495
23	なし	0.329	0.911	0.973	0.482
25	なし	0.322	0.938	0.972	0.479

#### 4.1.3.11 モルフォロジー演算と二値化，小領域削除，深層学習の尤度を用いた後処理の実験

##### 4.1.3.11.1 実験条件

4.1.3.10 節で深層学習を適用して得られた椎間板候補領域に対して，以下の手順で後処理を行った．深層学習のテスト段階で，各パッチが正例に属する尤度が得られたため，尤度に10000を掛けた値をパッチの中心画素の画素値とする尤度画像を作成した．尤度画像の画素値が1以上の画素のみを対象として判別分析法を適用し，二値化を行った．本論文で使用したVFデータセットには各症例最大6本の目視可能な椎間板が含まれていたため，二値化後の画像の各フレームにラベリングを適用し，面積が7番目以下の領域を削除した．ただし，各フレームで面積が100ピクセル以下の領域は，面積の順位に関係なく削除した．VFフレーム画像に半径2ピクセルの円形のmedianフィルタを適用後，半径5ピクセルの円形のTop-hat変換フィルタを適用した．小領域削除後の椎間板候補領域をマスクに，Top-hat変換フィルタを適用後のフレーム画像に判別分析法を適用し，二値化を行った．生成された画像に半径2ピクセルの円形のclosingフィルタを2回適用した．続いて半径4ピクセルの円形のmedianフィルタを適用し，最終結果画像とした．最終結果画像の評価を行う際には，半径2ピクセルの誤差を許容して画素単位のF値を求めた．

##### 4.1.3.11.2 実験結果

患者テスト症例のVFフレーム画像を図4.78に，Ground truthを図4.79に， $S_p = 19$ として画素値正規化補正後に深層学習を適用して得られた椎間板候補領域を図4.80に，深層学習の尤度画像のカラーマップを図4.81に，図4.81を二値化して得られた画像を図4.82に，小領域を削除して得られた画像を図4.83に，図4.78にTop-hat変換フィルタを適用して得られた画像を図4.84に，図4.84を二値化して得られた画像を図4.85に，closingフィルタを適用して得られた画像を図4.86に，最終結果画像を図4.87に示す． $S_p$ を15から25まで2刻みで変更して画素値正規化補正を適用して得られた各後処理段階の平均F値を表4.11に示す．

表 4.11: 4.1.3.11 節で  $S_p$  を 15 から 25 まで 2 刻みに変更して画素値正規化補正を適用して得られた各後処理段階の平均 F 値

$S_p$	15	17	19	21	23	25
深層学習を適用後の椎間板候補領域	0.498	0.532	0.552	0.548	0.537	0.528
尤度画像を二値化後	0.491	0.528	0.550	0.547	0.537	0.528
小領域削除後	0.579	0.587	0.591	0.580	0.558	0.545
Top-hat 変換画像を二値化後	0.598	0.608	0.605	0.606	0.593	0.590
closing フィルタを適用後	0.616	0.626	0.629	0.625	0.605	0.596
最終結果 (誤差を許容しない)	0.635	0.644	0.647	0.644	0.624	0.616
最終結果 (半径 2 ピクセルの誤差を許容)	<b>0.794</b>	<b>0.805</b>	<b>0.807</b>	<b>0.807</b>	<b>0.786</b>	<b>0.778</b>





図 4.78: 患者テスト症例の VF フレーム画像

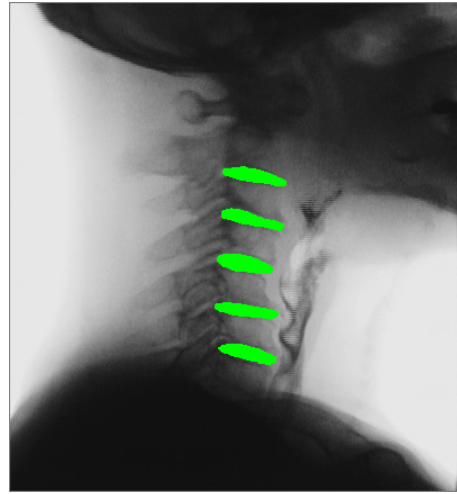


図 4.79: 患者テスト症例の Ground truth

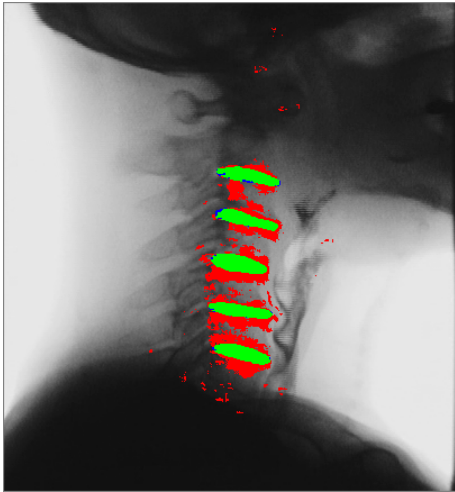


図 4.80: 4.1.3.10 節で  $S_p = 19$  として画素値正規化補正後に深層学習を適用して得られた椎間板候補領域 ( $F$  値 = 0.568)

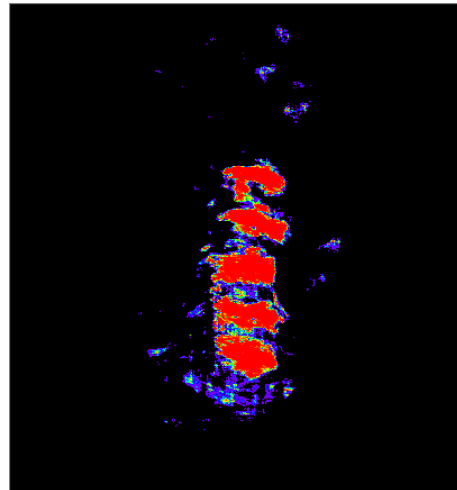


図 4.81: 4.1.3.11 節で作成した深層学習の尤度画像のカラーマップ



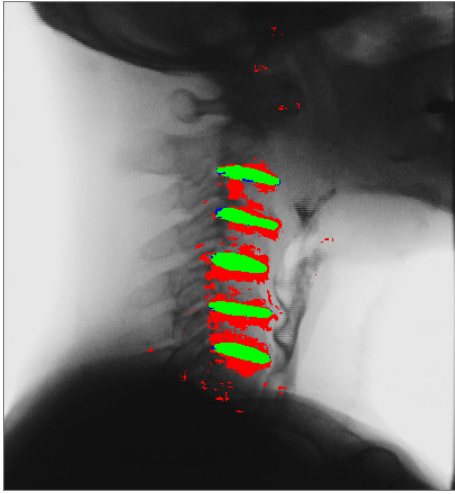


図 4.82: 4.1.3.11 節で図 4.81 を二値化して得られた画像 ( $F$  値 = 0.567)

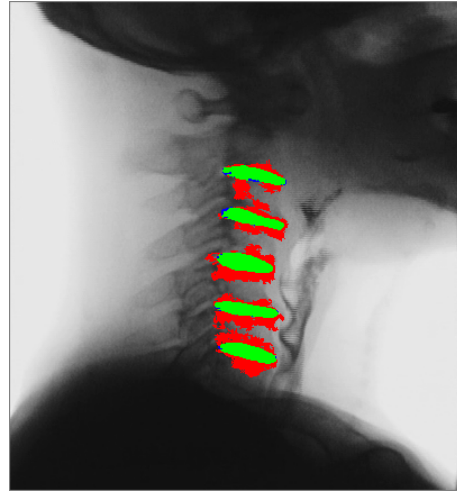


図 4.83: 4.1.3.11 節で小領域を削除して得られた画像 ( $F$  値 = 0.596)

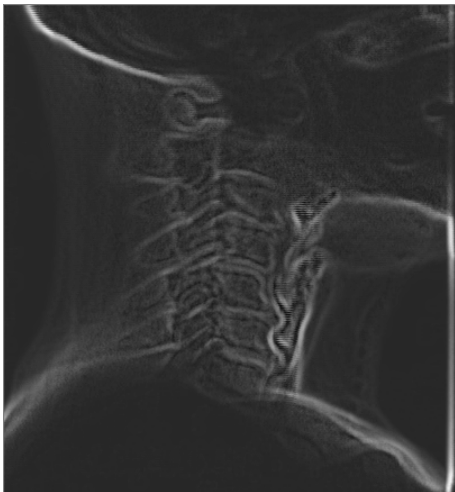


図 4.84: 4.1.3.11 節で VF フレーム画像に Top-hat 変換フィルタを適用して得られた画像

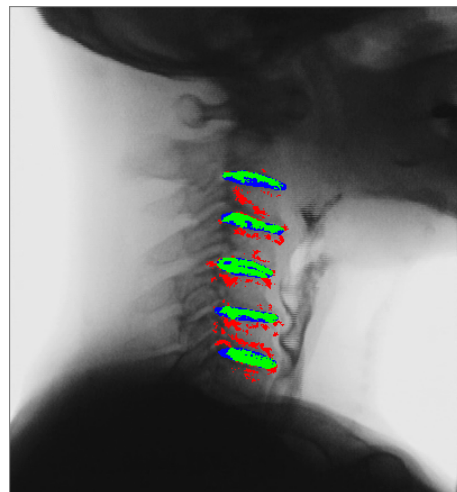


図 4.85: 4.1.3.11 節で図 4.84 を二値化して得られた画像 ( $F$  値 = 0.642)

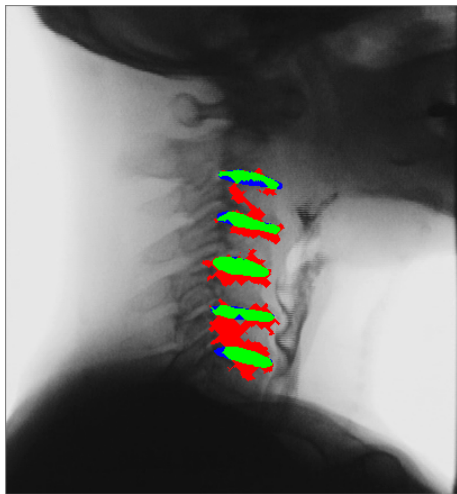


図 4.86: 4.1.3.11 節で closing フィルタを適用して得られた画像 ( $F$  値 = 0.610)

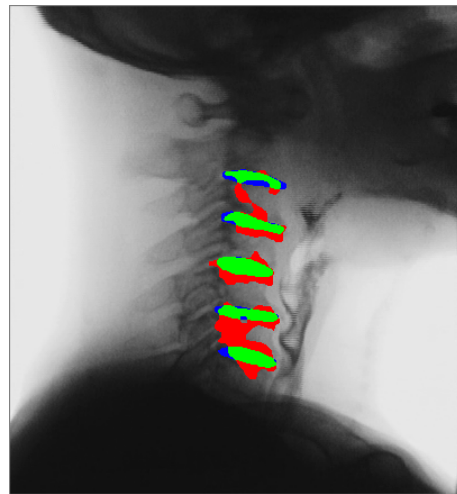


図 4.87: 4.1.3.11 節の最終結果画像 (半径 2 ピクセルの誤差を許容した  $F$  値 = **0.765**)

#### 4.1.3.12 モルフォロジー演算と二値化，深層学習の尤度を用いた後処理の実験

##### 4.1.3.12.1 実験条件

4.1.3.10 節で深層学習を適用して得られた椎間板候補領域に対して，4.1.3.11 節と別の手順で後処理を行い，結果を比較した。まず，4.1.3.11 節に示した手順で深層学習の尤度画像を作成した。尤度画像に対して半径 3 ピクセルの 3 次元 erosion フィルタを適用した。判別分析法で二値化を行い，半径 3 ピクセルの 3 次元メディアンフィルタを適用することで，最終結果画像とした。

##### 4.1.3.12.2 実験結果

尤度画像に erosion フィルタを適用後のカラーマップを図 4.88 に，二値化後して得られた画像を図 4.89 に，最終結果画像を図 4.90 に示す。 $S_p$  を 15 から 25 まで 2 刻みで変更して画素値正規化補正を適用して得られた各後処理段階の平均  $F$  値を表 4.12 に示す。

表 4.12: 4.1.3.12 節で  $S_p$  を 15 から 25 まで 2 刻みに変更して画素値正規化補正を適用して得られた各後処理段階の平均  $F$  値

$S_p$	15	17	19	21	23	25
深層学習を適用後の椎間板候補領域	0.498	0.532	0.552	0.548	0.537	0.528
二値化後	0.487	0.545	0.587	0.670	0.661	0.669
最終結果 (半径 2 ピクセルの誤差を許容)	<b>0.725</b>	<b>0.769</b>	<b>0.799</b>	<b>0.880</b>	<b>0.856</b>	<b>0.861</b>

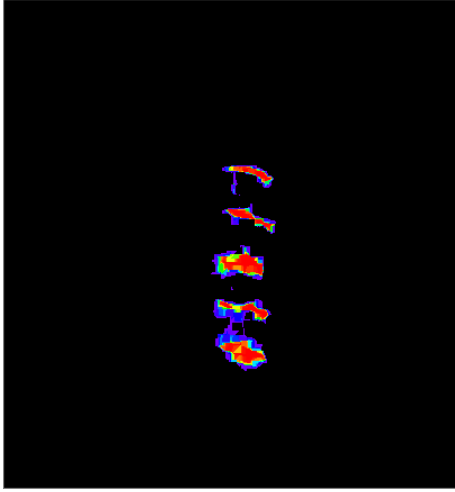


図 4.88: 4.1.3.12 節で尤度画像に erosion フィルタを適用後のカラーマップ

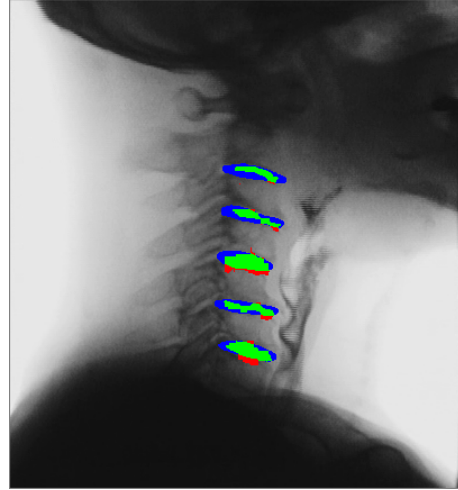


図 4.89: 4.1.3.12 節で二値化して得られた画像 ( $F$  値 = 0.627)

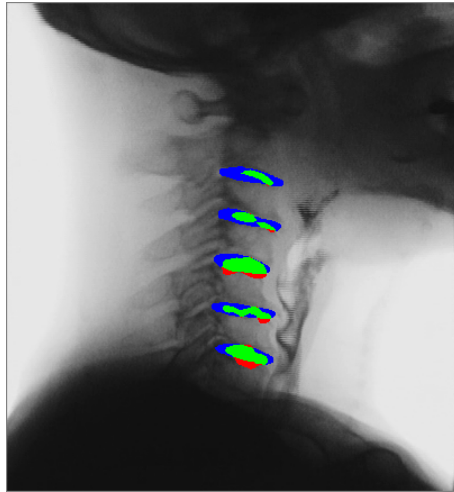


図 4.90: 4.1.3.12 節の最終結果画像 (半径 2 ピクセルの誤差を許容した  $F$  値 = **0.850**)

## 4.2 マルチチャンネル化による前処理を用いた方法

4.1 節の手法では、深層学習の出力画像にモルフォロジー演算、判別分析法、メディアンフィルタを適用することで高い精度で椎間板を抽出することができたが、これらの画像フィルタの種類、適用順、パラメータは経験的に決定していた。経験的な処理手順の決定は抽出対象やモダリティ、データセットが変わる度に設計者が試行錯誤を繰り返す必要があり、効率的ではない。そこで本節では、アドホックな後処理を廃止し、マルチチャンネル化 (multi-channelization, MC) という新しい前処理技術を導入することによって、抽出精度を向上させることを目指す。なお、後処理として比較的安定して椎間板を包含する領域を抽出することができる頸部マスクの作成はそのまま使用する。

### 4.2.1 深層学習とマルチチャンネル化、特徴画像の選択による椎間板領域の抽出

VF の各フレーム画像は 8 ビットのグレースケール画像である。各フレーム画像に  $M$  種類の画像フィルタをそれぞれ適用し、同数の特徴画像を生成する。 $M$  種類の特徴画像から重複を許して  $N$  種類を選択し、マルチチャンネル化画像 (MC 画像) と呼ばれる  $N$  個のチャンネルから構成される画像の  $c^1, c^2, \dots, c^n, \dots, c^N$  チャンネルにそれぞれ保存する。それぞれの特徴画像を  $(f^1_{i_1}, f^2_{i_2}, \dots, f^n_{i_n}, \dots, f^N_{i_N})$  と表す。ただし、 $i_1, i_2, \dots, i_n, \dots, i_N$  は特徴画像のインデックスを表し、それぞれ  $i_1, i_2, \dots, i_n, \dots, i_N \in \{1, 2, \dots, m, \dots, M\}$  を満たすものとする。MC 画像に対して深層学習を適用し、椎間板領域を抽出する。特徴画像の組合せ  $(f^1_{i_1}, \dots, f^N_{i_N})$  における抽出精度をエネルギー値  $e(f^1_{i_1}, \dots, f^N_{i_N})$  で表す。ただし、 $e(f^1_{i_1}, \dots, f^N_{i_N})$  は 1 から画素単位の F 値を引いた値で、 $0 \leq e(f^1_{i_1}, \dots, f^N_{i_N}) \leq 1$  を満たす。エネルギー値が小さいほど望ましい組合せとなる。 $N$  種類の特徴画像の組合せを変更する度に新しい深層学習モデルを作成する。 $M$  種類の特徴画像から  $N$  種類を選択する組合せの数は  $M$  が大きくなるにつれ膨大になるが、実行可能な時間内に最適な組合せを見つけるために焼きなまし法を用いる。焼きなまし法は山上り法、遺伝的アルゴリズムと同じく最適化手法の一種であり、組合せ最適化問題の解法となる。エネルギー値の最小化を目的として、組合せが変わる毎に温度  $T$  を減温し、組合せ変更後のエネルギー値が変更前より小さい場合には必ず組合せを更新し、大きい場合にも一定確率で更新する。これにより、局所最適解に陥りにくく大域的最適解を見つけやすくなる [45]。理論上、焼きなまし法で  $T$  を十分緩やかに冷却すれば、確率 1 で大域的最適解に辿り着く。本節では、焼きなましの初期化段階で、 $c^1, \dots, c^N$  チャンネル全てに同一の組合せ画像  $(f^1_i, \dots, f^N_i)$  を保存して  $e(f^1_i, \dots, f^N_i)$  を求めたうえで、特徴画像をエネルギー関数の昇順に順位付けする。 $(i_1, \dots, i_N)$  に  $(1, \dots, 1)$  を保存し、焼きなましの初期状態とする。焼きなましの段階では、 $i_1, \dots, i_N$  のそれぞれについて、乱数を用いてインデックスを  $[-step, +step]$  の範囲内で一時的に変更する。変更したインデックスが 0 以下あるいは  $M+1$  以上となった場合にはもう一度インデックスを変更する。変更後のインデックスを  $(i'_1, \dots, i'_N)$  と表す。 $e(f^1_{i_1}, \dots, f^N_{i_N}) > e(f^1_{i'_1}, \dots, f^N_{i'_N})$  を満たした場合、もしくは 0 以上 1 未満の一樣乱数が確率変数  $p = \exp(\frac{e-e'}{T})$  より小さい値になった場合にインデックスの更新を採択する。焼きなましで得られた特徴画像の組合せとエネルギー値を全て保存し、既に探索を行った組

合せが再度出現した場合には、チャンネルの順番にかかわらず、同じエネルギー値を再利用する。焼きなましの終了後、エネルギー値が最小となる特徴画像の組合せを最適解とみなす。手法の流れを図 4.91 に示す。

#### 4.2.2 画像前処理を用いた従来研究

マルチチャンネル化の関連研究として、画像前処理を用いた従来研究について述べる。原画像に前処理を適用し、1枚の特徴画像を生成し、それを識別器に入力する手法が提案されている。Jadoon ら [55] は、マンモグラムから乳房の腫瘍を分類するために、コントラスト調整ヒストグラム平坦化、2次元ウェーブレット変換、カーブレット変換、dense scale invariant feature transform (DSIFT) を組み合わせた前処理手法を提案した。前処理後の画像をパッチ化して CNN へ入力し、一部の実験でソフトマックス層を SVM 層に置換した。Jiang ら [56] は、マンモグラムから乳房の腫瘍を分類するために、原画像にヒストグラム平坦化フィルタを適用したうえで、SVM と Rough set theory を組み合わせた Improved SVM に入力する手法を提案した。Nanni ら [57] は、新生児の顔画像データセットからの疼痛予測、蛍光顕微鏡で撮影した細胞画像データセットからのフェノタイプ分類、パパニコロウ染色を行った子宮頸部細胞データセットからの異常分類を行うために、原画像に Local binary pattern (LBP) フィルタを適用して1枚の特徴画像を生成したうえで SVM に入力する手法を提案し、LBP フィルタのパラメータの比較を行った。Ramteke ら [58] は、頭部 CT 画像から脳の正常、異常を識別するために、原画像にメディアンフィルタを適用して1枚の特徴画像を生成したうえで識別器に入力し、後処理として頭蓋骨領域と灰白質領域を除去する手法を提案した。識別器として、K 最近傍法と SVM を別個に用いた。Chi-Hoon Lee [59] らは、頭部 MR 画像から脳腫瘍を抽出するために、原画像に平滑化フィルタとコントラスト強調フィルタを適用したうえで、再尤推定、ロジスティック回帰、SVM、マルコフ確率場、条件付き確率場にそれぞれ入力する手法を提案した。

原画像に前処理を適用し、複数枚の特徴画像を生成し、それらを並列に識別器に入力する手法が提案されている。Jeon ら [60] は、小腸カプセル内視鏡における小腸の病変を検出するために、RGB 画像を CIELab 画像に変換し、両者を並列にパッチベースの CNN へ入力する手法を提案した。CNN として、ImageNet [61] を用いて転移学習を行った GoogLeNet を使用した。RGB 画像と CIELab 画像それぞれにおけるソフトマックス層の出力結果を統合するために、SVM 層を用いた。Shen ら [62] は、胸部 CT 画像から肺の腫瘍を分類するために、大きさの異なる3種類のパッチを並列に CNN に入力する multi-scale convolutional neural networks (MCNN) を提案した。Histogram of oriented gradients あるいは local binary pattern (LBP) を用いて前処理を行った際の分類精度を上回った。Huang ら [63] は、CNN の最初の層で解像度の異なる複数の画像を生成し、層の深さ方向と特徴マップの次元方向の両方で相互結合を行う multi-scale dense networks (MSDNets) を提案した。平野ら [64] は、皮膚写真からメラノーマを識別するために、原画像の RGB チャンネルの内1チャンネルを二値化画像に差し替えた画像を深層学習に入力する手法と、輪郭からのユークリッド距離画像を4チャンネル目に追加して深層学習に入力する手法を提案した。深層学習として、ImageNet を用いて転移学習を行っ

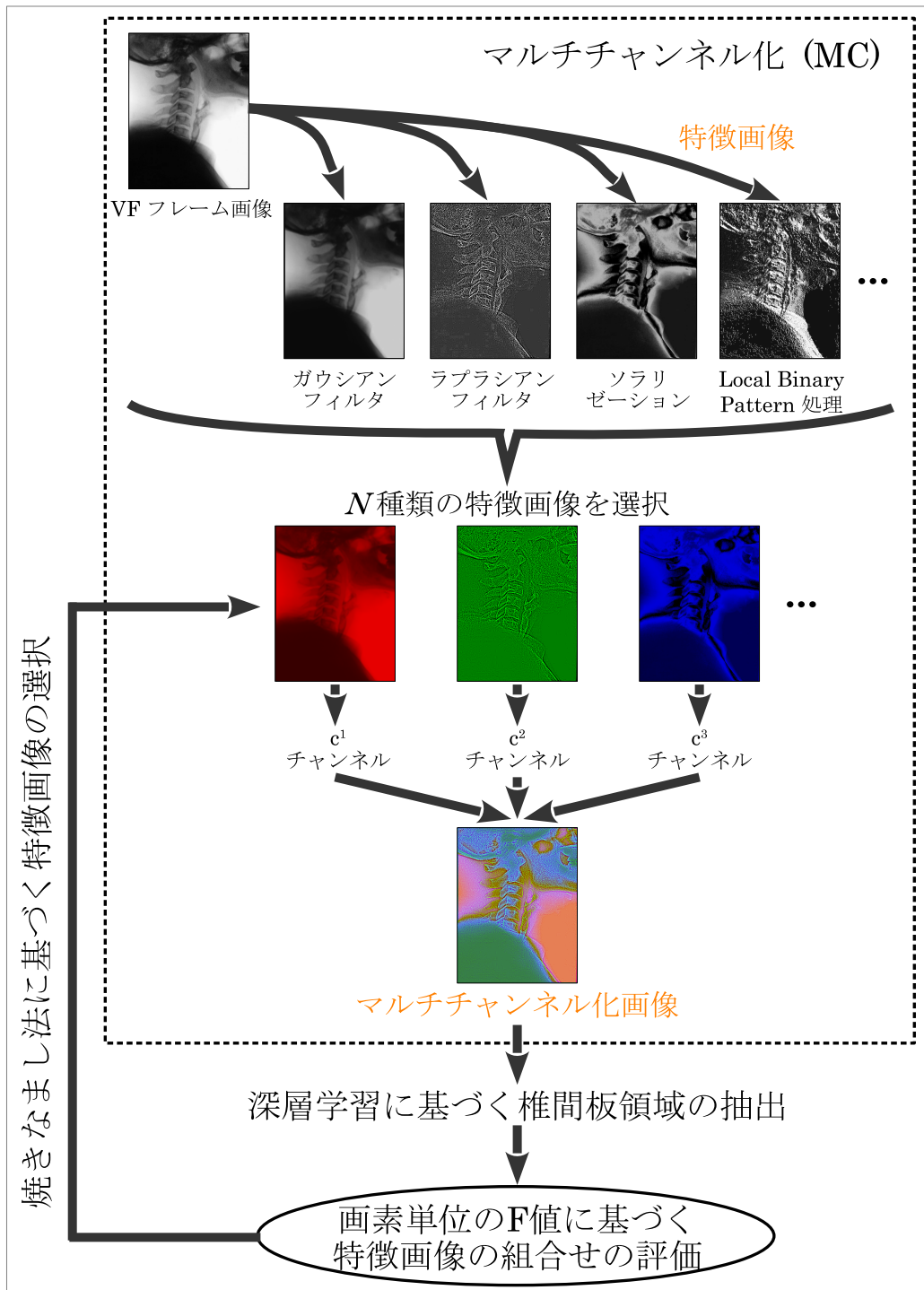


図 4.91: 深層学習と MC, 焼きなまし法に基づく特徴画像の選択による椎間板の領域抽出

た VGG-16 を使用した。Heidari ら [65] は、胸部 X 線画像から COVID-19 の識別を行うために、原画像にヒストグラム平坦化フィルタとバイラテラルフィルタをそれぞれ適用して得られた特徴画像を原画像と併せて、深層学習の 3 チャンネルの入力として用いる手法を提案した。数尾ら [66] は、低線量 CT 画像やスパースビュー CT 画像の画質を改善するために、トータルバリエーション正則化項が異なる複数の圧縮センシング再構成画像を生成し、それらを並列に U-Net に入力する手法を提案した。松山ら [67] は、胸部 CT 画像から肺の腫瘍を分類するために、ウェーブレット変換を前処理に用いた特徴画像の選択手法を提案した。CT 原画像に低周波、高周波領域の水平、垂直方向の 2 次元ウェーブレット変換を施すことで 4 種類の特徴画像を生成し、最適な 3 種類を選択したうえで CNN の入力ベクトルとした。

本論文でも 1 枚の原画像から複数の特徴画像を生成し、深層学習に適用する手法を提案するが、特徴画像の枚数はより多く、さらに線形フィルタから非線形フィルタまで様々な画像フィルタを用いて特徴画像を生成する。特徴画像の組合せ数は膨大であるが、焼きなまし法を用いて組合せ最適化を行っている。

なお、本論文では特徴画像の選択を行うが、原画像から複数の特徴量を抽出したうえで、最適な特徴量を選択し、識別器に入力する手法が提案されている。これらの手法は一般的に特徴量選択と呼ばれる。原画像から対象物を識別するとき、画素値のコントラストや領域の面積、円形度など、あらゆる特徴量を抽出できるが、全ての特徴量が識別に役立つわけではない。そこで、識別に有用な特徴量を選択肢で識別器に入力する手法として、フィルタ法、ラッパー法、組み込み法が開発されている。フィルタ法は、それぞれの特徴量の予測に使える度合を点数化し、点数が高い特徴量のみを選択する手法である。ラッパー法は、それぞれの特徴量にランクを付与した後、複数の特徴量を組合せて識別器に入力し、精度が最大となる組合せを選択する手法である。ラッパー法には、ランクが高い特徴量の一つずつ加え、最適な特徴量の数を探索する前進法や、全ての特徴量からランクが低い特徴量の一つずつ除外し、最適な特徴量の数を探索する後退法が含まれる。組み込み法は、機械学習モデルの内部で特徴量選択を行う手法であり、ラッソ回帰や決定木、ランダムフォレストなどが含まれる [68]。Kursa ら [69] は、偽の特徴量と真の特徴量にランダムフォレストを適用して重要度を計算し、偽の特徴量より重要度の高い真の特徴量のみを選択する Boruta という特徴量選択手法を提案し、Madelon という人工的に作成した画像データセットを用いた識別問題に応用した。Zhang ら [70] は、global sensitivity analysis (GSA) に基づく重要度の順位付けとランダムフォレストを組合わせた特徴量選択手法を提案し、PCA による次元削減に対する優位性を示したうえで、トンネルによる地盤沈下の予測問題に応用した。Fahad ら [71] は、アントコロニー最適化に基づく特徴量選択手法を提案し、Iris データセットや SPECT 画像データセットの分類問題に応用した。根本ら [72] は、マンモグラムから乳房の悪性腫瘍影を検出するために、腫瘍の形状、濃度分布に関する 490 個の特徴量の中から判別に有用な特徴量を 2 段階選択法を用いて選択し、adaptive ring filter を用いて検出する手法を提案した。識別器としてマハラノビス距離を用いている。Ganesan ら [73] は、頭部 MR 画像から脳の腫瘍を分類するために、tolerance rough set firefly based quick reduct (TRSFFQR) という特徴量選択手法を提案した。識別器として J48, naïve byes, IBk を用いて性能比較を行った。Khan ら [74] は、頭部 MR 画像から脳の腫瘍を

分類するために、correntropy via mutual learning and extreme learning machine (CML-ELM) という特徴量選択手法と VGG-16, VFF-19 を組合わせた. Toğaçar ら [75] は, 胸部 X 線画像から肺炎の識別を行うために, minimum redundancy maximum relevance (mRMR) という特徴量選択手法と AlexNet[79], VGG-16, VGG-19 を組合わせた. Sartakhti ら [76] は, 年齢, 肥満の有無, ビリルビン値など 19 の特徴量から肝炎の予測を行うために, SVM と焼きなまし法を組合わせた特徴量選択手法を提案した. Asuntha ら [77] は, CT 画像, MR 画像, 超音波画像から癌を検出するために, 原画像にガボールフィルタ, グレースケール化, エッジ強調フィルタ, モルフォロジカルフィルタを順に適用して得られた画像から, 腫瘍の面積, 形状などの特徴量を抽出し, SVM で識別を行う手法を提案した. 最適な特徴量の組合せを選択するために, 粒子群最適化と遺伝的アルゴリズムをそれぞれ用いた.

### 4.2.3 実験

#### 4.2.3.1 特徴画像の作成

本論文では  $N = 46$  種類の特徴画像を作成し, それぞれの特徴画像に英数字 3 文字から 5 文字の略語を当てた. 特徴画像の種類と略語の一覧を表 4.13 に示すとともに, 各画像フィルタについて説明を加える. ただし, 特徴画像の画素値を 0 から 255 までの範囲内に収めるために, 後述する GRY と NML1 - 3 以外の 42 種類の特徴画像に NML2 を適用した.

表 4.13: 特徴画像の種類と略語の一覧

特徴画像の種類	略語
グレースケールの原画像そのもの	GRY
ヒストグラム平坦化フィルタ	HGE
ネガ・ポジ反転フィルタ	NGP
ソラリゼーションフィルタ	SOL
全ての画素値を 0 にするフィルタ	ZSP
平均化フィルタ (3 × 3 ピクセル)	MEA1
平均化フィルタ (5 × 5 ピクセル)	MEA2
ガウシアンフィルタ (3 × 3 ピクセル)	GAU1
ガウシアンフィルタ (5 × 5 ピクセル)	GAU2
メディアンフィルタ (3 × 3 ピクセル)	MED1
メディアンフィルタ (5 × 5 ピクセル)	MED2
K 最近傍平均化フィルタ (周囲 3 × 3 ピクセルから 3 画素)	KNN1
K 最近傍平均化フィルタ (周囲 5 × 5 ピクセルから 5 画素)	KNN2
バイラテラルフィルタ (3 × 3 ピクセル)	BLT1
バイラテラルフィルタ (5 × 5 ピクセル)	BLT2
ラプラシアンフィルタ (4 近傍)	LPL1



ラプラシアンフィルタ (8 近傍)	LPL2
E 字トーンカーブフィルタ ( $\gamma = 2$ )	ETC1
E 字トーンカーブフィルタ ( $\gamma = \frac{1}{2}$ )	ETC2
S 字トーンカーブフィルタ ( $a = 0$ )	STC1
S 字トーンカーブフィルタ ( $a = 1$ )	STC2
画素値正規化フィルタ ( $a = 1$ )	NML1
画素値正規化フィルタ ( $a = 2$ )	NML2
画素値正規化フィルタ ( $a = 3$ )	NML3
MED1 を適用後の local binary pattern フィルタ	LBP1
Local binary pattern フィルタ	LBP2
Dilated local binary pattern フィルタ ( $5 \times 5$ ピクセル)	LBP3
ポスタライゼーションフィルタ (10 値化)	POS1
ポスタライゼーションフィルタ (20 値化)	POS2
ポスタライゼーションフィルタ (30 値化)	POS3
Dilated エンボスフィルタ ( $3 \times 5$ ピクセル)	EMB1
Dilated エンボスフィルタ ( $5 \times 3$ ピクセル)	EMB2
エンボスフィルタ ( $3 \times 3$ ピクセル)	EMB3
MED1 を適用後の Top-hat 変換フィルタ ( $9 \times 9$ ピクセル)	TOP1
Top-hat 変換フィルタ ( $9 \times 9$ ピクセル)	TOP2
メディアンフィルタ ( $9 \times 9$ ピクセル) を適用後の Top-hat 変換フィルタ ( $9 \times 9$ ピクセル)	TOP3
メディアンフィルタ ( $9 \times 9$ ピクセル) を適用後の Top-hat 変換フィルタ ( $3 \times 3$ ピクセル)	TOP4
メディアンフィルタ ( $3 \times 3$ ピクセル) を 3 回適用後の Top-hat 変換フィルタ ( $9 \times 9$ ピクセル)	TOP5
Sobel フィルタ (水平方向)	SBLX1
Dilated Sobel フィルタ ( $3 \times 5$ ピクセル, 水平方向)	SBLX2
Dilated Sobel フィルタ ( $5 \times 5$ ピクセル, 水平方向)	SBLX3
Sobel フィルタ (垂直方向)	SBLY1
Dilated Sobel フィルタ ( $5 \times 3$ ピクセル, 垂直方向)	SBLY2
Dilated Sobel フィルタ ( $5 \times 5$ ピクセル, 垂直方向)	SBLY3
Sobel フィルタ (勾配)	SBLM
Sobel フィルタ (角度)	SBLD

#### ア) グレースケールの原画像そのもの (GRY)

原画像そのものに深層学習を適用する対照実験を行う目的で、本特徴画像を作成した。線形フィルタの一種である。

### イ) ヒストグラム平坦化フィルタ (HGE)

入力画像の画素値を横軸に，画素数を縦軸にとる関数を作成すると，通常は図 4.92 の左図のように画素値毎に画素数が異なる．ヒストグラム平坦化フィルタでは，画像内のコントラストの強調を目的とし，0 から 255 までの 256 段階の画素値における画素数を均等にする．人為的なパラメータの設定が不要であるという利点をもつ [78]．非線形フィルタの一種である．正常症例の VF フレーム画像を図 4.93 に，ヒストグラム平坦化フィルタを適用して得られた特徴画像を図 4.94 に示す．

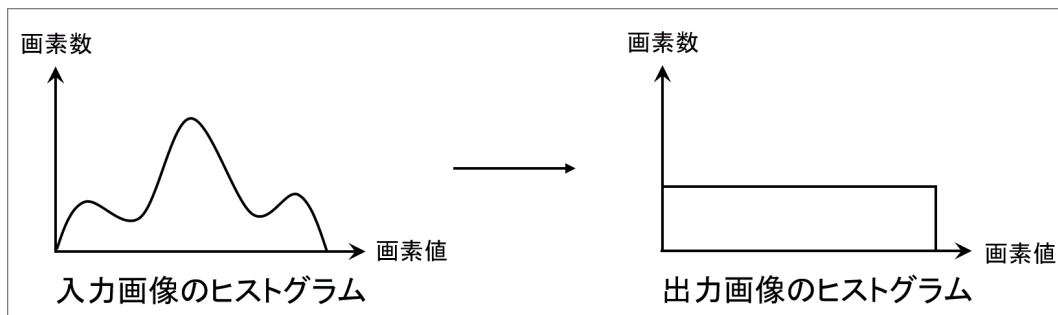


図 4.92: ヒストグラム平坦化フィルタ



図 4.93: 正常症例の VF フレーム画像



図 4.94: 特徴画像 HGE

### ウ) ネガ・ポジ反転フィルタ (NGP)

図 4.95 に示すように，入力画像の濃淡を反転させてネガフィルムのような効果を得る．入力画像の画素値を  $x$ ，出力画像の画素値を  $y$  と置いたとき， $y = 255 - x$  という関係が成立する．非線形フィルタの一種である．図 4.93 にネガ・ポジ反転フィルタを適用して得られた特徴画像を図 4.96 に示す．

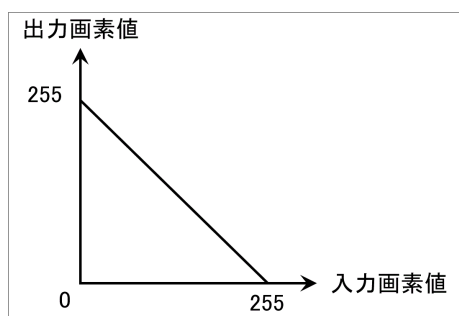


図 4.95: ネガ・ポジ反転フィルタ



図 4.96: 特徴画像 NGP

### エ) ソラリゼーションフィルタ (SOL)

図 4.97 に示すように，画像内のコントラストの強調を目的とし，画素値の一部を反転させることにより，ネガ画像とポジ画像が混ざり合ったような特殊な効果が生じる．入力画像，出力画像の画素値  $x, y$  に対し， $y = 255 \sin \frac{2\pi x}{255}$  という関係が成立する．非線形フィルタの一種である．図 4.93 にソラリゼーションフィルタを適用して得られた特徴画像を図 4.98 に示す．

### オ) 全ての画素値を 0 にするフィルタ (ZSP)

次元削減を行ったうえで深層学習を適用する対照実験を行う目的で，本特徴画像を作成した．線形フィルタの一種である．

### カ) 平均化フィルタ (MEA)

画像の平滑化を目的とし，入力画像の各画素を中心とする  $k \times k$  ピクセルのカーネルを作成し，カーネル内の画素値の平均値を出力画像の画素値とする．線形フィルタの一種である．

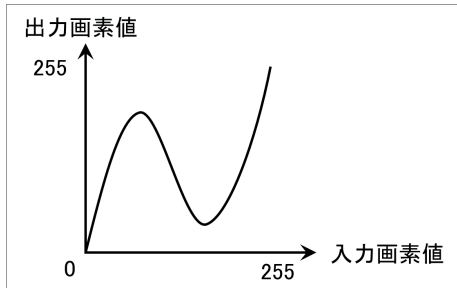


図 4.97: ソラリゼーションフィルタ



図 4.98: 特徴画像 SOL

MEA1 では  $k = 3$ , MEA2 では  $k = 5$  とした. 図 4.93 に MEA1, MEA2 をそれぞれ適用して得られた特徴画像を図 4.99, 図 4.100 に示す.



図 4.99: 特徴画像 MEA1

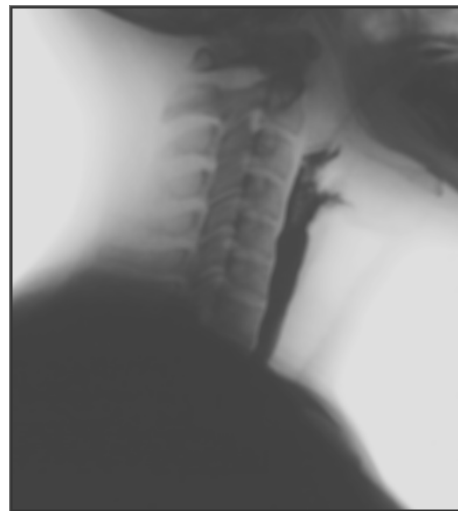


図 4.100: 特徴画像 MEA2

#### キ) ガウシアンフィルタ (GAU)

画像の平滑化を目的とし, 入力画像の各画素を中心とする  $k \times k$  ピクセルのカーネルを作成し, ガウス分布に従ってカーネルの中央に近いほど大きな重みをつける. 線形フィルタの一種である. GAU1 では  $k = 3$ , GAU2 では  $k = 5$  とした. 図 4.101 に GAU1 と GAU2 のカーネル





図 4.104: 特徴画像 MED1

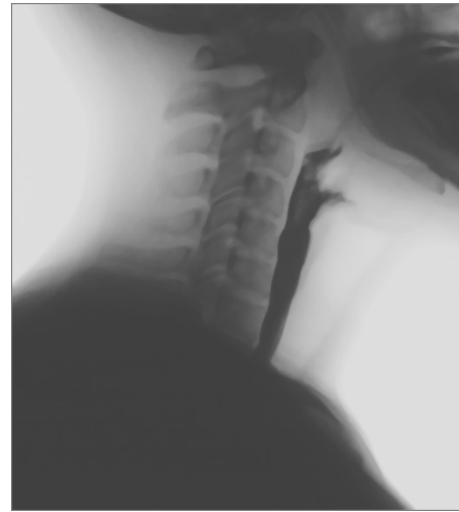


図 4.105: 特徴画像 MED2

#### ケ) **K** 最近傍平均化フィルタ (**KNN**)

画像の平滑化を目的とし，入力画像の各画素を中心とする  $k \times k$  ピクセルのカーネルを作成する．カーネル内の画素値をソートしたうえで，中心画素と近い値の画素値を  $a$  個選択し，それらの平均値を出力画像の画素値とする． $a \leq \frac{k^2}{2}$  とすることで画像内のエッジが保存されやすくなる．非線形フィルタの一種である．KNN1 では  $k = 3$ ， $a = 3$  とし，KNN2 では  $k = 5$ ， $a = 5$  とした．図 4.106 に KNN1，KNN2 の処理の具体例を示す．図 4.93 に KNN1，KNN2 をそれぞれ適用して得られた特徴画像を図 4.107，図 4.108 に示す．

#### コ) バイラテラルフィルタ (**BLT**)

画像の平滑化を目的とし，入力画像の各画素を中心とする  $k \times k$  ピクセルのカーネルを作成し，中心画素からの距離と画素値の差に応じて，ガウシアン分布に従う重みをつける．ガウシアンフィルタと異なり，画像内のエッジ情報を保存することが可能である．非線形フィルタの一種である．BLT1 では  $k = 3$ ，BLT2 では  $k = 5$  とした．図 4.93 に BLT1，BLT2 をそれぞれ適用して得られた特徴画像を図 4.109，図 4.110 に示す．

#### サ) ラプラシアンフィルタ (**LPL**)

画像内のエッジの強調を目的とし，カーネル内の水平方向の 2 次微分成分と垂直方向の 2 次微分成分を足し合わせることで，方向に依存しないエッジを得ることが可能である．線形フィルタの一種である．LPL1 では  $k = 3$  とし，中心画素の周囲 4 近傍に着目した．LPL2 では  $k = 3$  とし，中心画素の周囲 8 近傍に着目した．図 4.111 に LPL1 と LPL2 のカーネルを示す．図 4.93 に LPL1，LPL2 をそれぞれ適用して得られた特徴画像を図 4.112，図 4.113 に示す．

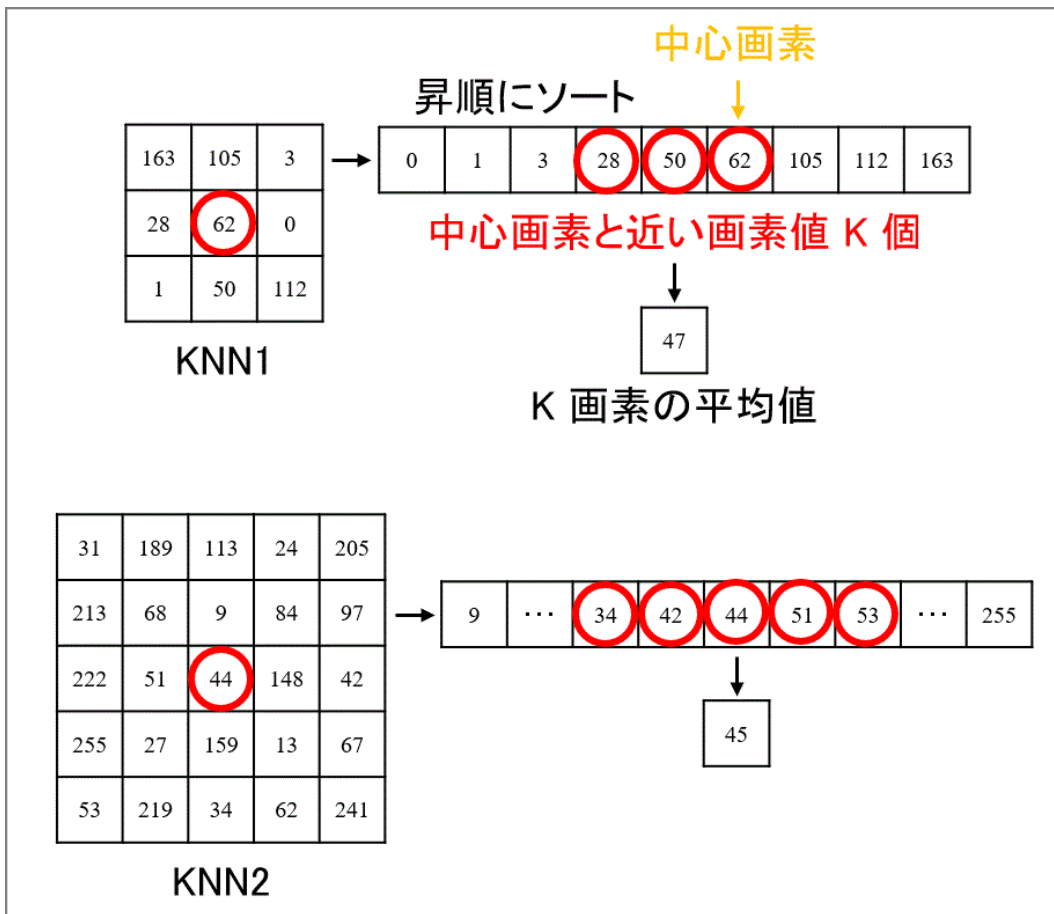


図 4.106: K 最近傍平均化フィルタの処理の具体例



図 4.107: 特徴画像 KNN1



図 4.108: 特徴画像 KNN2



図 4.109: 特徴画像 BLT1



図 4.110: 特徴画像 BLT2

0	1	0
1	-4	1
0	1	0

LPL1

1	1	1
1	-8	1
1	1	1

LPL2

図 4.111: ラプラシアンフィルタのカーネル

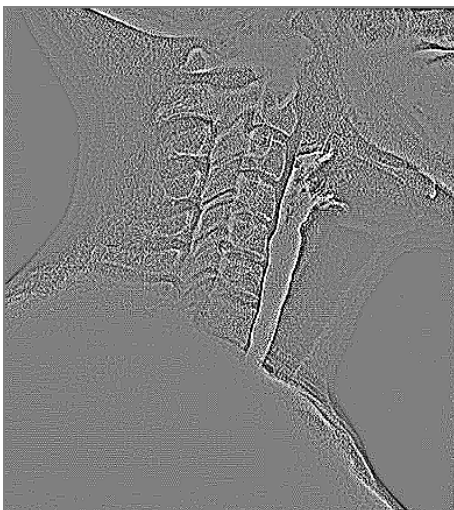


図 4.112: 特徴画像 LPL1

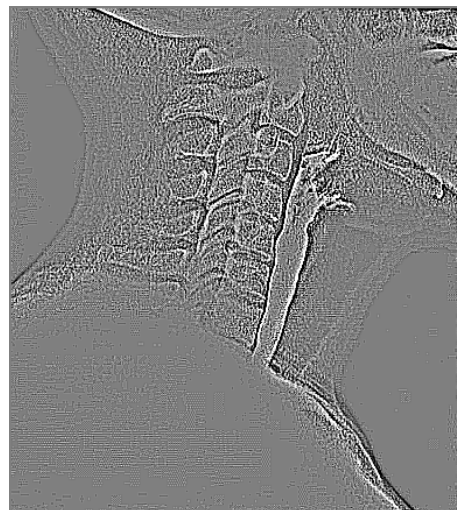


図 4.113: 特徴画像 LPL2



### シ) E字トーンカーブフィルタ (ETC)

画像内のコントラストの強調を目的とし，入力画像と出力画像の画素値を  $x, y$  に対して， $y = 255 \left(\frac{x}{255}\right)^{\frac{1}{\gamma}}$  という式で表される変換を行う．曲線を用いているため，画像内のコントラストを滑らかに変換することが可能である． $\gamma > 1$  としたとき上に凸， $\gamma < 1$  としたとき下に凸の曲線となる．非線形フィルタの一種である．ETC1 では  $\gamma = 2$ ，ETC2 では  $\gamma = \frac{1}{2}$  とした．図 4.93 に ETC1，ETC2 をそれぞれ適用して得られた特徴画像を図 4.114，図 4.115 に示す．



図 4.114: 特徴画像 ETC1



図 4.115: 特徴画像 ETC2

### ス) S字トーンカーブフィルタ (STC)

画像内のコントラストの強調を目的とし，入力画像と出力画像の画素値を  $x, y$  に対して， $y = \frac{255(\sin(\frac{\pi x}{255} - a) + 1)}{2}$  という式で表される変換を行う．中間調の画素値を引き伸ばす効果がある．非線形フィルタの一種である．STC1 では  $a = 0$ ，STC2 では  $a = 1$  とした．図 4.93 に STC1，STC2 をそれぞれ適用して得られた特徴画像を図 4.116，図 4.117 に示す．

### セ) 画素値正規化フィルタ (NML)

画像内のコントラストの強調を目的とし，画像内の画素値の平均  $n$  と分散  $\sigma$  を求め， $n - a\sigma$  から  $n + a\sigma$  までの画素値を 0 から 255 に線形変換する．非線形フィルタの一種である．NML1 では  $a = 1$ ，NML2 では  $a = 2$ ，NML3 では  $a = 3$  とした．図 4.93 に NML1，NML2，NML3 をそれぞれ適用して得られた特徴画像を図 4.118，図 4.119，図 4.120 に示す．



图 4.116: 特徴画像 STC1



图 4.117: 特徴画像 STC2



图 4.118: 特徴画像 NML1



图 4.119: 特徴画像 NML2

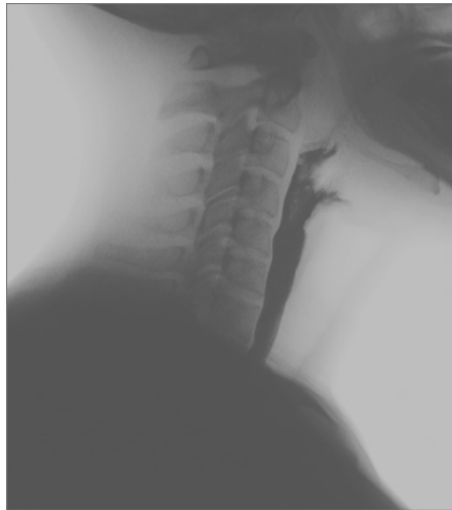


図 4.120: 特徴画像 NML3

### ソ) Local binary pattern フィルタ (LBP)

中心画素の画素値を周囲 8 近傍それぞれの画素値と比較し、大小を 0 か 1 で表す。0 と 1 の値を時計回りに並べ、8 桁の 2 進数とみなす。2 進数を 10 進数に変換し、出力画像の画素値とする。撮影時の明るさに依存せず、中心画素と周囲の画素の画素値の大小を比較することが可能である。非線形フィルタの一種である。図 4.121 に local binary pattern フィルタの処理の具体例を示す。カーネルサイズが  $3 \times 3$  ピクセルのメディアンフィルタを適用した後 local binary pattern フィルタを適用した特徴画像を LBP1、前処理を行わずに local binary pattern フィルタを適用した特徴画像を LBP2、前処理を行わずに  $5 \times 5$  ピクセルの dilated local binary pattern フィルタを適用した特徴画像を LBP3 と表す。なお、dilated フィルタとは、図 4.122 に示す黒塗りの画素を無視して  $3 \times 3$  ピクセルのカーネルを生成することを指す。広範囲の情報を出力画像の画素値に反映することができる。図 4.93 に LBP1、LBP2、LBP3 をそれぞれ適用して得られた特徴画像を図 4.123、図 4.124、図 4.125 に示す。

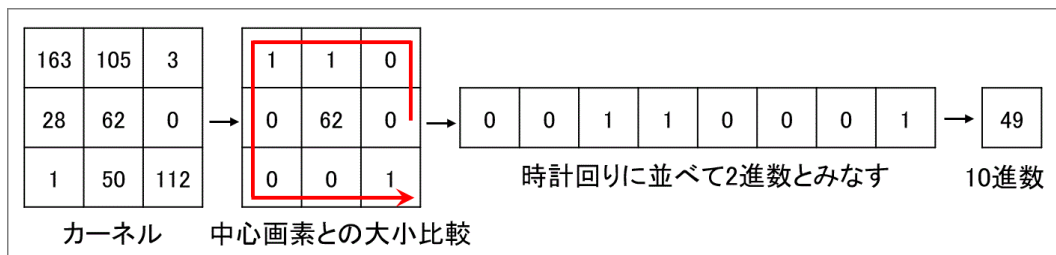


図 4.121: Local binary pattern フィルタの処理の具体例

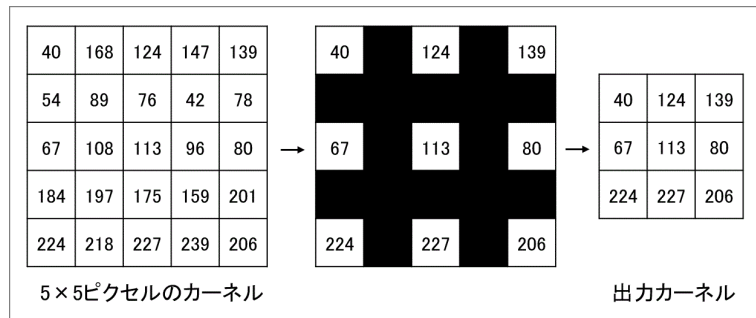


図 4.122: Dilated フィルタの処理の具体例



図 4.123: 特徴画像 LBP1



図 4.124: 特徴画像 LBP2



図 4.125: 特徴画像 LBP3

#### タ) ポスタライゼーションフィルタ (POS)

画像の情報量の削減を目的とし、0 から 255 までの画素値を  $n$  段階に丸め込む。特に  $n = 2$  のときには二値化と呼ばれる。非線形フィルタの一種である。POS1 では  $n = 10$ ，POS2 では  $n = 20$ ，POS3 では  $n = 30$  とした。図 4.93 に POS1，POS2，POS3 をそれぞれ適用して得られた特徴画像を図 4.126，図 4.127，図 4.128 に示す。



図 4.126: 特徴画像 POS1



図 4.127: 特徴画像 POS2



図 4.128: 特徴画像 POS3

### チ) エンボスフィルタ (EMB)

画像内のエッジの強調を目的とし、入力画像  $img_{in}$  にネガ・ポジ反転フィルタ、平行移動を順に適用した画像を  $img_{fil}$  と置いたとき、 $img_{in} + img_{fil} - 128$  を出力画像の画素値とする。エンボスフィルタのカーネルは図 4.129 のように表される。線形フィルタの一種である。カーネルサイズが  $3 \times 5$  ピクセルの dilated エンボスフィルタを EMB1 とし、カーネルサイズが  $5 \times 3$  ピクセルの dilated エンボスフィルタを EMB2 とし、カーネルサイズが  $3 \times 3$  ピクセルのエンボスフィルタを EMB3 とした。図 4.93 に EMB1, EMB2, EMB3 をそれぞれ適用して得られた特徴画像を図 4.130, 図 4.131, 図 4.132 に示す。

-2	-1	0
-1	1	1
0	1	2

図 4.129: エンボスフィルタのカーネル



図 4.130: 特徴画像 EMB1



図 4.131: 特徴画像 EMB2



図 4.132: 特徴画像 EMB3

### ツ) Top-hat 変換フィルタ (TOP)

モルフォロジー演算の一種であり、ノイズの削除を目的とする。入力画像  $img_{in}$  にモルフォロジー演算の erosion フィルタを適用した画像を  $img_{ero}$  と置いたとき、 $img_{in} - img_{ero}$  を出力画像の画素値とする。非線形フィルタの一種である。カーネルサイズが  $3 \times 3$  ピクセルのメディアンフィルタを適用した後、カーネルサイズが  $9 \times 9$  ピクセルの Top-hat 変換フィルタを適用して得られた特徴画像を TOP1 とする。前処理を行わずにカーネルサイズが  $9 \times 9$  ピクセルの Top-hat 変換フィルタを適用して得られた特徴画像を TOP2 とする。カーネルサイズが  $9 \times 9$  ピクセルのメディアンフィルタを適用した後、カーネルサイズが  $9 \times 9$  ピクセルの Top-hat 変換フィルタを適用して得られた特徴画像を TOP3 とする。カーネルサイズが  $9 \times 9$  ピクセルのメディアンフィルタを適用した後、カーネルサイズが  $3 \times 3$  ピクセルの Top-hat 変換フィルタを適用して得られた特徴画像を TOP4 とする。カーネルサイズが  $3 \times 3$  ピクセルのメディアンフィルタを 3 回繰り返し適用した後、カーネルサイズが  $9 \times 9$  ピクセルの Top-hat 変換フィルタを適用して得られた特徴画像を TOP5 とする。図 4.93 に TOP1 - 5 をそれぞれ適用して得られた特徴画像を図 4.133, 図 4.134, 図 4.135, 図 4.136, 図 4.137 に示す。



図 4.133: 特徴画像 TOP1



図 4.134: 特徴画像 TOP2

### テ) Sobel フィルタ (SBL)

画像内のエッジの強調を目的とし、微分と平滑化を同時に行う。水平方向、垂直方向それぞれの Sobel フィルタのカーネルを図 4.138 に示す。微分フィルタと比較してノイズが抑えられ、滑らかなエッジを抽出することができる。入力画像に水平方向、垂直方向の Sobel フィルタをそれぞれ適用して得られた画素値を  $x_1, x_2$  と置いたとき、勾配の Sobel フィルタの出力画素値  $y_m$  は  $y_m = \sqrt{x_1^2 + x_2^2}$  と表され、角度の Sobel フィルタの出力画素値  $y_d$  は  $y_d = \frac{255(\arctan \frac{x_2}{x_1} + \pi)}{2\pi}$

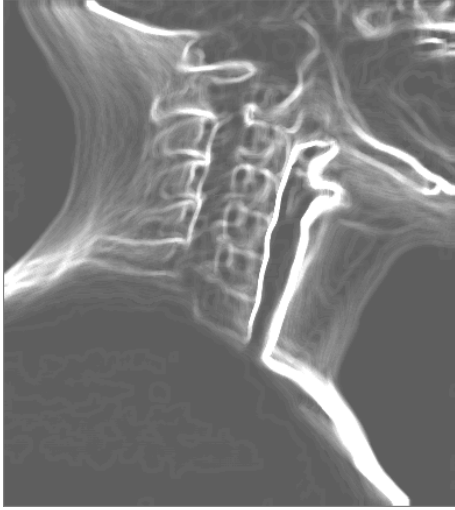


図 4.135: 特徴画像 TOP3

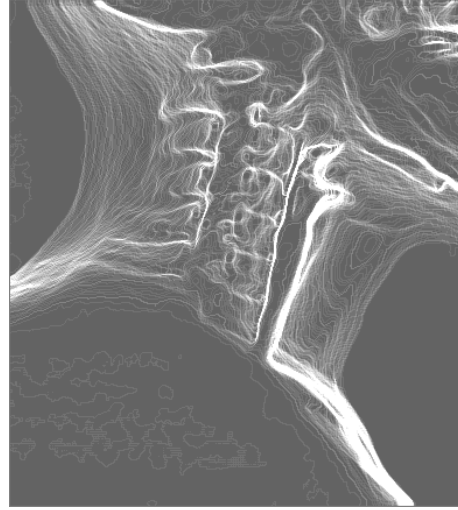


図 4.136: 特徴画像 TOP4



図 4.137: 特徴画像 TOP5



と表される。カーネルサイズが $3 \times 3$ ピクセルの水平方向の Sobel フィルタを SBLX1, カーネルサイズが $3 \times 5$ ピクセルの水平方向の dilated Sobel フィルタを SBLX2, カーネルサイズが $5 \times 5$ ピクセルの水平方向の dilated Sobel フィルタを SBLX3 と表す。カーネルサイズが $3 \times 3$ ピクセルの垂直方向の Sobel フィルタを SBLY1, カーネルサイズが $5 \times 3$ ピクセルの垂直方向の dilated Sobel フィルタを SBLY2, カーネルサイズが $5 \times 5$ ピクセルの垂直方向の dilated Sobel フィルタを SBLY3 と表す。カーネルサイズが $3 \times 3$ ピクセルの勾配の Sobel フィルタを SBLM, カーネルサイズが $3 \times 3$ ピクセルの角度の Sobel フィルタを SBLD と表す。SBLX1-3, SBLY1-3 は線形フィルタであり, SBLM, SBLD は非線形フィルタである。図 4.93 に SBLX1-3, SBLY1-3, SBLM, SBLD をそれぞれ適用して得られた特徴画像を図 4.139, 図 4.140, 図 4.141, 図 4.142, 図 4.143, 図 4.144, 図 4.145, 図 4.146 に示す。

-1	0	1	-1	-2	-1
-2	0	2	0	0	0
-1	0	1	1	2	1
水平方向の Sobel カーネル			垂直方向の Sobel カーネル		

図 4.138: 水平, 垂直方向の Sobel フィルタのカーネル



図 4.139: 特徴画像 SBLX1



図 4.140: 特徴画像 SBLX2



図 4.141: 特徴画像 SBLX3



図 4.142: 特徴画像 SBLY1



図 4.143: 特徴画像 SBLY2



図 4.144: 特徴画像 SBLY3



図 4.145: 特徴画像 SBLM

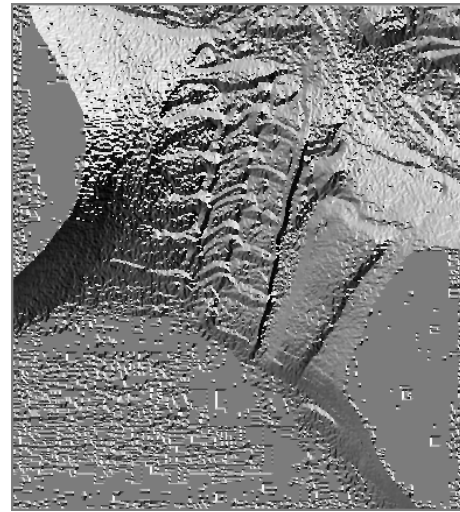


図 4.146: 特徴画像 SBLD

#### 4.2.3.2 ガウシアンフィルタを基準に特徴画像の組合せを変更した実験

##### 4.2.3.2.1 実験条件

マルチチャンネル化の効果を検証するため、チャンネル数を3とし、3チャンネルに全てに同一の特徴画像を入力した場合と、異なる特徴画像を入力した場合の椎間板抽出精度を比較する実験を行った。まず、4.2.3.1節で定義した特徴画像の内、MEA1, GAU1, MED1, LPL1, NML2, LBP2, TOP5, SBLX1, SBLY1, SBLM, SBLDの11種類を作成した。11種類のそれぞれの特徴画像を3チャンネル全てに保存して得られたMC画像に深層学習を適用し、画素単位のF値の降順に特徴画像をソートした。次に、GAU1を必ず含み、3チャンネル全ての特徴画像の種類が異なるように、全通りの組合せを深層学習に入力した。焼きなまし法に基づく最適な特徴画像の探索は行わなかった。ただし、4.1.3.10節と同一の患者群11症例と正常群24症例を合わせた35症例を学習症例に割り当て、本節では正常群3症例をテスト症例に割り当てた。 $S_p$ を21,  $prob$ を0.5とし、前処理としてパッチ毎の画素値正規化補正, 回転補正, データ拡張は行わず, 4.1.3.7節で作成した椎間板領域のデータと頸部マスクCを用い, 4.1.3.1節と同一の深層学習モデル, 計算機, フレームワークを用いて実験を行った。

##### 4.2.3.2.2 実験結果

3チャンネル全てに同一の特徴画像を保存して得られた3症例の平均F値を表4.14に示す。GAU1を含む全通りの特徴画像の組合せを保存して得られた3症例の平均F値を表4.15に示す。ただし、表4.15の特徴画像の順位とは、表4.14で特徴画像をF値の降順にソートして得られた順位を表す。表4.14の最大F値より表4.15の最大F値の方が大きい値となった。また、表4.14で得られた上位3つの特徴画像の組合せが表4.15で1位になるわけではないことがわかった。正常症例のVFフレーム画像を図4.147に、Ground truthを図4.148に、(GAU1, GAU1, GAU1)と(GAU1, TOP5, LBP2)をそれぞれ適用して得られた結果画像と同一症例のF

値を図 4.149, 図 4.150 に示す.

表 4.14: 4.2.3.2 節で得られた 11 種類の特徴画像の順位

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(GAU1, GAU1, GAU1)	1	0.611
(MEA1, MEA1, MEA1)	2	0.607
(MED1, MED1, MED1)	3	0.605
(SBLM, SBLM, SBLM)	4	0.600
(TOP5, TOP5, TOP5)	5	0.600
(NML2, NML2, NML2)	6	0.598
(LBP2, LBP2, LBP2)	7	0.585
(SBLY1, SBLY1, SBLY1)	8	0.573
(SBLD, SBLD, SBLD)	9	0.569
(LPL1, LPL1, LPL1)	10	0.560
(SBLX1, SBLX1, SBLX1)	11	0.543

表 4.15: GAU1 を含む全通りの特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(GAU1, TOP5, LBP2)	(1, 5, 7)	0.646
(GAU1, TOP5, LPL1)	(1, 5, 10)	0.645
(GAU1, SBLM, TOP5)	(1, 4, 5)	0.645
(GAU1, TOP5, SBLX1)	(1, 5, 11)	0.638
(GAU1, SBLM, LPL1)	(1, 4, 10)	0.632
(GAU1, TOP5, SBLD)	(1, 5, 9)	0.631
(GAU1, MEA1, SBLM)	(1, 2, 4)	0.628
(GAU1, SBLM, SBLD)	(1, 4, 9)	0.624
(GAU1, SBLM, SBLX1)	(1, 4, 11)	0.623
(GAU1, MED1, SBLM)	(1, 3, 4)	0.622
(GAU1, SBLM, LBP2)	(1, 4, 7)	0.621
(GAU1, MED1, SBLY1)	(1, 3, 8)	0.620
(GAU1, MED1, TOP5)	(1, 3, 5)	0.620
(GAU1, SBLM, NML2)	(1, 4, 6)	0.619
(GAU1, MEA1, TOP5)	(1, 2, 5)	0.613
(GAU1, TOP5, SBLY1)	(1, 5, 8)	0.613
(GAU1, MEA1, MED1)	(1, 2, 3)	0.612

(GAU1, LPL1, SBLX1)	(1, 10, 11)	0.611
(GAU1, MED1, NML2)	(1, 3, 6)	0.608
(GAU1, NML2, SBLX1)	(1, 6, 11)	0.608
(GAU1, LBP2, SBLX1)	(1, 7, 11)	0.607
(GAU1, MEA1, NML2)	(1, 2, 6)	0.606
(GAU1, TOP5, NML2)	(1, 5, 6)	0.604
(GAU1, MEA1, SBLX1)	(1, 2, 11)	0.604
(GAU1, MED1, LBP2)	(1, 3, 7)	0.602
(GAU1, LBP2, LPL1)	(1, 7, 10)	0.601
(GAU1, NML2, LBP2)	(1, 6, 7)	0.600
(GAU1, SBLM, SBLY1)	(1, 4, 8)	0.600
(GAU1, NML2, SBLD)	(1, 6, 9)	0.597
(GAU1, MEA1, LBP2)	(1, 2, 7)	0.597
(GAU1, SBLD, SBLX1)	(1, 9, 11)	0.595
(GAU1, MED1, SBLX1)	(1, 3, 11)	0.592
(GAU1, SBLY1, SBLX1)	(1, 8, 11)	0.589
(GAU1, NML2, LPL1)	(1, 6, 10)	0.587
(GAU1, LBP2, SBLD)	(1, 7, 9)	0.586
(GAU1, MED1, LPL1)	(1, 3, 10)	0.585
(GAU1, MED1, SBLD)	(1, 3, 9)	0.582
(GAU1, MEA1, LPL1)	(1, 2, 10)	0.580
(GAU1, LBP2, SBLY1)	(1, 7, 8)	0.574
(GAU1, MEA1, SBLD)	(1, 2, 9)	0.573
(GAU1, NML2, SBLY)	(1, 6, 8)	0.571
(GAU1, SBLD, LPL1)	(1, 9, 10)	0.569
(GAU1, SBLY1, LPL1)	(1, 8, 10)	0.569
(GAU1, MEA1, SBLY1)	(1, 2, 8)	0.557
(GAU1, SBLY1, SBLD)	(1, 8, 9)	0.554

#### 4.2.3.3 21 種類の特徴画像に焼きなまし法を適用した実験

##### 4.2.3.3.1 実験条件

マルチチャンネル化と焼きなまし法に基づく特徴画像の選択により、椎間板抽出精度が向上するか否かを検討する実験を行った。まず、4.2.3.1 節で定義した特徴画像の内、GRY, HGE, SOL, ZSP, MEA1, GAU1, MED1, KNN1, BLT1, LPL1, ETC1, STC1, NML2, LBP2, POS1, EMB3, TOP5, SBLX1, SBLY1, SBLM, SBLD の 21 種類を作成した。チャンネル



図 4.147: 正常症例の VF フレーム画像

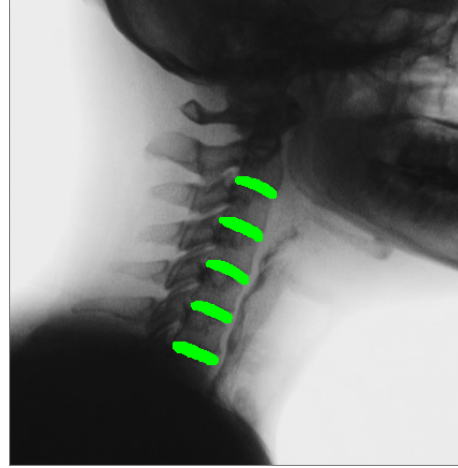


図 4.148: 正常症例の Ground truth

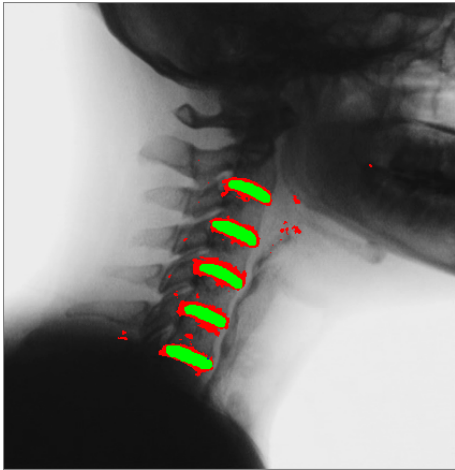


図 4.149: (GAU1, GAU1, GAU1) を適用して得られた結果画像 (F 値 =0.664)

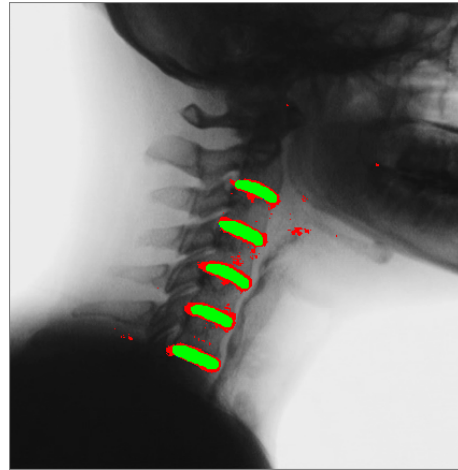


図 4.150: (GAU1, TOP5, LBP2) を適用して得られた結果画像 (F 値 =0.733)

数を3とし、21種類の特徴画像のそれぞれを3チャンネル全てに保存して得られたMC画像に深層学習を適用し、画素単位のF値の降順に特徴画像をソートした。次に、4.2.1節に示した手順で、焼きなまし法に基づく最適な特徴画像の組合せの探索を行った。ただし、4.2.1節では $i, j, k \in \{1, 2, \dots, N\}$ としていたのに対し、本節ではソート後の1位の特徴画像が中心になるようにインデックスを折り返して $i, j, k \in \{1, 2, \dots, 2N-1\}$ とした。変更したインデックスが0以下あるいは $2N$ 以上となった場合には、1あるいは $2N-1$ を基点にインデックスを折り返した。焼きなましの初期温度を $T = 10000$ 、冷却スケジュールを $T = T \times 0.99$ 、終了条件を $T \leq 0.0001$ 、特徴画像のインデックスの遷移範囲を $step = 2$ とした。ただし、4.1.3.7節で作成した椎間板領域のデータと頸部マスクCを4.1.3.10節と同一の学習、テスト症例に適用した。 $S_p$ を21、 $prob$ を0.5とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張は行わず、4.1.3.1節と同一の深層学習モデル、計算機、フレームワークを用いて実験を行った。[45][51][52]を参考にプログラムを作成した。

#### 4.2.3.3.2 実験結果

3チャンネル全てに同一の特徴画像を保存して得られた21種類の特徴画像の順位を表4.16に、焼きなまし法を適用して得られた上位10種類の特徴画像の組合せを表4.17に示す。焼きなまし法で得られた最適な特徴画像の組合せには、非線形フィルタであるSBLMとSOLが含まれていた。また、表4.16の最大F値より表4.17の最大F値の方が大きい値となった。図4.147に(GRY, GRY, GRY)と(SBLM, SOL, LPL1)をそれぞれ適用して得られた結果画像と同一症例のF値を図4.151、図4.152に示す。

表 4.16: 4.2.3.3 節で得られた 21 種類の特徴画像の順位

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(GRY, GRY, GRY)	1	0.568
(KNN1, KNN1, KNN1)	2	0.565
(BLT1, BLT1, BLT1)	3	0.563
(GAU1, GAU1, GAU1)	4	0.561
(MEA1, MEA1, MEA1)	5	0.560
(TOP5, TOP5, TOP5)	6	0.553
(NML2, NML2, NML2)	7	0.553
(ETC1, ETC1, ETC1)	8	0.553
(MED1, MED1, MED1)	9	0.552
(SBLY1, SBLY1, SBLY1)	10	0.551
(EMB3, EMB3, EMB3)	11	0.543
(SBLM, SBLM, SBLM)	12	0.542
(STC1, STC1, STC1)	13	0.535
(SOL, SOL, SOL)	14	0.528

(HGE, HGE, HGE)	15	0.527
(SBLD, SBLD, SBLD)	16	0.527
(LPL1, LPL1, LPL1)	17	0.520
(LBP2, LBP2, LBP2)	18	0.514
(POS1, POS1, POS1)	19	0.511
(SBLX1, SBLX1, SBLX1)	20	0.502
(ZSP, ZSP, ZSP)	21	0.0785

表 4.17: 4.2.3.3 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLM, SOL, LPL1)	(12, 14, 17)	0.605
(TOP5, SOL, SBLM)	(6, 14, 12)	0.605
(SOL, TOP5, LBP2)	(14, 6, 18)	0.599
(TOP5, SOL, SBLY1)	(6, 14, 10)	0.598
(TOP5, SOL, TOP5)	(6, 14, 6)	0.595
(SBLM, SBLD, ZSP)	(12, 16, 21)	0.592
(GRY, SOL, SBLM)	(1, 14, 12)	0.592
(SBLM, SBLX1, SOL)	(12, 20, 14)	0.590
(TOP5, SBLM, EMB3)	(6, 12, 11)	0.589
(ZSP, SOL, SBLY1)	(21, 14, 10)	0.589

#### 4.2.3.4 学習，テスト症例と焼きなまし条件を変更して 21 種類の特徴画像に焼きなまし法を適用した実験

##### 4.2.3.4.1 実験条件

4.2.3.3 節から学習，テスト症例の分割と焼きなまし条件を変更し，マルチチャンネル化と焼きなまし法に基づく特徴画像の選択により，椎間板抽出精度が向上するか否かを検討する実験を行った．本節では，患者群 13 症例と正常群 26 症例を合わせた 39 症例を学習症例として用い，残りの患者群 6 症例と正常群 13 症例を合わせた 19 症例をテスト症例として用いた．また，4.2.3.3 節では，特徴画像のインデックスを  $i, j, k \in \{1, 2, \dots, 2N - 1\}$  とし，変更後のインデックスが 0 以下あるいは  $2N$  以上となった場合にはインデックスを折り返していたのに対し，本節では，4.2.1 節に示す手順で， $i, j, k \in \{1, 2, \dots, N\}$  とし，変更したインデックスが 0 以下あるいは  $N + 1$  以上となった場合にはもう一度インデックスを変更した．ただし，特徴画像の種類，チャンネル数，焼きなましの初期温度，冷却スケジュール，終了条件，*step*



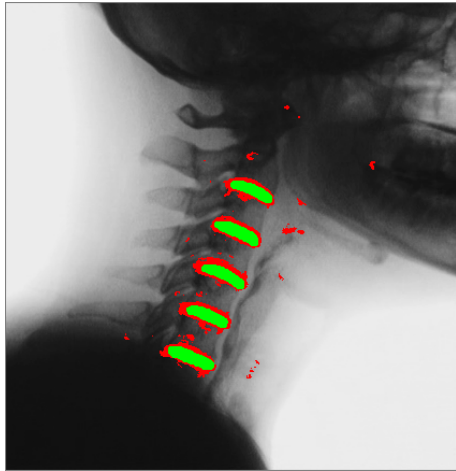


図 4.151: (GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.650)

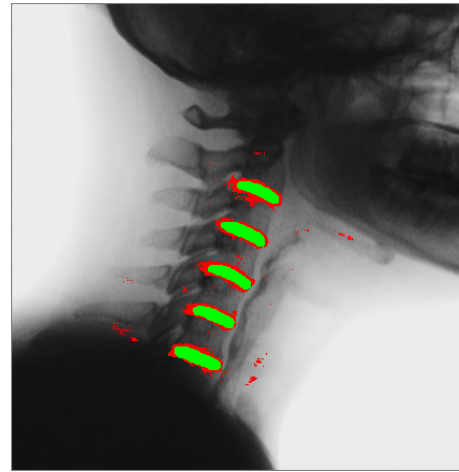


図 4.152: (SBLM, SOL, LPL1) を適用して得られた結果画像 (F 値=0.701)

は 4.2.3.3 節と統一し、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、 $S_p$  を 21,  $prob$  を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張は行わず、4.1.3.1 節と同一の深層学習モデル、計算機、フレームワークを用いて実験を行った。

#### 4.2.3.4.2 実験結果

3 チャンネル全てに同一の特徴画像を保存して得られた 21 種類の特徴画像の順位を表 4.18 に、焼きなまし法を適用して得られた上位 10 種類の特徴画像の組合せを表 4.19 に示す。焼きなまし法で得られた最適な特徴画像の組合せには、非線形フィルタである SBLM と TOP5 が含まれていた。また、表 4.18 の最大 F 値より表 4.19 の最大 F 値の方が大きい値となった。焼きなまし法で遷移が採択された回数は 1833 回であり、採択されなかった回数は 1354 回であった。4.2.3.3 節と本節とでは、同一の特徴画像の順位も最適な特徴画像の組合せも異なっていた。したがって、特徴画像の順位はデータセットや焼きなまし条件に依存することが示唆される。正常症例の VF フレーム画像を図 4.153 に、Ground truth を図 4.154 に、(GRY, GRY, GRY) と (SBLM, TOP5, SBLM) をそれぞれ適用して得られた結果画像と同一症例の F 値を図 4.155, 図 4.156 に示す。

表 4.18: 4.2.3.4 節で得られた 21 種類の特徴画像の順位

特徴画像の組合せ	特徴画像の順位	平均 F 値
(GRY, GRY, GRY)	1	0.562
(MED1, MED1, MED1)	2	0.552
(MEA1, MEA1, MEA1)	3	0.551

(NML2, NML2, NML2)	4	0.547
(GAU1, GAU1, GAU1)	5	0.544
(BLT1, BLT1, BLT1)	6	0.544
(KNN1, KNN1, KNN1)	7	0.544
(STC1, STC1, STC1)	8	0.543
(ETC1, ETC1, ETC1)	9	0.539
(SBLM, SBLM, SBLM)	10	0.533
(SBLY1, SBLY1, SBLY1)	11	0.530
(TOP5, TOP5, TOP5)	12	0.525
(EMB3, EMB3, EMB3)	13	0.523
(HGE, HGE, HGE)	14	0.511
(POS1, POS1, POS1)	15	0.501
(SOL, SOL, SOL)	16	0.499
(LPL1, LPL1, LPL1)	17	0.498
(SBLX1, SBLX1, SBLX1)	18	0.494
(SBLD, SBLD, SBLD)	19	0.493
(LBP2, LBP2, LBP2)	20	0.487
(ZSP, ZSP, ZSP)	21	0.0732

表 4.19: 4.2.3.4 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLM, TOP5, SBLM)	(10, 12, 10)	0.593
(SBLM, SOL, TOP5)	(10, 16, 12)	0.593
(TOP5, SBLM, GAU1)	(12, 10, 5)	0.592
(GRY, SBLM, SBLD)	(1, 10, 19)	0.590
(STC1, TOP5, SBLM)	(8, 12, 10)	0.590
(SBLM, TOP5, TOP5)	(10, 12, 12)	0.589
(SBLD, SBLM, GAU1)	(19, 10, 5)	0.589
(STC1, SBLM, KNN1)	(8, 10, 7)	0.589
(STC1, SBLM, BLT1)	(8, 10, 6)	0.588
(GRY, SBLM, SBLX1)	(1, 10, 18)	0.586



図 4.153: 正常症例の VF フレーム画像

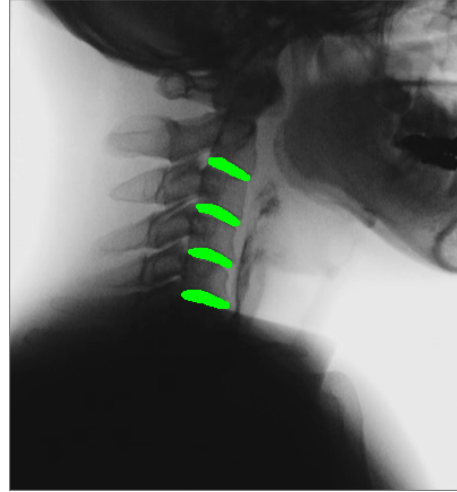


図 4.154: 正常症例の Ground truth



図 4.155: (GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.521)

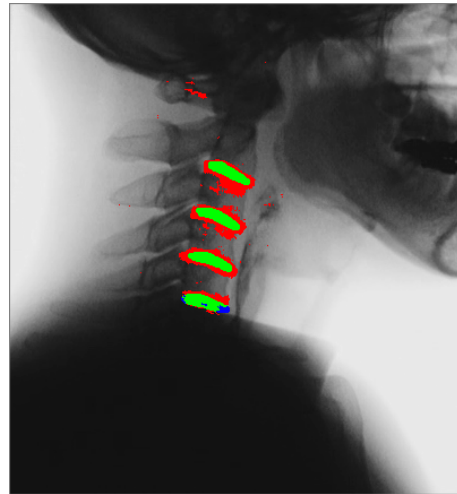


図 4.156: (SBLM, TOP5, SBLM) を適用して得られた結果画像 (F 値=0.596)

#### 4.2.3.5 焼きなまし条件を変更して 45 種類の特徴画像に焼きなまし法を適用した実験

##### 4.2.3.5.1 実験条件

4.2.3.4 節から、使用する特徴画像の種類と焼きなまし条件を変更し、マルチチャンネル化と焼きなまし法に基づく特徴画像の選択により、椎間板抽出精度が向上するか否かを検討する実験を行った。本節では、4.2.3.1 節で定義した特徴画像の内、TOP5 を除く 45 種類を作成した。焼きなましの初期温度を  $T = 10000$ ，冷却スケジュールを  $T = T \times 0.995$ ，終了条件を  $T \leq 0.0001$  とした。Step の初期値を  $StepMax = 5$  とし、 $Step = \frac{StepMax \times T}{10000}$  (ただし、 $Step < 1$  のときは  $Step = 1$ ) とした。ただし、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を 4.2.3.4 節と同一の学習，テスト症例に適用し，チャンネル数を 3， $S_p$  を 21， $prob$  を 0.5 とし，前処理としてパッチ毎の画素値正規化補正，回転補正，データ拡張は行わず，4.1.3.1 節と同一の深層学習モデル，計算機，フレームワークを用いて実験を行った。

##### 4.2.3.5.2 実験結果

3 チャンネル全てに同一の特徴画像を保存して得られた上位 10 種類の特徴画像を表 4.20 に、焼きなまし法を適用して得られた上位 10 種類の特徴画像の組合せを表 4.21 に示す。焼きなまし法で得られた最適な特徴画像の組合せには、非線形フィルタである STC1, TOP1, POS2 が含まれていた。また、表 4.20 の最大 F 値より表 4.21 の最大 F 値の方が大きい値となった。焼きなまし法で遷移が採択された回数は 3675 回であり、採択されなかった回数は 1685 回であった。なお、4.2.3.4 節では 21 種類の特徴画像を用いたのに対し、本節では 45 種類の特徴画像を用いたが、4.2.3.4 節では、3 チャンネル全てに同一の特徴画像を保存して得られた最大 F 値が 0.562 であったのに対し、焼きなまし法で得られた最大 F 値は 0.593 となった。一方、本節では、3 チャンネル全てに同一の特徴画像を保存して得られた最大 F 値が 0.568 であったのに対し、焼きなまし法で得られた最大 F 値は 0.588 となった。4.2.3.4 節と本節の焼きなまし条件の差異を無視すれば、特徴画像を 21 種類に限定した方が 45 種類を用いるより F 値の上昇幅が大きくなった。理由として、特徴画像の枚数が少なければ探索空間内を満遍なく探索することができ、大域的最適解を見つけやすくなるを考える。なお、4.2.3.4 節と本節で同じ学習，テスト症例を用いて (GRY, GRY, GRY) を適用する実験を行ったところ、4.2.3.4 節の平均 F 値は 0.562，本節の平均 F 値は 0.568 と異なる値となった。理由として、DIGITS をフレームワークとして用いて CNN の学習を行う際に、ミニバッチ法や dropout で乱数を用いており、random seed を固定しても乱数を完全には固定できないことが挙げられる。図 4.153 に (GRY, GRY, GRY) と (STC1, TOP1, POS2) をそれぞれ適用して得られた結果画像と同一症例の F 値を図 4.157，図 4.158 に示す。

表 4.20: 4.2.3.5 節で得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の順位	平均 F 値
(GRY, GRY, GRY)	1	0.568
(GAU1, GAU1, GAU1)	2	0.560

(ETC2, ETC2, ETC2)	3	0.557
(KNN1, KNN1, KNN1)	4	0.552
(MEA1, MEA1, MEA1)	5	0.552
(MED1, MED1, MED1)	6	0.549
(BLT1, BLT1, BLT1)	7	0.549
(NML2, NML2, NML2)	8	0.548
(NML1, NML1, NML1)	9	0.546
(BLT2, BLT2, BLT2)	10	0.546

表 4.21: 4.2.3.5 節で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(STC1, TOP1, POS2)	(22, 11, 29)	0.588
(TOP1, ETC2, NML3)	(11, 3, 15)	0.587
(STC1, NGP, EMB1)	(22, 13, 38)	0.586
(KNN1, ETC2, TOP1)	(4, 3, 11)	0.586
(TOP1, KNN1, STC2)	(11, 4, 24)	0.585
(KNN2, MEA1, STC2)	(12, 5, 24)	0.583
(NML3, KNN1, STC1)	(15, 4, 22)	0.581
(STC1, MEA2, EMB1)	(22, 17, 38)	0.581
(GAU2, NGP, EMB1)	(19, 13, 38)	0.580
(NML3, NGP, EMB1)	(15, 13, 38)	0.579

#### 4.2.3.6 3-fold 3-phase cross validation を用いて LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet を比較した実験

##### 4.2.3.6.1 実験条件

本節では、深層学習モデルとして LeNet, AlexNet, U-Net, Feature pyramid network (FPN)[81], LinkNet[82], Pyramid scene parsing network (PSPNet)[83] の 6 種類をそれぞれ用いて、深層学習とマルチチャンネル化、特徴画像の選択に基づく椎間板領域の抽出を行った。AlexNet は、2012 年に開催された画像認識コンペティション“ImageNet Large Scale Visual Recognition Competition (ILSVRC)”にて、SIFT, Fisher vector, SVM を組合わせた従来手法に大差をつけて優勝した CNN モデルである。計算機環境に対応するため、特徴マップをチャンネル方向に分割し、2 台の GPU を用いて並列計算を行うモデルを提唱した。論文内で、確率的勾配降下法や L2 正則化の概念も提唱されている。U-Net は、全層畳み込みネットワーク (Fully convolutional

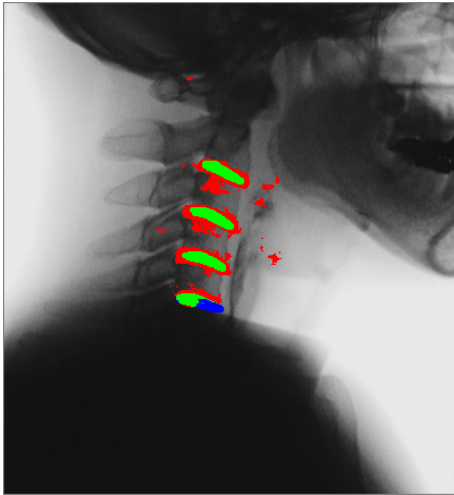


図 4.157: (GRY, GRY, GRY) を適用して得られた結果画像 (F 値=0.533)

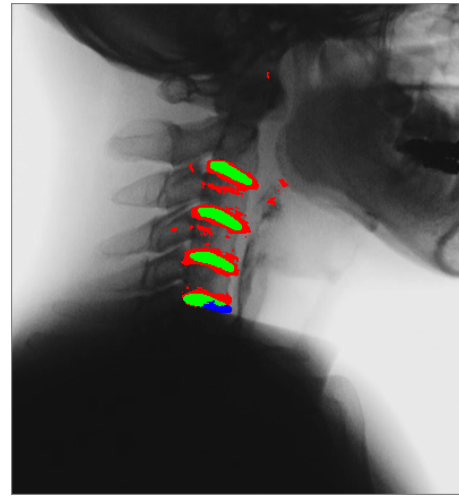


図 4.158: (STC1, TOP1, POS2) を適用して得られた結果画像 (F 値=0.559)

network, FCN) に位置づけられるとともに、画像内の領域抽出を行う semantic segmentation モデルの一種である。ネットワークの構造は decoder 部分と encoder 部分に分かれており、decoder に encoder の特徴マップを連結することで、画像内の位置情報を保持しながら領域抽出を行うことができる。FPN も、同じく semantic segmentation モデルの一種である。異なる階層の特徴マップを連結し、ヒートマップとすることで、低レイヤのコンテキストを活用しながら領域抽出を行うことができる。LinkNet は、ストライド畳み込み層という概念を取り入れた semantic segmentation モデルである。ハイパーパラメータの数を削減することで、メモリの消費を抑えながら学習効率を上げることに成功した。PSPNet は、一枚の画像から複数の解像度の画像を作成し、それぞれに畳み込み層とプーリング層を適用したうえで連結することで、画像内の領域抽出を行う semantic segmentation モデルである。ハイパーパラメータの数を削減し、なおかつ、低レイヤのコンテキストを活用しながら領域抽出を行うことができる。本節では、LeNet, AlexNet を用いてパッチベースの分類を行い、U-Net, FPN, LinkNet, PSPNet を用いて semantic segmentation を行った。実験環境に対応するために、AlexNet のカーネルサイズとストライドを一部修正した。表 4.22 に、6 つの深層学習モデルそれぞれにおける画像サイズ、転移学習の有無、エポック数を示す。ただし、深層学習モデルの P はパッチベースを意味し、S は semantic segmentation を意味する。パッチベースの 2 つのモデルでは、画像サイズはパッチサイズを意味し、semantic segmentation の 4 つのモデルでは、画像サイズは入力画像をリサイズして得られた画像の大きさを意味する。LeNet, AlexNet, U-Net では転移学習を行わずにスクラッチから学習を行い、FPN, LinkNet, PSPNet では ImageNet を用いて転移学習を行った。転移学習のバックボーンアーキテクチャとして residual network-34 (ResNet-34)[84] を使用した。エポック数はそれぞれの深層学習モデルで実験的に決定した。

深層学習による評価実験の手順は、学習 (検証を含む) とテストの二つのフェーズから構成

表 4.22: 深層学習のパラメータ

深層学習モデル		画像サイズ	転移学習の有無	エポック数
LeNet	P	21 × 21	なし	30
AlexNet	P	21 × 21	なし	50
U-Net	S	256 × 256	なし	90
FPN	S	256 × 256	あり	135
LinkNet	S	256 × 256	あり	165
PSPNet	S	288 × 288	あり	175

される。本節の実験では、それに加え、特徴画像の組合せを評価する手順が必要となる。そこで、①学習 (検証を含む)、②特徴評価、③テストの3つのフェーズから構成される 3-fold 3-phase cross validation を提案する。データセットを3つのグループ A, B, C に分割し、①から③のフェーズを順番に適用する。グループとフェーズの組み合わせを入れ替えて3回実験を行い、評価値 (F 値) の平均値を求めて評価する。なお、各グループで患者群と正常群がなるべく均等な人数になるようにする。表 4.23 に 3-fold 3-phase cross validation を用いた実験の流れを示す。

表 4.23: 3-fold 3-phase cross validation を用いた実験の流れ

	①学習	②特徴評価	③テスト
実験 1	A	B	C
実験 2	B	C	A
実験 3	C	A	B

本節では、4.2.3.1 節で定義した特徴画像の内、TOP5 を除く 45 種類を使用し、3-fold 3-phase cross validation を用いて実験を行った。グループ A に患者群 6 症例と正常群 13 症例を合わせた 19 症例を、グループ B に患者群 6 症例と正常群 13 症例を合わせた 19 症例を、グループ C に患者群 7 症例と正常群 13 症例を合わせた 20 症例を割り当てた。4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わなかった。チャンネル数を 3 とした。焼きなましの初期温度を  $T = 10000$  とし、冷却スケジュールは、深層学習の学習時間の関係で LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet の順に  $T = T \times 0.95$ ,  $T = T \times 0.8$ ,  $T = T - 100$ ,  $T = T - 100$ ,  $T = T - 100$ ,  $T = T - 100$  とした。終了条件を  $T \leq 1$  とした。探索空間内の広範囲を探索できるように、step の初期値を順に  $StepMax = 23$ ,  $StepMax = 23$ ,  $StepMax = 5$ ,  $StepMax = 5$ ,  $StepMax = 5$ ,  $StepMax = 5$  とし、 $Step = \frac{StepMax \times T}{10000}$  (ただし、 $Step < 1$  のときは  $Step = 1$ ) とした。prob を 0.5 とした。本節では、U-Net, FPN, LinkNet, PSPNet の結果画像として 3 チャンネルのカラー画像が出力されたため、後処理として、画像サイズの変更、グレースケール化、判別分析法による二値化を行い、頸部マスク外の領域を除外したうえで画

素単位の F 値を求めた。3-fold 3-phase cross validation の実験 1, 実験 2, 実験 3 で得られた最適な特徴画像の組合せは異なるものの, 3 グループの合計 58 症例の F 値を平均し, テスト結果とした。本節の実験に使用した計算機は, OS Linux 64 ビット Ubuntu 16.04, メモリ 62.8 GB, プロセッサ Intel® Core™ i9-9900K CPU @ 3.60GHz × 16, グラフィック GeForce RTX 2080 Ti/PCIe/SSE2 であり, 使用したフレームワークは Tensorflow-1.14.0 GPU[85], Keras-2.3.1[86] であった。本フレームワークを用いることで, 深層学習モデル内の乱数を固定することが可能となり, 再現性が得られた。文献 [53][54] を参考にプログラムを作成した。

#### 4.2.3.6.2 実験結果

3-fold 3-phase cross validation の実験 1 に LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet をそれぞれ適用して得られた上位 10 種類の特徴画像を表 4.24, 表 4.25, 表 4.26, 表 4.27, 表 4.28, 表 4.29 に示す。実験 1 に焼きなまし法と LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet をそれぞれ適用して得られた上位 10 種類の特徴画像の組合せを表 4.30, 表 4.31, 表 4.32, 表 4.33, 表 4.34, 表 4.35 に示す。6 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せをそれぞれ入力して得られた平均 F 値と, 3 グループの最適な特徴画像の組合せをまとめて表 4.36 に示す。ただし, P はパッチベースの深層学習モデルを, S は semantic segmentation モデルを表す。MC なしとはテストで (GRY, GRY, GRY) を深層学習に入力したもので, MC ありとはテストで最適な特徴画像の組合せを深層学習に入力したものを指す。\* は  $p < 0.05$ , \*\* は  $p < 0.01$  を表す。AlexNet, U-Net, FPN では, 最適な特徴画像の組合せを入力して得られた平均 F 値が (GRY, GRY, GRY) を入力して得られた平均 F 値と比較して有意に大きかった。6 つの深層学習モデルに図 4.153 の (GRY, GRY, GRY) をそれぞれ適用して得られた結果画像と同一症例の F 値を図 4.159, 図 4.160, 図 4.161, 図 4.162, 図 4.163, 図 4.164 に示す。6 つの深層学習モデルに最適な特徴画像の組合せをそれぞれ適用して得られた結果画像と同一症例の F 値を図 4.165, 図 4.166, 図 4.167, 図 4.168, 図 4.169, 図 4.170 に示す。3-fold 3-phase cross validation の実験 1 に焼きなまし法を適用した際, 遷移が採択された回数は深層学習モデルの順に 180 回, 42 回, 100 回, 100 回, 100 回, 100 回であり, 遷移が採択されなかった回数は 6 つの深層学習モデル全てで 0 回であった。同じく実験 1 に焼きなまし法を適用した際, 特徴画像の組合せは順に 208 通り, 83 通り, 136 通り, 142 通り, 140 通り, 138 通り生成され, 計算時間は順に約 8.7 日, 約 33.8 日, 約 16.4 日, 約 13.4 日, 約 5.4 日, 約 4.2 日となった。

#### 4.2.3.7 焼きなまし条件を変更して U-Net, FPN, LinkNet, PSPNet を比較した実験

##### 4.2.3.7.1 実験条件

4.2.3.6 節では, 焼きなましの温度が高すぎたため, 特徴画像の変更前後のエネルギー値の大小にかかわらず, 遷移が毎回採択されてしまった。本節では, 4.2.3.6 節から焼きなまし条件を変更し, U-Net, FPN, LinkNet, PSPNet を用いて, 深層学習とマルチチャンネル化, 特徴画像の選択に基づく椎間板領域の抽出を行った。焼きなましの初期温度を  $T = 500$ , 冷却スケ



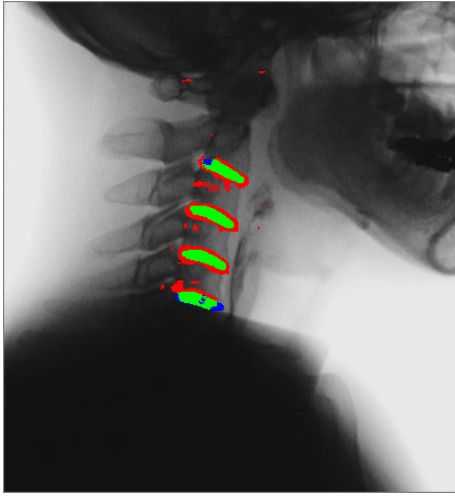


図 4.159: (GRY, GRY, GRY) に LeNet を適用して得られた結果画像 (F 値 =0.634)

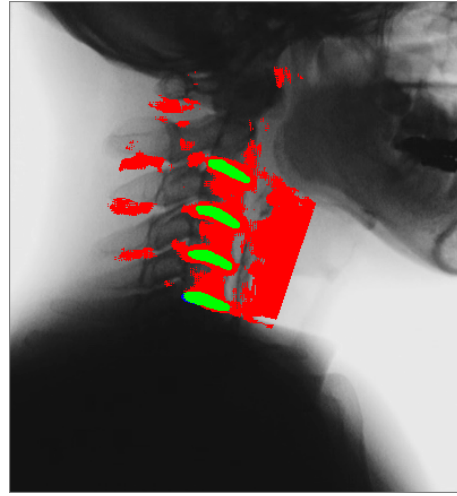


図 4.160: (GRY, GRY, GRY) に AlexNet を適用して得られた結果画像 (F 値=0.214)

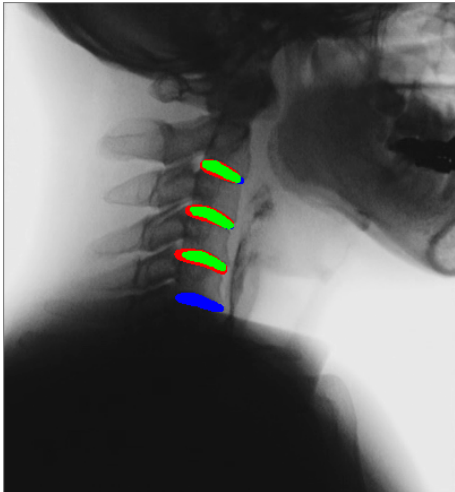


図 4.161: (GRY, GRY, GRY) に U-Net を適用して得られた結果画像 (F 値 =0.680)

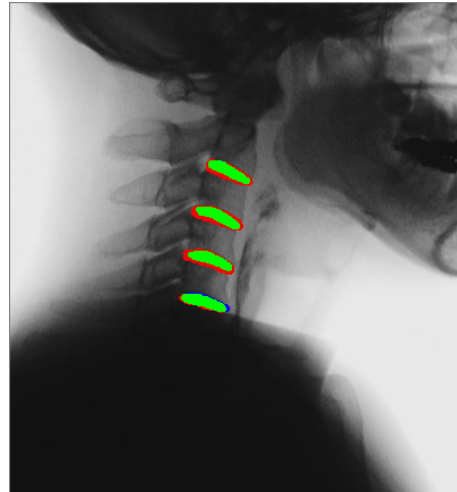


図 4.162: (GRY, GRY, GRY) に FPN を適用して得られた結果画像 (F 値 =0.791)

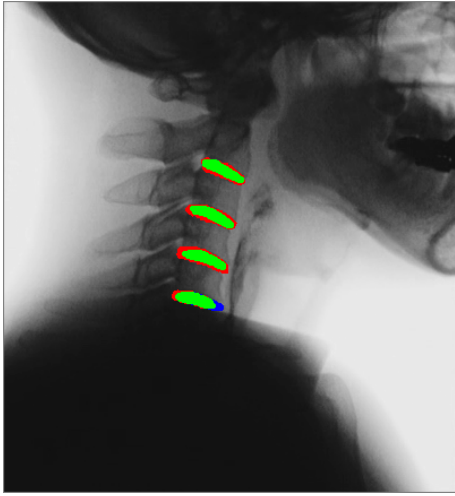


図 4.163: (GRY, GRY, GRY) に LinkNet を適用して得られた結果画像 (F 値=0.800)

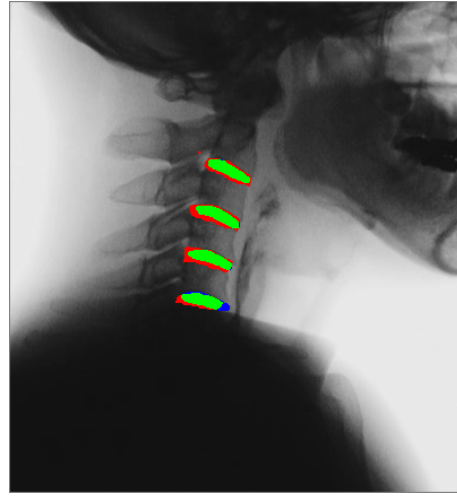


図 4.164: (GRY, GRY, GRY) に PSP-Net を適用して得られた結果画像 (F 値=0.765)

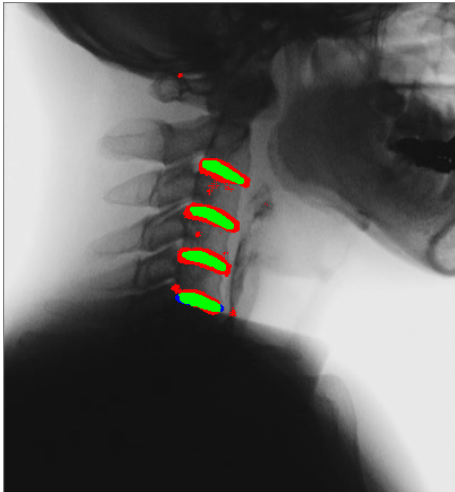


図 4.165: (SBLX2, SBLY1, POS3) に LeNet を適用して得られた結果画像 (F 値=0.657)

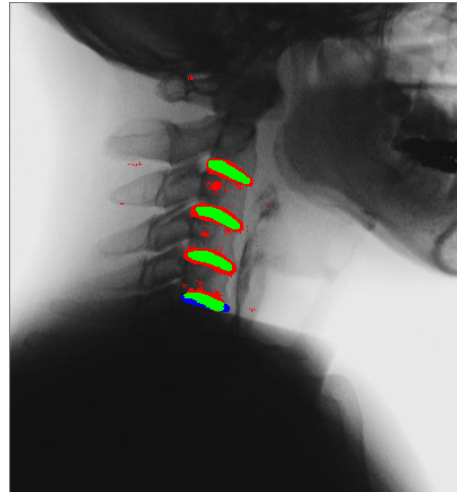


図 4.166: (SBLY1, SBLX2, SBLY3) に AlexNet を適用して得られた結果画像 (F 値=0.666)

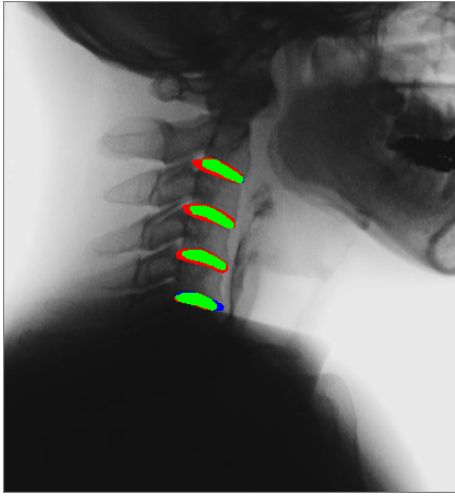


図 4.167: (HGE, HGE, HGE) に U-Net を適用して得られた結果画像 (F 値=0.792)

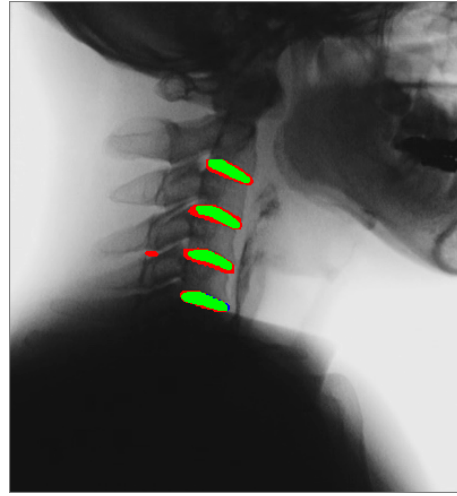


図 4.168: (EMB3, EMB3, GAU1) に FPN を適用して得られた結果画像 (F 値=0.782)

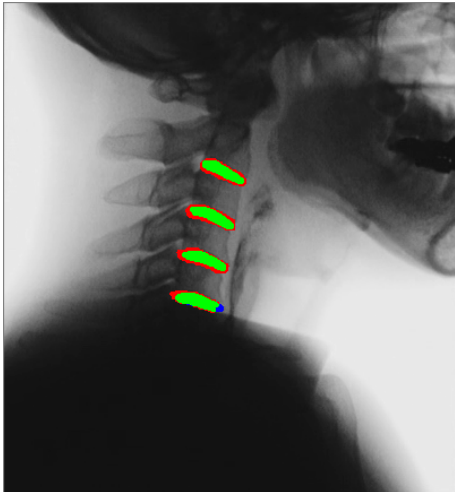


図 4.169: (EMB3, ETC1, BLT1) に LinkNet を適用して得られた結果画像 (F 値=0.782)

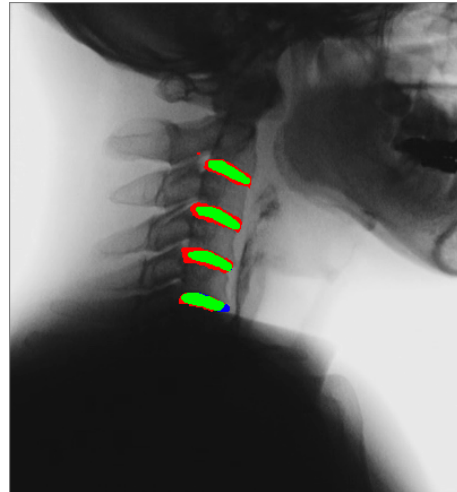


図 4.170: (ETC2, MEA2, MED2) に PSPNet を適用して得られた結果画像 (F 値=0.766)

表 4.24: 4.2.3.6 節で LeNet を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(SBLY1, SBLY1, SBLY1)	1	0.632
(SBLY3, SBLY3, SBLY3)	2	0.622
(SBLD, SBLD, SBLD)	3	0.621
(EMB2, EMB2, EMB2)	4	0.604
(POS3, POS3, POS3)	5	0.594
(SBL2, SBL2, SBL2)	6	0.594
(LBP1, LBP1, LBP1)	7	0.592
(ETC2, ETC2, ETC2)	8	0.591
(HGE, HGE, HGE)	9	0.590
(POS2, POS2, POS2)	10	0.589

ジュールを  $T = T \times 0.9$ ,  $step$  の初期値を  $StepMax = 23$  とし, 遷移が採択された回数が 100 回を超えた場合に焼きなましを終了した. 理論上, 遷移が採択される回数が 100 回目のとき,  $T = 0.0148$  となり, エネルギー差  $e - e'$  が  $-0.5, -0.4, -0.3, -0.2, -0.1, -0.05, -0.01$  となった場合に, 遷移確率  $p$  が順に  $1.93 \times 10_{-15}, 1.69 \times 10_{-12}, 1.48 \times 10_{-9}, 1.30 \times 10_{-6}, 1.14 \times 10_{-3}, 3.38 \times 10_{-2}, 0.508$  となる試算となった. また, 4.2.3.6 節では後処理として判別分析法による二値化を行ったが, 本節では画素値 128 をしきい値とする二値化を行った. ただし, 4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い, 4.2.3.6 節と同一の学習, テスト症例, チャンネル数, 深層学習パラメータ, 計算機, フレームワークを用い,  $prob$  を 0.5 とし, 前処理としてパッチ毎の画素値正規化補正, 回転補正, データ拡張を行わなかった.

#### 4.2.3.7.2 実験結果

3-fold 3-phase cross validation の実験 1 に焼きなまし法と U-Net, FPN, LinkNet, PSPNet をそれぞれ適用して得られた上位 10 種類の特徴画像の組合せを表 4.37, 表 4.38, 表 4.39, 表 4.40 に示す. 4 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せをそれぞれ入力して得られた平均 F 値と, 3 グループの最適な特徴画像の組合せをまとめて表 4.41 に示す. U-Net, FPN, LinkNet では, 最適な特徴画像の組合せを入力して得られた平均 F 値が (GRY, GRY, GRY) を入力して得られた平均 F 値と比較して有意に大きかった. 4 つの深層学習モデルに図 4.153 の最適な特徴画像の組合せをそれぞれ適用して得られた結果画像と同一症例の F 値を図 4.171, 図 4.172, 図 4.173, 図 4.174 に示す. 3-fold 3-phase cross validation の実験 1 に焼きなまし法を適用した際, 遷移が採択された回数は 4 つの深層学習モデル全てで 101 回であり, 遷移が採択されなかった回数は順に 7 回, 30 回, 8 回, 5 回であった.

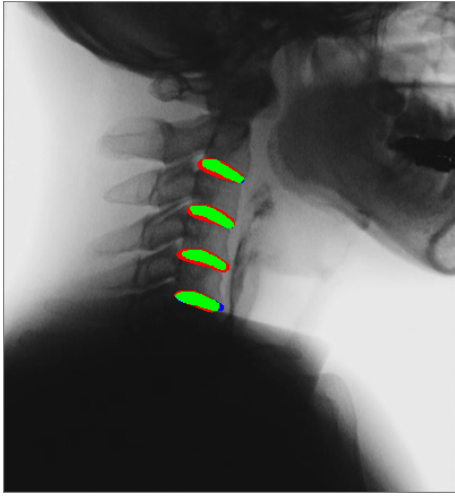


図 4.171: (EMB1, SBLD, SBLY1) に U-Net を適用して得られた結果画像 (F 値=0.808)

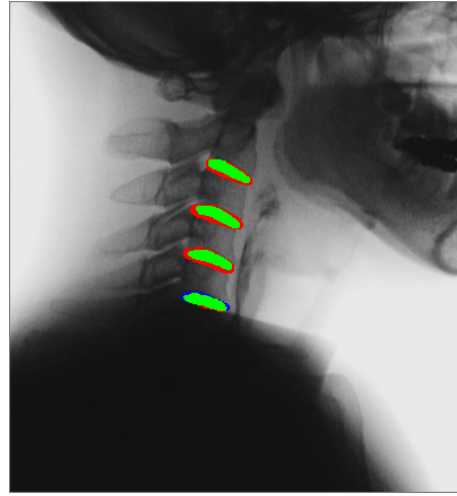


図 4.172: (ETC2, NML1, MED2) に FPN を適用して得られた結果画像 (F 値=0.783)

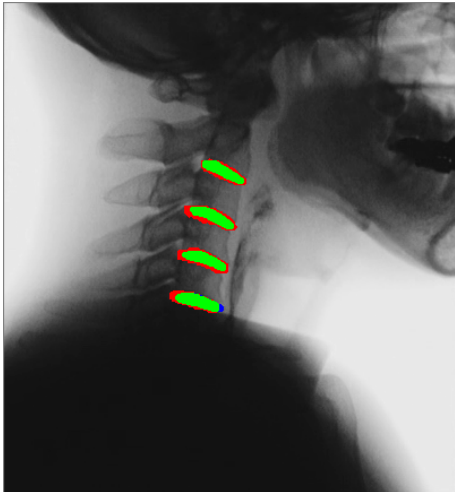


図 4.173: (BLT1, BLT1, BLT1) に LinkNet を適用して得られた結果画像 (F 値=0.787)

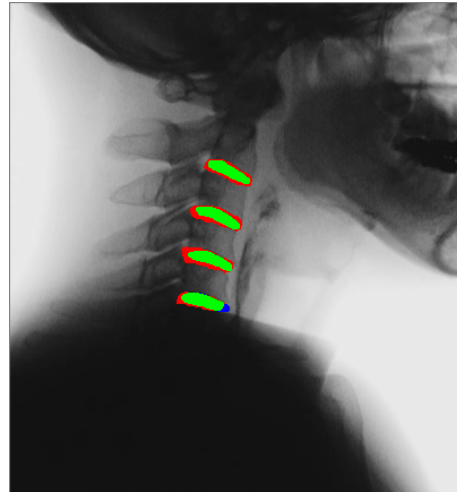


図 4.174: (NML3, NML3, NML3) に PSPNet を適用して得られた結果画像 (F 値=0.769)

表 4.25: 4.2.3.6 節で AlexNet を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(MEA1, MEA1, MEA1)	1	0.656
(GRY, GRY, GRY)	2	0.634
(HGE, HGE, HGE)	3	0.611
(POS2, POS2, POS2)	4	0.598
(SBLY2, SBLY2, SBLY2)	5	0.589
(EMB2, EMB2, EMB2)	6	0.589
(NML1, NML1, NML1)	7	0.588
(SBLD, SBLD, SBLD)	8	0.585
(NML2, NML2, NML2)	9	0.580
(SBLY1, SBLY1, SBLY1)	10	0.573

#### 4.2.3.8 アウトライアを除外した実験

##### 4.2.3.8.1 実験条件

4.2.3.7 節では、症例ごとに F 値が大きく異なっていた。この問題を解決するため、目視で椎間板と非椎間板の境界を明確に判別できない症例を除外して、U-Net とマルチチャンネル化、特徴画像の選択に基づく椎間板領域の抽出を行った。除外した症例は 2 症例で、どちらも患者群の症例であった。これらの 2 症例を図 4.175 と図 4.176 に示す。2 症例を除外したところ、グループ A の症例数は患者群 6 症例と正常群 13 症例を合わせた 19 症例、グループ B の症例数は患者群 5 症例と正常群 13 症例を合わせた 18 症例、グループ C の症例数は患者群 6 症例と正常群 13 症例を合わせた 19 症例となったため、既存の症例のグループを変更することなく、3-fold 3-phase cross validation を適用した。焼きなましの初期温度を  $T = 10000$ 、冷却スケジュールを  $T = T \times 0.9$ 、終了条件を  $T \leq 1$ 、 $stepMax$  を 23 とした。ただし、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、4.2.3.6 節と同一のチャンネル数、深層学習パラメータ、計算機、フレームワークを用い、 $prob$  を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わず、4.2.3.7 節と同一の後処理を行った。

##### 4.2.3.8.2 実験結果

3-fold 3-phase cross validation の実験 1 に焼きなまし法と U-Net を適用して得られた上位 10 種類の特徴画像の組合せを表 4.42 に示す。4.2.3.7 節と本節のそれぞれで (GRY, GRY, GRY) と最適な特徴画像の組合せを U-Net に入力して得られた平均 F 値をまとめて表 4.43 に示す。4.2.3.7 節と本節のそれぞれで、3-fold 3-phase cross validation の実験 2 に U-Net を適用して得られた症例ベースの F 値を表 4.44 に示す。4.2.3.7 節と本節で焼きなまし条件が異なるものの、アウトライアを除外しても症例間の F 値の差異は改善されず、症例によっては F 値が下がってしまった。理由として、学習データの枚数が 19 枚と少なく、一枚あたりの影響が大きいこ

表 4.26: 4.2.3.6 節で U-Net を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の順位	平均 F 値
(SBLY2, SBLY2, SBLY2)	1	0.774
(NML1, NML1, NML1)	2	0.773
(SBLY3, SBLY3, SBLY3)	3	0.771
(MED2, MED2, MED2)	4	0.763
(SBLY1, SBLY1, SBLY1)	5	0.756
(NGP, NGP, NGP)	6	0.755
(TOP1, TOP1, TOP1)	7	0.754
(LBP3, LBP3, LBP3)	8	0.751
(EMB3, EMB3, EMB3)	9	0.750
(GAU1, GAU1, GAU1)	10	0.750

とが挙げられる。本節で 3-fold 3-phase cross validation の実験 1 に焼きなまし法を適用した際、遷移が採択された回数は 88 回であり、遷移が採択されなかった回数は 0 回であった。

#### 4.2.3.9 21 種類の特徴画像に焼きなまし法と PSPNet を適用した実験

##### 4.2.3.9.1 実験条件

4.2.3.4 節と 4.2.3.5 節で特徴画像の枚数を変更した比較実験を行ったが、焼きなまし条件が異なっており、枚数の純粋な比較はできなかった。本節では、4.2.3.7 節から特徴画像の枚数のみを変更し、PSPNet とマルチチャンネル化、特徴画像の選択に基づく椎間板領域の抽出を行った。4.2.3.7 節では 45 種類の特徴画像を用いたのに対し、本節では、GRY, HGE, SOL, ZSP, MEA1, GAU1, MED1, KNN1, BLT1, LPL1, ETC1, STC1, NML1, LBP1, POS1, EMB1, SBLX1, SBLY1, SBLM, SBLD, TOP1 の合計 21 種類の特徴画像を用いた。4.2.3.4 節と本節で特徴画像の枚数は同じであるが、NML, LBP, EMB, TOP の番号が異なっている。本節では、特徴画像の枚数の変更に伴い、*StepMax* を 11 とした。ただし、4.2.3.7 節と同一の焼きなましの初期温度、冷却スケジュール、終了条件を用い、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、4.2.3.6 節と同一の学習、テスト症例、チャンネル数、深層学習パラメータ、計算機、フレームワークを用い、*prob* を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わなかった。

##### 4.2.3.9.2 実験結果

3-fold 3-phase cross validation の実験 1 に PSPNet を適用して得られた上位 10 種類の同一の特徴画像を表 4.45 に、実験 1 に焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せを表 4.46 に示す。4.2.3.7 節と本節のそれぞれで (GRY, GRY, GRY)

表 4.27: 4.2.3.6 節で FPN を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の順位	平均 F 値
(EMB2, EMB2, EMB2)	1	0.815
(POS3, POS3, POS3)	2	0.802
(NML1, NML1, NML1)	3	0.800
(BLT2, BLT2, BLT2)	4	0.794
(ETC1, ETC1, ETC1)	5	0.786
(HGE, HGE, HGE)	6	0.785
(KNN2, KNN2, KNN2)	7	0.783
(GAU1, GAU1, GAU1)	8	0.778
(MED1, MED1, MED1)	9	0.773
(MED2, MED2, MED2)	10	0.771

と最適な特徴画像の組合せを PSPNet に入力して得られた平均 F 値をまとめて表 4.47 に示す。4.2.3.7 節と本節の比較から、特徴画像を 21 種類に限定した方が 45 種類を用いるより F 値の上昇幅が大きくなった。これは 4.2.3.4 節と 4.2.3.5 節の比較結果と合致する。

#### 4.2.3.10 深層学習の層数を変更した実験

##### 4.2.3.10.1 実験条件

これまでの実験で U-Net を用いた際には、エンコーダの層数を 5 層としていた。本節では、エンコーダの層数を 2 層, 3 層, 4 層, 5 層と変更し、U-Net とマルチチャンネル化、特徴画像の選択に基づく椎間板領域の抽出を行った。4.2.3.7 節と同一の焼きなまし条件を用い、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、4.2.3.6 節と同一の学習、テスト症例、チャンネル数、深層学習パラメータ、計算機、フレームワークを用い、 $prob$  を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わなかった。

##### 4.2.3.10.2 実験結果

3-fold 3-phase cross validation の実験 1 に 2 層, 3 層, 4 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像を順に表 4.48, 表 4.49, 表 4.50 に、実験 1 に焼きなまし法と 2 層, 3 層, 4 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せを順に表 4.51, 表 4.52, 表 4.53 に示す。5 層の U-Net を用いた実験結果として、表 4.26 と表 4.37 の結果を再利用した。表 4.54 に、2 層, 3 層, 4 層, 5 層の U-Net を用いて得られた MC なし, ありの F 値と最適な特徴画像の組合せを示す。表 4.54 より、U-Net の層数を変更してもマルチチャンネル化の効果は認められた。また、3 層の MC なしの F 値より 2 層の MC ありの F 値が有意に大きく、4 層の MC なしの F 値より 3 層の MC ありの F 値が有意に大きく、5 層の MC なし



表 4.28: 4.2.3.6 節で LinkNet を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の順位	平均 F 値
(GAU1, GAU1, GAU1)	1	0.820
(HGE, HGE, HGE)	2	0.811
(EMB2, EMB2, EMB2)	3	0.811
(MED1, MED1, MED1)	4	0.800
(BLT1, BLT1, BLT1)	5	0.799
(NGP, NGP, NGP)	6	0.798
(EMB3, EMB3, EMB3)	7	0.797
(MEA2, MEA2, MEA2)	8	0.797
(KNN1, KNN1, KNN1)	9	0.794
(GRY, GRY, GRY)	10	0.793

の F 値より 4 層の MC ありの F 値が有意に大きく、5 層の MC なしの F 値より 3 層の MC ありの F 値が有意に大きく、5 層の MC なしの F 値より 2 層の MC ありの F 値が有意に大きかった。このことから、U-Net の層数を増やすよりも MC を導入した方が効果が大きいことがわかった。なお、3 層の MC ありの実験で 0.777 という F 値を得られたことから、使用した VF データセットに対しては、エンコーダの層数が 3 層で充分であったと考えられる。

#### 4.2.3.11 チャンネル数を変更した実験

##### 4.2.3.11.1 実験条件

これまでの実験ではチャンネル数を 3 としていた。これは、深層学習のフレームワークがカラー画像あるいはグレースケール画像を入力する仕組みになっており、特徴画像を 3 種類に限定することによって、それらを大きな改造を施すことなく利用することができるからである。さらに、転移学習に用いられる画像は 3 チャンネルのカラー画像であることが多いので、その利用も容易になるという利点もある。ただし、実装上の拡張によりチャンネル数を変更することも可能である。本節では、チャンネル数を 1, 2, 3, 4 と変更し、U-Net とマルチチャンネル化、特徴画像の選択に基づく椎間板領域の抽出を行った。4.2.3.7 節と同一の焼きなまし条件を用い、4.1.3.7 節で作成した椎間板領域のデータと頸部マスク C を用い、4.2.3.6 節と同一の学習、テスト症例、深層学習パラメータ、計算機、フレームワークを用い、U-Net のエンコーダの層数を 5 層とし、*prob* を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わなかった。

表 4.29: 4.2.3.6 節で PSPNet を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の順位	平均 F 値
(HGE, HGE, HGE)	1	0.787
(EMB2, EMB2, EMB2)	2	0.778
(NML3, NML3, NML3)	3	0.775
(MED2, MED2, MED2)	4	0.772
(KNN1, KNN1, KNN1)	5	0.771
(NGP, NGP, NGP)	6	0.771
(ETC1, ETC1, ETC1)	7	0.771
(MEA1, MEA1, MEA1)	8	0.771
(ETC2, ETC2, ETC2)	9	0.771
(KNN2, KNN2, KNN2)	10	0.770

#### 4.2.3.11.2 実験結果

3-fold 3-phase cross validation の実験 1 で得られた上位 10 種類の同一の特徴画像を 1 チャンネル, 2 チャンネル, 4 チャンネルの順に表 4.55, 表 4.56, 表 4.57 に, 実験 1 に焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せを 2 チャンネル, 4 チャンネルの順に表 4.58, 表 4.59 に示す. 3 チャンネルの実験結果として, 表 4.26 と表 4.37 の結果を再利用した. 表 4.60 に, 1 チャンネル, 2 チャンネル, 3 チャンネル, 4 チャンネルそれぞれの MC なし, ありの F 値と最適な特徴画像の組合せを示す. 表 4.60 より, チャンネル数を変更してもマルチチャンネル化の効果は認められた. 計算時間は順に約 4.7 日, 約 9.2 日, 約 13.9 日, 約 16.2 日となり, 生成された特徴画像の組合せ数は順に 45 通り, 104 通り, 135 通り, 153 通りとなった.

#### 4.2.3.12 焼きなましの開始位置を変更した実験

##### 4.2.3.12.1 実験条件

これまでの実験では, 焼きなまし法の初期状態を  $(i_1, i_2, \dots, i_n, \dots, i_N) = (1, 1, \dots, 1, \dots, 1)$  としていたが, 最適解が初期状態に依存するか検討する必要がある. 本節では, 3 チャンネルの画像で 45 種類の特徴画像の順位付けを行った後, 上位 28 種類の特徴画像を選択し, (1 位, 1 位, 1 位) と (14 位, 14 位, 14 位) のそれぞれから焼きなましを開始して特徴選択を行うことで, 得られた最適な特徴画像の組合せと F 値を比較した. 本節では, 4.2.3.6 節で定義した 3-fold 3-phase cross validation のグループ A を学習に, グループ B を特徴選択に用いて特徴選択のみを行い, テストは行わなかった. 焼きなましの初期温度を  $T = 500$ , 冷却スケジュールを  $T = T \times 0.95$ ,  $step$  の初期値を  $StepMax = 14$  とし, 遷移が採択された回数が 200 回を超えた場合に焼きなましを終了した. 4.1.3.7 節で作成した椎間板領域のデータと頸部マスク

表 4.30: 4.2.3.6 節で焼きなまし法と LeNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(NGP, EMB3, TOP4)	(26, 15, 27)	0.641
(SBLM, ETC1, TOP4)	(23, 18, 27)	0.640
(MED2, TOP4, SBLM)	(20, 27, 23)	0.639
(ETC1, NML1, TOP4)	(18, 19, 27)	0.636
(POS1, EMB3, TOP4)	(25, 15, 27)	0.633
(SBLY1, SBLY1, SBLY1)	(1, 1, 1)	0.632
(TOP2, POS1, SBLM)	(16, 25, 23)	0.628
(SBLM, TOP2, TOP4)	(23, 16, 27)	0.628
(SBLD, TOP1, GAU1)	(3, 33, 30)	0.627
(POS2, MEA2, SBLX3)	(10, 21, 29)	0.624

C を用い、 $prob$  を 0.5 とし、前処理としてパッチ毎の画素値正規化補正、回転補正、データ拡張を行わなかった。keras をフレームワークに用い、エンコーダの層数を 5 層とした U-Net を用い、エポック数を 100、バッチサイズを 1 とし、平均二乗誤差を損失関数に、確率的勾配降下法を最適化関数に用いた。

#### 4.2.3.12.2 実験結果

上位 10 種類の同一の特徴画像を表 4.61 に、焼きなまし法を (1 位, 1 位, 1 位), (14 位, 14 位, 14 位) から開始して得られた上位 10 種類の最適な特徴画像の組合せを順に表 4.62, 表 4.63 に示す。焼きなましの開始位置を変更することにより、ほぼ同じ F 値ではあるが異なる最適解を得た。これは、焼きなまし回数に上限を設けていることに起因すると考えられる。

#### 4.2.4 考察

マルチチャンネル化と焼きなまし法に基づく特徴画像の選択を導入することにより、8 ビットの医用画像をそのまま深層学習モデルに入力する手法と比較し、構造物の抽出精度を向上させることができた。マルチチャンネル化で用いた画像フィルタは線形フィルタと非線形フィルタに分けられる。最適な 3 種類の特徴画像の組合せの内、その多くで非線形フィルタが含まれる傾向にあった。従来の深層学習の内部でも畳み込み層に線形フィルタが、プーリング層と活性化関数に非線形関数が用いられている。マルチチャンネル化で用いた線形フィルタは畳み込み層で代用できる可能性があるが、非線形フィルタはプーリング層と活性化関数で代用できないと考える。すなわち、プーリング層の適用後には中間画像が小さくなるため、厳密な意味での非線形フィルタではない。また、活性化関数も ReLU やハイパボリックタンジェ

表 4.31: 4.2.3.6 節で焼きなまし法と AlexNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(MEA1, MEA1, MEA1)	(1, 1, 1)	0.656
(MEA1, NML3, MED2)	(1, 25, 27)	0.646
(HGE, TOP2, EMB3)	(3, 23, 21)	0.621
(HGE, MEA2, MEA2)	(3, 12, 12)	0.618
(MEA1, TOP3, GAU2)	(1, 24, 28)	0.616
(HGE, LBP3, TOP2)	(3, 23, 20)	0.615
(POS2, NML1, GAU2)	(4, 7, 28)	0.614
(HGE, LBP3, TOP2)	(3, 20, 23)	0.611
(MEA1, TOP4, NML3)	(1, 16, 25)	0.608
(SBLY3, GAU1, EMB3)	(11, 13, 21)	0.598

ントなどの数種類に限定されるが、マルチチャンネル化ではさらに多くの非線形フィルタを用いることができる。

以上の考察を補強するため、線形フィルタと非線形フィルタを深層学習の内部で近似できるか否かを検討する実験を行った。まず、58 症例の VF を学習症例 40 症例と検証症例 9 症例とテスト症例 9 症例に分割した。学習症例と検証症例に画像フィルタを適用して得られた画像を教師データとし、U-Net で学習、テストを行った。テスト症例に U-Net を適用して得られた近似画像と、テスト症例に画像フィルタを直接適用して得られた Ground truth との間の差分画像を求めた。評価指標として、近似画像と Ground truth との二乗平均平方根誤差 (root mean square error, RMSE) を求めた。画像フィルタとして、線形フィルタである EMB2 と、非線形フィルタである TOP2, LBP1, SBLD をそれぞれ用いた。keras をフレームワークに用い、U-Net のエポック数を 100, バッチサイズを 1 とし、損失関数として平均二乗誤差を、最適化関数として確率的勾配降下法を用いた。正常症例の VF フレーム画像を図 4.177 に、エンボスフィルタ、Top-hat 変換フィルタ、LBP フィルタ、勾配の Sobel フィルタそれぞれの近似画像、Ground truth, 差分画像を図 4.178 から図 4.189 に示す。各症例にそれぞれの画像フィルタを適用して得られた RMSE を表 4.64 に示す。全体的な傾向として、近似画像では、Ground truth と比較し、隣接する画素間の画素値の変化が滑らかなものとなった。この現象は椎間板領域と非椎間板領域の境界でも確認された。また、エンボスフィルタの近似画像では、画像の右下に本来であれば存在しないはずの縦長の楕円が生成された。U-Net を用いても線形フィルタや非線形フィルタを完全に近似できないことが示された。また、実験前には線形フィルタより非線形フィルタの方が RMSE が大きくなると予想したが、必ずしもそうではないことが示された。本章では椎間板領域の抽出を目的としたが、画像内の面積が小さい椎間板を正確に抽出するためには、椎体との境界が重要となるため、マルチチャンネル化で高周波成分を

表 4.32: 4.2.3.6 節で焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLY1, GAU1, GAU2)	(5, 10, 13)	0.794
(SBLY1, TOP1, SBLX2)	(5, 7, 15)	0.789
(SBLY2, GAU1, MEA1)	(1, 10, 16)	0.788
(SBLY2, SBLX3, TOP2)	(1, 11, 12)	0.785
(SBLY3, EMB3, SBLX2)	(3, 9, 15)	0.785
(MED2, SBLY1, EMB2)	(4, 5, 19)	0.785
(SBLY2, MED2, SBLY3)	(1, 4, 3)	0.785
(NGP, SBLY1, LBP3)	(6, 5, 8)	0.785
(SBLY3, SBLY2, TOP2)	(3, 1, 12)	0.784
(SBLY3, MED2, TOP3)	(3, 4, 20)	0.782

強調することが有効であると考えられる。

従来の深層学習の内部でも損失関数を最小化するために確率的勾配降下法などの最適化手法を用いており、その過程で誤差逆伝搬法を用いて重みや畳み込み層のカーネルを更新する。深層学習の学習段階でソフトマックス層の出力結果に応じて畳み込み層のカーネルを更新することは広義の特徴選択に相当する。しかしながら、前処理段階でマルチチャンネル化を用いて特徴選択を行うことで従来の深層学習手法をブーストすることができた。マルチチャンネル化による特徴選択は従来の深層学習の最適化を拡張したものであるとも言えるが、最適化の手法が異なる。深層学習の最適化では損失関数を最小化するように連続的にカーネルの成分を更新するのに対し、マルチチャンネル化による特徴選択では画像フィルタの各画素値は変更せず離散的に画像フィルタの種類を更新する。マルチチャンネル化による特徴選択により、従来の連続的な更新では成し得なかった大きな更新ができる可能性がある。

本節では、特徴画像の組合せの探索手法として焼きなまし法を用いた。組合せ最適化の代替手法として山上り法と全探索法が挙げられる。山上り法をマルチチャンネル化に適用する場合、遷移後の F 値が遷移前より高くなれば遷移を採択し、そうでなければ遷移前の組合せにとどまる。山上り法と焼きなまし法の効果を比較するために、チャンネル数を 3 とする以下の簡易実験を行った。45 種類の特徴画像を F 値の降順にソートし、3 種類の特徴画像のインデックスを 3 次元の軸とし、(1 位, 1 位, 1 位) とその周囲 26 近傍の (1 位, 1 位, 2 位), (1 位, 2 位, 2 位), (2 位, 2 位, 2 位), (1 位, 1 位, 3 位), (1 位, 2 位, 3 位), (1 位, 3 位, 3 位), (2 位, 2 位, 3 位), (2 位, 3 位, 3 位), (3 位, 3 位, 3 位) の特徴画像をそれぞれ 3 チャンネルに保存して U-Net を適用し、F 値を求めた。ただし、各チャンネルに保存する特徴画像の順序が異なっても、組合せが同じであれば同一の F 値を再利用した。特徴画像の組合せ以外全ての実験条件を 4.2.3.12 節と統一した。表 4.65 に (1 位, 1 位, 1 位) とその周囲 26 近傍の特徴画像の組合



図 4.175: 除外した患者症例 1

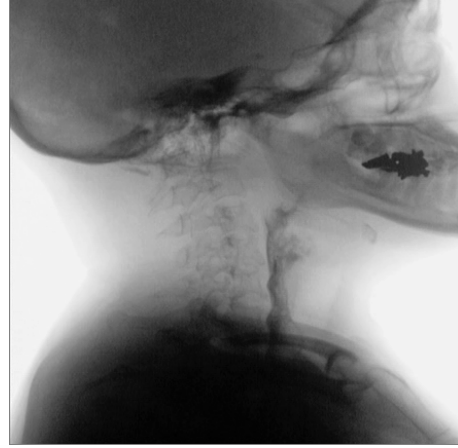


図 4.176: 除外した患者症例 2



図 4.177: 正常症例の VF フレーム画像

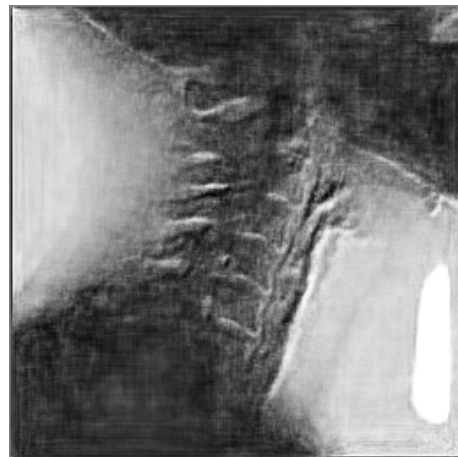


図 4.178: U-Net を適用して得られた EMB2 の近似画像



図 4.179: EMB2 の Ground truth

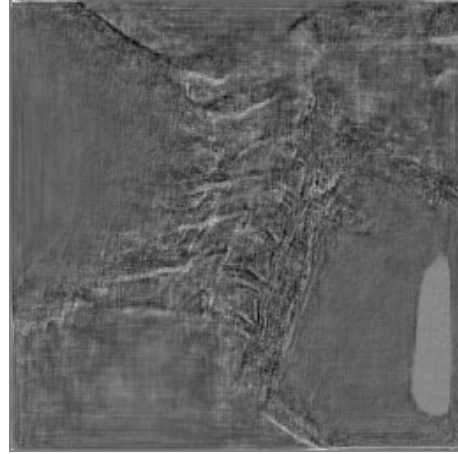


図 4.180: EMB2 の差分画像



図 4.181: U-Net を適用して得られた TOP2 の近似画像



図 4.182: TOP2 の Ground truth

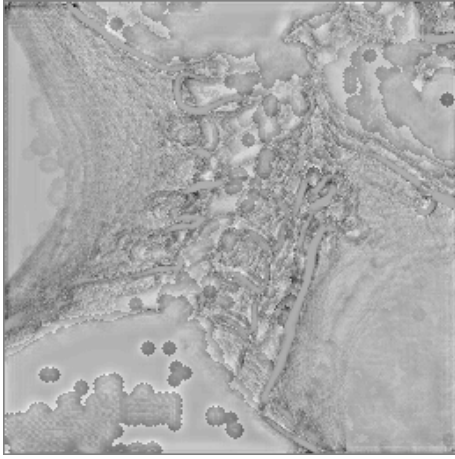


図 4.183: TOP2 の差分画像



図 4.184: U-Net を適用して得られた  
LBP1 の近似画像

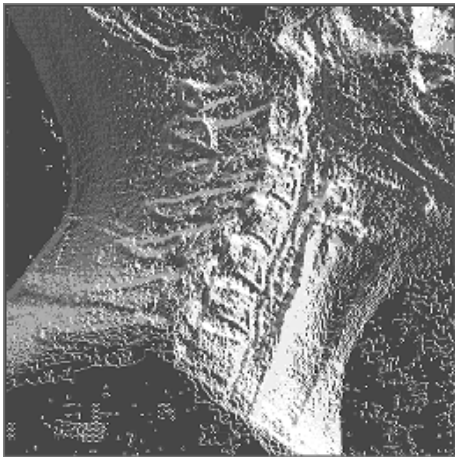


図 4.185: LBP1 の Ground truth

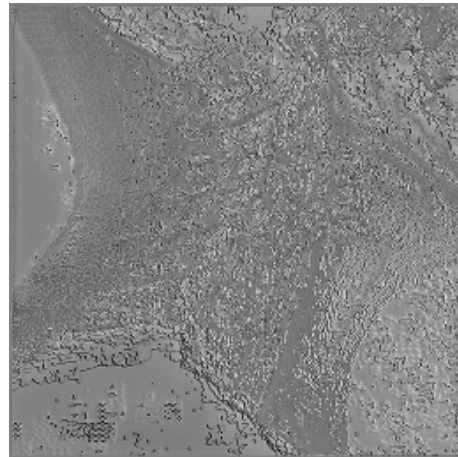


図 4.186: LBP1 の差分画像





図 4.187: U-Net を適用して得られた SBLD の近似画像

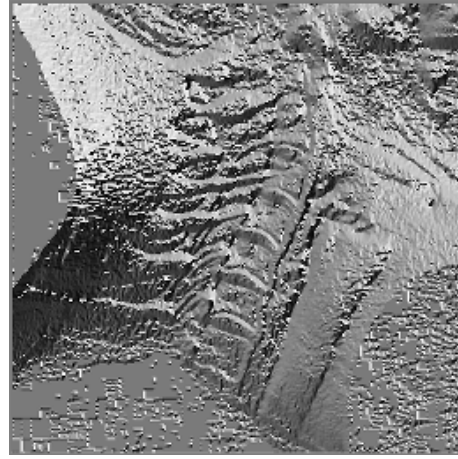


図 4.188: SBLD の Ground truth

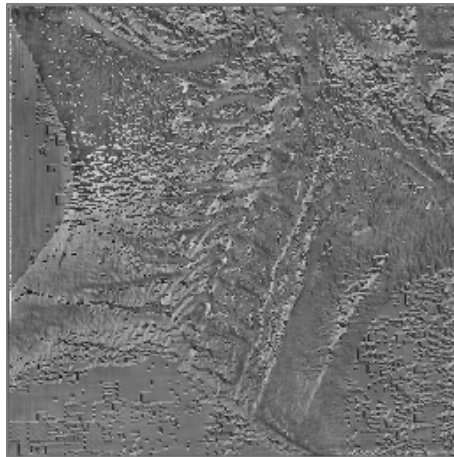


図 4.189: SBLD の差分画像

表 4.33: 4.2.3.6 節で焼きなまし法と FPN を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(NML2, ETC2, EMB1)	(13, 30, 11)	0.816
(EMB2, EMB2, EMB2)	(1, 1, 1)	0.815
(GRY, GAU2, EMB1)	(16, 23, 11)	0.813
(BLT1, GRY, BLT2)	(15, 16, 4)	0.807
(KNN1, MEA1, KNN2)	(14, 12, 7)	0.803
(TOP1, BLT1, POS3)	(20, 15, 2)	0.801
(NML2, EMB1, MEA1)	(13, 11, 12)	0.801
(MED1, GAU1, EMB1)	(9, 8, 11)	0.797
(HGE, HGE, ETC1)	(6, 6, 5)	0.797
(POS2, GAU2, POS3)	(18, 23, 2)	0.796

せを示す。表 4.62, 表 4.65 より, (1 位, 1 位, 1 位) から *step* が 1 の山上市法を用いて探索を開始した場合, (1 位, 2 位, 2 位) という局所最適解にとどまり, F 値は 0.821 となるが, 焼きなまし法で求めた最適解は (SBLY3, EMB3, KNN1) で, それぞれの順位は (2 位, 4 位, 25 位) であり, F 値は 0.830 となった。したがって, 単純な山上市法よりも焼きなまし法の方がより良い組合せを探索できることがわかった。さらに, 全探索法をマルチチャンネル化に適用する場合の組合せ数を試算する。チャンネル数を 3 とし, 各チャンネルに保存する特徴画像の順序が異なっても組合せが同じであれば同一のエネルギー値を再利用する場合, 特徴画像の種類を 1 種類, 2 種類, 3 種類, 4 種類, 5 種類, 6 種類, 7 種類, 8 種類, 9 種類, 10 種類, …としたとき, 組合せ数は順に 1 通り, 4 通り, 10 通り, 20 通り, 35 通り, 76 通り, 119 通り, 176 通り, 249 通り, 340 通り, …となる。4.2.3.7 節では, 焼きなましの遷移の採択回数を 101 回としたため, 単純計算で 7 種類以上の特徴画像を用いる際には全探索法より焼きなまし法の方が効率が良いことになる。

本節のテストでは, グループ A, B, C を合わせた合計 58 症例の平均 F 値により椎間板の抽出精度を評価している。グループ A, B, C で最適な特徴画像の組合せが異なるものの, 最適な特徴画像の組合せを入力に用いた際の F 値が (GRY, GRY, GRY) を入力に用いた際の F 値より大きい値となった。本節の実験を実際の医療現場に応用する際, グループ A, B, C で得られたいずれの最適な特徴画像の組合せを採用することもでき, ある程度の椎間板抽出精度向上が担保されている。

本節ではマルチチャンネル化の効果のみを検証するために後処理を行わなかった。後処理を行えば抽出精度の向上は期待できるが, 処理の選定や組み合わせがアドホックになる。

本節では 21 種類あるいは 45 種類の特徴画像を用いたが, アルゴリズムとしては他の特徴画像を組み込むことや, 特徴画像の種類を減らすことが可能である。最適な特徴画像の枚数

表 4.34: 4.2.3.6 節で焼きなまし法と LinkNet を適用して得られた上位 10 種類の特最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(POS1, POS1, EMB2)	(34, 34, 3)	0.821
(GAU1, GAU1, GAU1)	(1, 1, 1)	0.820
(NML3, STC2, EMB2)	(24, 39, 3)	0.817
(POS1, MEA1, EMB2)	(34, 28, 3)	0.814
(GAU1, KNN1, HGE)	(1, 9, 2)	0.813
(KNN1, NML2, MEA2)	(9, 21, 8)	0.813
(KNN1, KNN2, KNN1)	(9, 18, 9)	0.812
(KNN1, NML1, EMB3)	(9, 15, 7)	0.809
(EMB3, POS3, MED1)	(7, 22, 4)	0.809
(KNN1, MEA1, EMB2)	(9, 28, 3)	0.807

はデータセットに依存すると予想されるため、21 種類あるいは 45 種類という枚数に拘らず、最適な特徴画像の枚数を探ることが望ましい。

本節では LeNet, AlexNet, U-Net, FPN, LinkNet, PSPNet という 6 つの深層学習モデルを用いたが、アルゴリズムとしては他の識別器を組み込むこともできる。現在、様々な機械学習モデルが登場し、精度向上を目指して日々改良が加えられている。医用画像や生物画像を対象とした機械学習モデルの例を挙げる。Liu ら [87] は、二方向から撮影したマンモグラムから乳房腫瘍の検出を行うために、bipartite graph convolutional network (BGN) を基にした機械学習モデルを提案した。Yang ら [88] は、3次元画像である心血管造影 CT 画像から複数の冠動脈を分離するために、座標情報と画素値情報の両方を組み込んだ conditional partial-residual graph convolutional network (CPR-GCN) を提案した。Kulikov ら [89] は、大腸菌画像の生物画像から重なり合う物体を分離して抽出するために、U-Net の各 upscaling block の最初の畳み込み層を SinConv layer で置き換えた深層学習モデルを提案した。Hashimoto ら [90] は、病理画像のアノテーションコスト、拡大倍率や染色条件の違いに対応するために、multiple instance learning (MIL), domain adversarial (DA), multi-scale (MS) learning を組み合わせたパッチベースの深層学習モデルを提案した。Maksoud ら [91] は、高解像度の病理画像に対してパッチベースの深層学習を適用すると計算資源を多く消費するため、原画像を低解像度に変換してパッチベースでスクリーニングを行い、所見が不明な位置のみを高解像度で精査する selective objective switch (SOS) を提案した。Dong ら [92] は、上部消化管内視鏡画像と下部消化管内視鏡画像の両方を用いて領域抽出を行うために、マルチモダルな教師なし転移学習モデルを提案した。Meng ら [93] は、様々な網膜疾患がある中で、新たな網膜疾患の眼底写真が学習データに加わるごとにスクラッチから学習を行わなければならないという問題に対応するために、教師データの特徴分布を保持し、予測時のラベルと属性を増分学習フレームワークで統合する ADINet

表 4.35: 4.2.3.6 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(MEA2, EMB3, KNN1)	(24, 19, 5)	0.799
(HGE, HGE, HGE)	(1, 1, 1)	0.787
(MEA2, POS3, EMB2)	(24, 15, 2)	0.785
(MEA2, BLT1, KNN2)	(24, 17, 10)	0.784
(EMB1, KNN2, ETC2)	(20, 10, 9)	0.784
(GAU2, MED1, HGE)	(23, 14, 1)	0.783
(HGE, ETC2, EMB2)	(1, 9, 2)	0.780
(POS2, BLT2, ETC2)	(21, 12, 9)	0.778
(MEA2, POS3, NML3)	(24, 15, 3)	0.777
(POS2, BLT1, HGE)	(21, 17, 1)	0.776

表 4.36: 4.2.3.6 節で 6 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せを入力して得られた平均 F 値

深層学習 モデル		MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
LeNet	P	0.604	<b>0.613</b>	(NGP, EMB3, TOP4)	(SBLX2, SBLY1, POS3)	(ETC1, SBLX1, SBLY2)
AlexNet	P	0.459	<b>0.602**</b>	(MEA1, MEA1, MEA1)	(SBLY1, SBLX2, SBLY3)	(SBLY2, TOP1, GAU2)
U-Net	S	0.660	<b>0.776**</b>	(SBLY1, GAU1, GAU2)	(HGE, HGE, HGE)	(LBP2, SBLY2, SBLX3)
FPN	S	0.752	<b>0.789*</b>	(NML2, ETC2, EMB1)	(EMB3, EMB3, GAU1)	(GRY, KNN1, ETC1)
LinkNet	S	0.803	<b>0.808</b>	(POS1, POS1, EMB2)	(EMB3, ETC1, BLT1)	(BLT1, BLT1, BLT1)
PSPNet	S	0.750	<b>0.768</b>	(MEA2, EMB3, KNN1)	(ETC2, MEA2, MED2)	(SBLX1, GAU1, LBP3)

(Attribute driven incremental network) を提案した。Qihang Yu ら [94] は、3次元医用画像から臓器や腫瘍を抽出するために、ネットワークをマクロレベルのトポロジとマイクロレベルの cell operation に分けて最適なネットワーク構造を探索する coarse-to-fine neural architecture search (C2FNAS) を提案した。Zhao ら [95] は、大腸がんの病理画像からリンパ節転移の有無を予測するために、GCN, multiinstance learning (MIL) と特徴量選択を組み込んだ深層学習モデルを提案した。Lee ら [96] は、医用画像内の構造物の境界があいまいな場合に抽出精度を向上させるため、boundary preserving block (BPB) による形状の学習と shape boundary-aware evaluator (SBE) による整合性の評価を組み込んだ深層学習モデルを提案した。Hanchao Yu ら [97] は、心臓 MR 画像を用いて動作推定を行う際に、学習データセットとテストデータセットの分布が異なる場合にも精度を保つために、メタ学習を組み込んだ fast online adaptive learning (FOAL) を提案した。さらに、機械学習モデルを自動的に設計する automatic machine learning (AutoML) [98] という手法が存在する。AutoML は、ハイパーパラメータの最適化、メタ学習、NAS な

表 4.37: 4.2.3.7 節で焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(TOP2, ETC2, SBLY1)	(12, 39, 5)	0.793
(LBP2, ETC2, SBLY1)	(14, 39, 5)	0.790
(LBP2, POS1, SBLY1)	(14, 43, 5)	0.789
(SBLY1, LPL1, LBP3)	(5, 41, 8)	0.786
(SBLM, LPL2, SBLX3)	(21, 31, 11)	0.783
(SBLX3, POS1, SBLY2)	(11, 43, 1)	0.781
(TOP2, SBLX1, MED2)	(12, 44, 4)	0.780
(TOP2, LPL1, SBLY1)	(12, 41, 5)	0.777
(SBLX3, MED1, SBLY1)	(11, 40, 5)	0.775
(SBLY2, SBLY2, SBLY2)	(1, 1, 1)	0.774

どの機能をもつ。識別器に AutoML を用いれば、計算時間は膨大になるものの、精度がより向上すると考えられる。勿論、SVM やランダムフォレスト、ロジスティック回帰などの古典的な機械学習を用いることもできる。本研究で SVM ではなく深層学習を使用した理由として、以下の二つを挙げる。一つ目は、深層学習と SVM の分類精度の大小は問題によって異なるが [99][100]、SVM で可能な抽出手法はパッチベースにとどまり、本節で、パッチベースと比較して semantic segmentation の抽出精度が高い傾向にあることを示したからである。二つ目は、上述の通り、SVM で可能な抽出手法はパッチベースにとどまるため、学習やテストに計算時間を費やすからである。

マルチチャンネル化の課題は、深層学習と特徴画像の選択を組み合わせしており、一回の実験に日単位あるいは週単位の計算時間を費やすことである。深層学習のテストの結果に依存して焼きなましの次の state への遷移を決定するため、複数の state の計算を並列に行うことはできない。計算時間の多くの割合を深層学習の学習、テストが占めるため、特徴画像の選択に費やすことができる時間が限られているなら、計算時間が短い識別器を用いて最適な特徴画像の組合せを選択し、それを U-Net などの深層学習に応用する手法が考えられる。ただし、最適な特徴画像の組合せは識別器ごとに異なるため、十分な計算時間を確保できれば、テストで使用する識別器を用いて特徴画像の選択を行うことが理想的である。また、計算時間の短縮のためには、使用する特徴画像の枚数を減らし、焼きなましで探索する特徴空間を狭くすることも効果的であると考えられる。本節では、Sobel フィルタ、top-hat 変換フィルタ、エンボスフィルタなどが最適な特徴画像の組合せ内に多く出現したため、それらの特徴画像を優先的に探索する手法が考えられる。あるいは、特徴画像間の類似度を計算し、類似度と画素単位の F 値に応じて特徴画像の選択を行う手法も考えられる。なお、計算時間のほぼ全てを特徴画像の順位付けと最適な特徴画像の組合せの選択に費やしており、テストに費やす時間は

表 4.38: 4.2.3.7 節で焼きなまし法と FPN を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(EMB2, EMB2, EMB2)	(1, 1, 1)	0.815
(EMB3, SBLY2, TOP4)	(21, 32, 27)	0.796
(EMB3, SBLX3, SBLY3)	(21, 33, 29)	0.789
(TOP1, SBLX1, SBLX3)	(20, 34, 33)	0.777
(GAU2, STC1, SBLY3)	(23, 31, 29)	0.774
(NGP, SBLX2, POS1)	(28, 25, 26)	0.773
(POS1, POS1, ETC2)	(26, 26, 30)	0.772
(GAU2, SBLY1, SBLY2)	(23, 37, 32)	0.770
(TOP1, STC2, SBLY2)	(20, 35, 32)	0.770
(SBLM, STC2, SBLY3)	(19, 35, 29)	0.766

数時間である。マルチチャンネル化を臨床医療現場に応用する際には、病院外部で事前に特徴画像の順位付けと最適な特徴画像の組合せの選択を済ませ、病院では一つの組合せの学習、テストのみを行うため、数時間で椎間板領域を抽出できる。

本論文の限界として、研究室で使用可能な VF データセットのみを用いたことが挙げられる。使用した VF データセットは 58 症例と少数であり、Ground truth の作成方法が椎間板抽出精度に大きく影響を与える。本論文では少数データセット問題の解決手法としてマルチチャンネル化による特徴画像の選択を用いたが、他にも様々な解決手法が存在する。Perez ら [101] は、tiny-imagenet-200 という少数の自然画像からクラス分類を行う際に、アフィン変換を中心としたデータ拡張、Generative adversarial networks (GAN)、入力画像のペアを同時に深層学習に入力する augmentation network を別個に用いた。Rajpurkar ら [102] は、少数の CT 画像から虫垂炎の識別を行うために、Kinetics と呼ばれる YouTube の大規模なビデオデータセットを用いて事前学習を行い、CT 画像を用いて転移学習を行う手法を提案した。Hagos ら [103] は、少数の眼底写真から糖尿病性網膜症の識別を行うために、ImageNet を用いて Inception-V3 モデルで事前学習を行い、眼底写真を用いて転移学習を行う手法を提案した。Liu ら [104] は、微生物が密集した少数の顕微鏡画像から重なり合う微生物を分離して検出するために、HOG-SVM detector を用いて微生物の候補領域を検出した後、大域的最適化を用いて偽陽性を削減する手法を提案した。Kataoka ら [105] は、実際の自然画像を必要としない事前学習モデルとして、数学的規則に基づいたフラクタルを生成し、画像パターンとカテゴリラベルを自動的に付与する手法を提案した。Shi ら [106] は、少数の超音波画像から乳房腫瘍や前立腺腫瘍を分類するために、異なる層の特徴量を連結して蓄積する stacked deep polynomial network (S-DPN) という深層学習モデルを提案した。本節で示したマルチチャンネル化の効果が他の医用画像や自然画像のデータセットでも確認できるかは現時点では不明であり、今後の研究に期待する。

表 4.39: 4.2.3.7 節で焼きなまし法と LinkNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(GAU1, GAU1, GAU1)	(1, 1, 1)	0.820
(SBLX1, STC2, SBL1)	(36, 39, 23)	0.801
(SBLX1, STC2, TOP1)	(36, 39, 27)	0.793
(POS1, LBP1, TOP1)	(34, 40, 27)	0.792
(SBLX1, SBLD, TOP4)	(36, 37, 29)	0.791
(SBLX3, SBLX1, TOP1)	(35, 36, 27)	0.789
(TOP3, LBP1, TOP3)	(32, 40, 32)	0.787
(SBLX3, STC2, SBL1)	(35, 39, 23)	0.786
(SBLX3, SOL, TOP1)	(35, 38, 27)	0.785
(SBLX3, SBLD, TOP1)	(35, 37, 27)	0.784

本研究の目的は、嚙下障害と頸椎の形動態の関連性を調べることであり、3 章では VF を用いた椎間板の形態解析手法を、本章では深層学習を用いた椎間板領域の自動抽出手法を提案した。今後、対象物を椎間板から頸椎椎体や頸椎棘突起へ拡大し、大規模な VF データセットを入手して本論文の提案手法を適用し、自動抽出した頸椎領域の形態解析を行う手法を開発し、嚙下障害と頸椎の形動態の関連性を評価することによって、最終的な目的を達成することが可能となる。なお、深層学習に改良を加えることで、椎間板領域の自動抽出と形動態解析による嚙下障害の識別を一つのネットワークで完結できると考える。また、X 線画像と VF では造影剤の有無やフレーム数、画像解像度などが異なるが、撮影方法は基本的に同じであることと、造影剤を用いても椎間板には影響がないことから、椎間板の抽出が目的であれば X 線画像を用いてもよい。しかしながら、今後、嚙下障害と頸椎の形動態の関連性を調べるにあたり、嚙下障害の有無や程度、頸椎の動態を調べるために VF が必須である。詳細な形動態解析を実現するために、嚙下障害の有無や頸椎疾患の診断名が明確な大規模な VF データセットを入手することが必要となる。医療機関での試料や情報の取り扱いと他機関への提供は、人を対象とする医学系研究に関する倫理指針 [107] に基づく。診療録、レセプト、健康診断の結果、ゲノム情報などの要配慮個人情報も新たに取得して研究を実施する場合には、原則研究対象者から同意を求める必要がある。自機関の既存情報のみを用いた学術研究では、個人情報であってもオプトアウトの手続きで研究を実施できる。他機関に既存試料や情報を提供する場合には、学術研究などの明確な目的があり、研究対象者に研究目的を通知し、研究対象者個人を直ちに判別できないように匿名化することで、機関を跨いだ研究を実施できる。医用画像処理研究は、法律や指針を遵守し、研究対象者の協力を得てはじめて可能となる。医用画像処理研究の展望として、現在盛んに行われている臓器の正確な抽出や医用画像診断支援システムの構築とともに、工学的な手法を取り入れた臓器の評価指標の作成と、それを用い

表 4.40: 4.2.3.7 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(HGE, HGE, HGE)	(1, 1, 1)	0.787
(GAU2, BLT1, EMB2)	(23, 17, 2)	0.782
(NML1, POS1, EMB2)	(13, 33, 2)	0.777
(POS2, GAU2, EMB3)	(21, 23, 19)	0.776
(GAU1, POS1, EMB2)	(11, 33, 2)	0.774
(KNN2, GRY, EMB3)	(10, 22, 19)	0.773
(NML1, POS1, HGE)	(13, 33, 1)	0.768
(MED1, POS1, EMB2)	(14, 33, 2)	0.765
(NML2, POS1, HGE)	(16, 33, 1)	0.765
(NML1, TOP2, EMB2)	(13, 32, 2)	0.764

表 4.41: 4.2.3.7 節で 6 つの深層学習モデルに (GRY, GRY, GRY) と最適な特徴画像の組合せを入力して得られた平均 F 値

深層学習 モデル	MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
U-Net	0.660	<b>0.747**</b>	(TOP2, ETC2, SBLY1)	(EMB1, SBLD, SBLY1)	(LBP3, SBLX2, LBP3)
FPN	0.752	<b>0.794*</b>	(EMB2, EMB2, EMB2)	(ETC2, NML1, MED2)	(NML2, NML2, NML2)
LinkNet	0.803	<b>0.813*</b>	(GAU1, GAU1, GAU1)	(BLT1, BLT1, BLT1)	(BLT1, BLT1, BLT1)
PSPNet	0.750	<b>0.765</b>	(HGE, HGE, HGE)	(NML3, NML3, NML3)	(POS2, SBLX1, BLT1)

た病因解明や治療法の確立を行うことで、研究対象者に研究成果を還元できると考える。



表 4.42: アウトライアを除外し, 焼きなまし法と U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(TOP1, SBLY1, GAU1)	(9, 1, 29)	0.835
(TOP1, SBLY2, POS2)	(9, 2, 31)	0.831
(TOP1, SBLX2, TOP4)	(9, 5, 26)	0.831
(STC2, TOP4, SBLY1)	(38, 26, 1)	0.829
(GAU2, SBLY1, LPL2)	(11, 1, 27)	0.827
(EMB3, SBLY2, POS2)	(10, 2, 31)	0.825
(KNN1, SBLY2, TOP3)	(8, 2, 30)	0.824
(SBLY3, BLT1, MEA1)	(12, 4, 28)	0.824
(SBLY3, SBLY1, EMB2)	(12, 1, 23)	0.823
(SBLY3, SBLY2, BLT2)	(12, 2, 24)	0.822

表 4.43: 4.2.3.7 節と 4.2.3.8 節の (GRY, GRY, GRY) と最適な特徴画像の組合せの比較

アウトライア 除外の有無	MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
なし (4.2.3.7 節)	0.660	<b>0.747**</b>	(TOP2, ETC2, SBLY1)	(EMB1, SBLD, SBLY1)	(LBP3, SBLX2, LBP3)
あり (4.2.3.8 節)	<b>0.690</b>	0.661	(TOP1, SBLY1, GAU1)	(KNN2, POS3, SBLY1)	(SBLY3, SBLY3, SBLY3)

表 4.44: 4.2.3.7 節と 4.2.3.8 節の症例ベースの比較

症例番号	4.2.3.7 節 (GRY, GRY, GRY)	4.2.3.7 節 (EMB1, SBLD, SBLY1)	4.2.3.8 節 (GRY, GRY, GRY)	4.2.3.8 節 (KNN2, POS3, SBLY1)
患者群 P16	0.405	0.880	0.351	0.0527
患者群 P17	0.477	0.490	0.368	0.147
患者群 P21	0.180	0.638	0.543	0.000
患者群 P4	0.833	0.789	0.855	0.391
患者群 P19	0.877	0.866	0.860	0.725
患者群 P26	0.563	0.540	0.566	0.000
正常群 N26	0.847	0.836	0.804	0.508
正常群 N24	0.737	0.872	0.741	0.626
正常群 N4	0.680	0.792	0.682	0.846
正常群 N23	0.805	0.872	0.772	0.000
正常群 N21	0.365	0.807	0.259	0.437
正常群 N10	0.606	0.852	0.669	0.639
正常群 N5	0.121	0.722	0.223	0.000998
正常群 N29	0.771	0.827	0.830	0.770
正常群 N22	0.827	0.867	0.836	0.496
正常群 N18	0.930	0.909	0.901	0.809
正常群 N12	0.583	0.870	0.643	0.604
正常群 N28	0.494	0.748	0.489	0.323
正常群 N38	0.484	0.845	0.631	0.632
平均値	0.610	0.791	0.633	0.421

表 4.45: 4.2.3.9 節で PSPNet を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(HGE, HGE, HGE)	1	0.787
(KNN1, KNN1, KNN1)	2	0.771
(ETC1, ETC1, ETC1)	3	0.771
(MEA1, MEA1, MEA1)	4	0.771
(GAU1, GAU1, GAU1)	5	0.770
(NML1, NML1, NML1)	6	0.768
(MED1, MED1, MED1)	7	0.766
(BLT1, BLT1, BLT1)	8	0.759
(EMB1, EMB1, EMB1)	9	0.757
(GRY, GRY, GRY)	10	0.752

表 4.46: 4.2.3.9 節で焼きなまし法と PSPNet を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(TOP1, NML1, NML1)	(13, 6, 6)	0.789
(HGE, HGE, HGE)	(1, 1, 1)	0.787
(TOP1, MED1, NML1)	(13, 7, 6)	0.782
(TOP1, POS1, BLT1)	(13, 15, 8)	0.781
(TOP1, POS1, BLT1)	(13, 15, 8)	0.781
(SBLD, SBLY1, BLT1)	(12, 11, 8)	0.778
(ZSP, KNN1, NML1)	(21, 2, 6)	0.776
(TOP1, EMB1, MED1)	(13, 9, 7)	0.774
(SBLY1, SBLY1, MED1)	(11, 11, 7)	0.773
(ZSP, HGE, GAU1)	(21, 1, 5)	0.771

表 4.47: 4.2.3.7 節と 4.2.3.9 節の (GRY, GRY, GRY) と最適な特徴画像の組合せの比較

特徴画像の 枚数	MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
45 種類 (4.2.3.7 節)	0.750	<b>0.765</b>	(HGE, HGE, HGE)	(NML3, NML3, NML3)	(POS2, SBLX1, BLT1)
21 種類 (4.2.3.9 節)	0.750	<b>0.771</b>	(TOP1, NML1, NML1)	(ZSP, BLT1, BLT1)	(SBLX1, SBLX1, GRY)

表 4.48: 2 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(SBLY1, SBLY1, SBLY1)	1	0.712
(LBP2, LBP2, LBP2)	2	0.652
(SBLY3, SBLY3, SBLY3)	3	0.650
(EMB3, EMB3, EMB3)	4	0.648
(NML2, NML2, NML2)	5	0.641
(SBLX2, SBLX2, SBLX2)	6	0.641
(SBLX3, SBLX3, SBLX3)	7	0.640
(MEA1, MEA1, MEA1)	8	0.637
(LBP3, LBP3, LBP3)	9	0.614
(HGE, HGE, HGE)	10	0.607

表 4.49: 3 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(EMB3, EMB3, EMB3)	1	0.762
(BLT2, BLT2, BLT2)	2	0.744
(SBLX1, SBLX1, SBLX1)	3	0.741
(SBLX2, SBLX2, SBLX2)	4	0.741
(SBLD, SBLD, SBLD)	5	0.740
(MED2, MED2, MED2)	6	0.739
(SBLY2, SBLY2, SBLY2)	7	0.737
(NGP, NGP, NGP)	8	0.736
(SBLX3, SBLX3, SBLX3)	9	0.735
(SBLY3, SBLY3, SBLY3)	10	0.723

表 4.50: 4 層の U-Net を適用して得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(TOP1, TOP1, TOP1)	1	0.789
(SBLY2, SBLY2, SBLY2)	2	0.775
(TOP2, TOP2, TOP2)	3	0.775
(SBLY3, SBLY3, SBLY3)	4	0.771
(NML2, NML2, NML2)	5	0.766
(SBLX3, SBLX3, SBLX3)	6	0.766
(MED1, MED1, MED1)	7	0.757
(SBLX1, SBLX1, SBLX1)	8	0.755
(GAU1, GAU1, GAU1)	9	0.750
(SBLY1, SBLY1, SBLY1)	10	0.749

表 4.51: 焼きなまし法と 2 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(TOP4, SBLY2, SBLY3)	(42, 31, 3)	0.764
(LPL1, TOP1, SBLY3)	(40, 35, 3)	0.750
(ETC2, SBLY2, SBLX3)	(44, 31, 7)	0.747
(ETC2, SBLY2, SBLX2)	(44, 31, 6)	0.742
(GAU2, SBLY2, SBLX1)	(30, 31, 13)	0.734
(TOP2, SBLY2, SBLX1)	(28, 31, 13)	0.724
(ETC2, SBLY2, SBLY3)	(44, 31, 3)	0.720
(POS1, POS2, SBLY3)	(41, 36, 3)	0.718
(ETC1, LPL2, SBLY1)	(38, 26, 1)	0.716
(SBLY1, SBLY1, SBLY1)	(1, 1, 1)	0.712

表 4.52: 焼きなまし法と 3 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(LPL2, LBP3, SBLY2)	(27, 14, 7)	0.782
(TOP4, KNN2, SBLY2)	(22, 35, 7)	0.780
(TOP2, POS2, EMB3)	(23, 40, 1)	0.768
(EMB3, EMB3, EMB3)	(1, 1, 1)	0.762
(NML2, NML1, EMB3)	(29, 42, 1)	0.757
(TOP3, POS2, EMB3)	(28, 40, 1)	0.756
(TOP1, MEA1, SBLX2)	(21, 41, 4)	0.756
(SBLM, SOL, SBLX1)	(24, 38, 3)	0.756
(POS3, GRY, EMB3)	(25, 39, 1)	0.753
(TOP4, STC1, SBLX2)	(22, 37, 4)	0.752

表 4.53: 焼きなまし法と 4 層の U-Net を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLX3, SBLD, TOP3)	(6, 27, 17)	0.795
(SBLY3, SBLD, BLT2)	(4, 27, 18)	0.790
(TOP1, TOP1, TOP1)	(1, 1, 1)	0.789
(NML2, SBLD, KNN2)	(5, 27, 20)	0.786
(SBLX3, SBLD, NGP)	(6, 27, 19)	0.785
(KNN1, EMB2, TOP4)	(12, 31, 16)	0.785
(MED1, SBLD, TOP3)	(7, 27, 17)	0.785
(SBLX3, SBLD, KNN2)	(6, 27, 20)	0.782
(SBLX1, NML3, TOP3)	(8, 33, 17)	0.780
(POS3, EMB2, BLT2)	(11, 31, 18)	0.776

表 4.54: 2 層, 3 層, 4 層, 5 層の U-Net を用いて得られた (GRY, GRY, GRY) と最適な特徴画像の組合せの比較

U-Net の 層数	MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
2	0.380	<b>0.717**</b>	(TOP4, SBLY2, SBLY3)	(SBLY2, SBLX2, TOP3)	(SBLY2, TOP1, EMB2)
3	0.636	<b>0.777**</b>	(LPL2, LBP3, SBLY2)	(SBLX1, SBLY1, LPL1)	(SBLY2, SBLY2, SBLY2)
4	0.720	<b>0.768**</b>	(SBLX3, SBLD, TOP3)	(SBLY3, SBLX2, ETC1)	(NGP, SBLD, LBP3)
5	0.660	<b>0.747**</b>	(TOP2, ETC2, SBLY1)	(EMB1, SBLD, SBLY1)	(LBP3, SBLX2, LBP3)

表 4.55: 1 チャンネルの実験で得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(SBLY2)	1	0.781
(SBLY3)	2	0.768
(KNN1)	3	0.765
(SBLX2)	4	0.760
(TOP1)	5	0.759
(GAU2)	6	0.757
(SBLY1)	7	0.755
(SBLX1)	8	0.751
(EMB3)	9	0.746
(GRY)	10	0.728

表 4.56: 2 チャンネルの実験で得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(SBLY2, SBLY2)	1	0.772
(SBLY1, SBLY1)	2	0.759
(SBLX2, SBLX2)	3	0.759
(NML3, NML3)	4	0.753
(EMB3, EMB3)	5	0.750
(SBLD, SBLD)	6	0.749
(SBLX3, SBLX3)	7	0.741
(TOP3, TOP3)	8	0.740
(POS2, POS2)	9	0.727
(TOP2, TOP2)	10	0.727

表 4.57: 4 チャンネルの実験で得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(SBLY2, SBLY2, SBLY2, SBLY2)	1	0.777
(TOP2, TOP2, TOP2, TOP2)	2	0.774
(SBLY3, SBLY3, SBLY3, SBLY3)	3	0.772
(TOP1, TOP1, TOP1, TOP1)	4	0.764
(EMB1, EMB1, EMB1, EMB1)	5	0.760
(SBLX3, SBLX3, SBLX3, SBLX3)	6	0.757
(EMB2, EMB2, EMB2, EMB2)	7	0.755
(SBLX1, SBLX1, SBLX1, SBLX1)	8	0.749
(SBLX2, SBLX2, SBLX2, SBLX2)	9	0.746
(KNN2, KNN2, KNN2, KNN2)	10	0.744

表 4.58: 2 チャンネルの実験で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLD, SBLY2)	(6, 1)	0.796
(SBLX1, SBLY1)	(13, 2)	0.794
(TOP1, SBLY2)	(14, 1)	0.789
(TOP1, SBLY1)	(14, 2)	0.783
(SBLY3, SBLX3)	(22, 7)	0.780
(SBLD, SBLX2)	(6, 3)	0.774
(SBLY2, SBLY2)	(1, 1)	0.772
(LBP3, SBLD)	(11, 6)	0.772
(SBLD, MEA2)	(6, 23)	0.772
(EMB3, SBLY1)	(5, 2)	0.770

表 4.59: 4 チャンネルの実験で焼きなまし法を適用して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(LPL1, NML1, SBLX3, SBLD)	(29, 13, 6, 25)	0.803
(SBLY1, MEA2, KNN2, SBLD)	(31, 17, 10, 25)	0.793
(SBLY1, GAU1, EMB2, TOP4)	(31, 14, 7, 28)	0.788
(SBLY1, GAU1, ETC1, EMB3)	(31, 14, 12, 23)	0.786
(SBLY1, BLT1, EMB1, SBLD)	(31, 11, 5, 25)	0.785
(BLT2, GAU1, SBLX3, LBP3)	(33, 14, 6, 26)	0.785
(GAU2, NML1, EMB2, SBLD)	(30, 13, 7, 25)	0.782
(SBLY1, MEA1, KNN2, KNN1)	(31, 16, 10, 27)	0.781
(SBLX2, MEA2, SBLX2, SBLY3)	(9, 17, 9, 3)	0.778
(NML2, NML1, KNN2, SBLD)	(32, 13, 10, 25)	0.778

表 4.60: チャンネル数を 1, 2, 3, 4 と変更して得られた (GRY, GRY, GRY) と最適な特徴画像の組合せの比較

チャンネル数	MC なし 平均 F 値	MC あり 平均 F 値	実験 1 の 最適な特徴画像	実験 2 の 最適な特徴画像	実験 3 の 最適な特徴画像
1	0.668	<b>0.720*</b>	(SBLY2)	(HGE)	(SBLX2)
2	0.600	<b>0.754**</b>	(SBLD, SBLY2)	(SBLY1, ETC1)	(SBLY3, SBLX2)
3	0.660	<b>0.747**</b>	(TOP2, ETC2, SBLY1)	(EMB1, SBLD, SBLY1)	(LBP3, SBLX2, LBP3)
4	0.687	<b>0.785**</b>	(LPL1, NML1, SBLX3, SBLD)	(HGE, HGE, HGE, HGE)	(SBLX3, BLT2, GAU1, SBLY3)



表 4.61: 4.2.3.12 節で得られた上位 10 種類の同一の特徴画像

特徴画像の組合せ	特徴画像の 順位	平均 F 値
(TOP2, TOP2, TOP2)	1	0.808
(SBLY3, SBLY3, SBLY3)	2	0.794
(NML1, NML1, NML1)	3	0.793
(EMB3, EMB3, EMB3)	4	0.791
(SBLX1, SBLX1, SBLX1)	5	0.784
(BLT2, BLT2, BLT2)	6	0.781
(MEA1, MEA1, MEA1)	7	0.778
(NML3, NML3, NML3)	8	0.776
(SBLY1, SBLY1, SBLY1)	9	0.776
(POS3, POS3, POS3)	10	0.774

表 4.62: 焼きなまし法を (1 位, 1 位, 1 位) から開始して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLY3, EMB3, KNN1)	(2, 4, 25)	0.830
(TOP2, BLT2, EMB2)	(1, 6, 20)	0.829
(TOP2, SBLY3, GRY)	(1, 2, 17)	0.827
(POS3, EMB3, GAU1)	(10, 4, 15)	0.820
(TOP2, SBLX1, LBP3)	(1, 5, 19)	0.819
(NML1, EMB3, GAU2)	(3, 4, 21)	0.819
(EMB3, NML1, EMB1)	(3, 4, 23)	0.819
(SBLY3, TOP2, GAU2)	(2, 1, 21)	0.818
(SBLY1, TOP2, EMB2)	(9, 1, 20)	0.816
(SBLX2, EMB3, LBP1)	(13, 4, 16)	0.815

表 4.63: 焼きなまし法を (14 位, 14 位, 14 位) から開始して得られた上位 10 種類の最適な特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(SBLX1, EMB2, GRY)	(5, 20, 17)	0.830
(TOP2, EMB1, EMB2)	(1, 23, 20)	0.820
(EMB3, LBP1, BLT1)	(4, 16, 18)	0.819
(BLT2, MEA1, BLT1)	(6, 7, 18)	0.818
(SBLY1, LBP1, SBL1)	(9, 16, 13)	0.815
(SBLX2, BLT2, STC1)	(13, 6, 26)	0.815
(LBP3, NML2, NML2)	(19, 12, 12)	0.814
(SBLX2, EMB3, EMB1)	(13, 4, 23)	0.812
(SBLX2, MEA1, SBLY2)	(13, 7, 24)	0.812
(NML1, LBP1, GAU1)	(3, 16, 15)	0.808

表 4.64: EMB2, TOP2, LBP1, SBLD を近似して得られた各症例の RMSE

症例番号	EMB2 RMSE	TOP2 RMSE	LBP1 RMSE	SBLD RMSE
患者群 P20	358	160	1580	2910
患者群 P21	499	298	1660	3070
患者群 P22	466	203	1700	2910
正常群 N27	268	110	1360	2750
正常群 N28	247	118	1400	3230
正常群 N29	319	155	1440	2910
正常群 N30	326	142	1420	3110
正常群 N31	398	128	1290	2870
正常群 N32	642	175	1310	3200
平均値	391	165	1460	3000

表 4.65: (1 位, 1 位, 1 位) とその周囲 26 近傍の特徴画像の組合せ

特徴画像の組合せ	特徴画像の 順位の組合せ	平均 F 値
(TOP2, TOP2, TOP2)	(1, 1, 1)	0.778
(TOP2, TOP2, SBLY3)	(1, 1, 2)	0.8206
(TOP2, SBLY3, SBLY3)	(1, 2, 2)	0.8209
(SBLY3, SBLY3, SBLY3)	(2, 2, 2)	0.716
(TOP2, TOP2, NML1)	(1, 1, 3)	0.770
(TOP2, SBLY3, NML1)	(1, 2, 3)	0.773
(TOP2, NML1, NML1)	(1, 3, 3)	0.741
(SBLY3, SBLY3, NML1)	(2, 2, 3)	0.803
(SBLY3, NML1, NML1)	(2, 3, 3)	0.781
(NML1, NML1, NML1)	(3, 3, 3)	0.748

## 第5章 結論

本論文では、医師がVFから手動抽出した椎間板領域の形態解析と、深層学習を用いた椎間板領域の抽出を行った。深層学習を用いた椎間板領域の抽出手法として、モルフォロジー演算による後処理を用いた方法と、マルチチャンネル化による前処理を用いた方法を提案した。患者群19名と正常群39名から構成されるVFデータセットに提案手法を適用した。医師がVFから手動抽出した椎間板領域の形態解析では、患者群の椎間板の形状が正常群と比較して不整であることが示された。深層学習とモルフォロジー演算による後処理を用いた椎間板領域の抽出では、学習、テスト症例の分割、椎間板領域のデータ、頸部マスク、パッチサイズ、パッチの画素値正規化補正と回転補正、深層学習の尤度、後処理を変更して、それぞれの条件で椎間板抽出精度が最大となるパラメータを模索した。深層学習とマルチチャンネル化による前処理を用いた椎間板領域の抽出では、学習、テスト症例の分割、特徴画像の種類、深層学習モデル、深層学習パラメータ、焼きなまし条件、計算機環境に伴う乱数固定の有無を変更して、マルチチャンネル化が椎間板抽出精度の向上に与える効果を多角的に検証した。本研究の目的は、嚙下障害と頸椎の形動態の関連性を調べることであり、そのためには嚙下障害の有無が明確な大規模なVFデータセットから椎間板領域を自動抽出し、形態解析を行う必要がある。本論文では、VFを用いて椎間板の形態解析を行う手法と、VFから椎間板領域を自動抽出する手法について述べた。今後、嚙下障害の有無が明確な大規模なVFデータセットの入手、提案手法の適用による椎間板を含む頸椎領域の自動抽出、自動抽出した嚙下障害患者群と正常嚙下群の頸椎領域に形態解析を行う手法の開発、嚙下障害と頸椎の形動態の関連性の評価を行うことによって、本研究の目的を達成できると考える。

## 謝辞

本研究を進めるにあたり、日々多大なるご指導、ご助言を頂きました筑波大学システム情報系情報工学域の滝沢穂高准教授に厚く御礼申し上げます。また、本研究に対してご助言、ご協力を頂きました筑波大学システム情報系情報工学域の工藤博幸教授に深く感謝致します。本論文の審査にあたっては、滝沢穂高准教授、工藤博幸教授、筑波大学システム情報系情報工学域の福井和広教授、同じく筑波大学システム情報系情報工学域の亀山啓輔教授、筑波大学システム情報系知能機能工学域系の黒田嘉宏教授にお願い致しました。ここに感謝を申し上げます。

研究遂行にあたり貴重なご協力を賜った神戸赤十字病院の目片幸二郎氏、京都大学大学院医学研究科人間健康科学系専攻の松林潤氏、岡山大学の瀧川朋亨氏、神戸赤十字病院の戸田一潔氏、伊藤康夫氏に感謝を申し上げます。

また、画像処理に関する様々な知見をくださった知能情報・生体工学研究室知的画像処理グループの皆様へ感謝致します。本研究では、齋藤祐季氏、八板駿平氏の整備したデータセットを一部使用しました。郡司絵莉華氏とは一部の実験を共同で進めました。

最後に、今日までの学生生活を支えてくださった家族と友人に感謝致します。

## 発表論文

### 雑誌論文

- [1] Ayano Fujinaka, Yuki Saito, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Automated Segmentation of Cervical Intervertebral Disks from Videofluorography Using a Convolutional Neural Network and its Performance Evaluation” *Journal of Signal Processing Systems*, Springer, Vol. 92, No. 3, pp. 299-305, 2020
- [2] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Segmentation of cervical intervertebral disks in videofluorography by CNN, multi-channelization and feature selection” *International Journal of Computer Assisted Radiology and Surgery*, Springer, pp. 1-8, 2020 (国際会議発表 [5] の予稿集論文を兼ねている)

### 国際会議発表

- [3] Ayano Fujinaka, Yuki Saito, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Segmentation of intervertebral disks from videofluorographic images using convolutional neural network” *International Forum on Medical Imaging in Asia 2019*, Vol. 11050, pp. 1-2, 2019
- [4] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Preliminary Study on Intervertebral Disk Segmentation from Videofluorography by Multi Channelization and CNN” *Proceedings of the 7th IIAE International Conference on Intelligent Systems and Image Processing 2019*, pp. 259-262, 2019
- [5] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Segmentation of cervical intervertebral disks in videofluorography by CNN, multi-channelization and feature selection” *Computer Assisted Radiology and Surgery 2020: 34th International Congress and Exhibition*, pp. 1-8, 2020
- [6] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Segmentation of cervical intervertebral disks from videofluorography based on semantic segmentation, multi-channelization, and feature selection: comparison of U-Net, FPN, LinkNet, and PSPNet” *The 16th Asian Conference on Computer Aided Surgery*, 2020

## 国内会議発表

- [7] 藤中彩乃, 齋藤祐季, 目片幸二郎, 滝沢穂高, 工藤博幸 「CNN を用いた嚙下時 X 線透視動画における頸椎椎間板抽出の基礎的検討」 電子情報通信学会技術報告書 (MI2018-32), Vol. 118, No. 225, pp. 1-2, 2018
- [8] 藤中彩乃, 目片幸二郎, 滝沢穂高, 工藤博幸 「マルチチャンネル化と CNN を用いた嚙下時 X 線透視動画における頸椎椎間板の抽出」 第 38 回日本医用画像工学会大会予稿集, pp. 377-378, 2019 (ポスター発表)
- [9] 藤中彩乃, 目片幸二郎, 滝沢穂高, 工藤博幸 「マルチチャンネル化の最適化と CNN による VF 画像からの椎間板の抽出」 マルチチャンネル化の最適化と CNN による VF 画像からの椎間板の抽出
- [10] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Preliminary Study on Segmentation of Cervical Intervertebral Disks in Videofluorography by Multi Channelization with Various Image Filters and CNN” Proceedings of 2nd Conference on Biological Image and Medical AI, 生体画像と医用人工知能研究会 (Biological Imaging and Medical AI), pp. 8, 2020
- [11] 藤中彩乃, 目片幸二郎, 滝沢穂高, 工藤博幸 「マルチチャンネル化と CNN を用いた嚙下時 X 線動画からの頸椎椎間板の抽出: LeNet, AlexNet, GoogLeNet の性能比較」 第 39 回日本医用画像工学会大会予稿集, pp. 196-200, 2020 (ポスター発表)

## ワークショップ発表

- [12] Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo “Segmentation of intervertebral disks from videofluorography based on CNN, multi-channelization, and feature selection” The 2020 UGA & UT Workshop, 2020 (口頭発表, ポスター発表)

## 採用

[1] 2020年4月より2022年3月まで日本学術振興会特別研究員DC2に採用

## 獲得研究費

[1] 日本学術振興会 科学研究費助成事業 (科学研究費補助金) 特別研究員奨励費 (JSPS 科研費 番号 20J10293)

## 受賞

[1] Best Presentation Award “Preliminary Study on Intervertebral Disk Segmentation from Videofluorography by Multi Channelization and CNN” The 7th IIAE International Conference on Intelligent Systems and Image Processing 2019, 2019



## 参考文献

- [1] 谷口洋「嚥下障害，診られますか？」羊土社，東京，pp. 13-35，2015
- [2] 進武幹，前山忠嗣，梅崎俊郎「嚥下障害」口腔・咽頭科，Vol. 1，pp. 93-101，1989
- [3] 木村彰男「嚥下障害の評価と治療」リハビリテーション医学，Vol. 28，No. 6，pp. 471-476，1991
- [4] Paul E. Marik, Danielle Kaplan “Aspiration Pneumonia and Dysphagia in the Elderly” Chest, Vol. 124, No. 1, pp. 328-336, 2003
- [5] Olle Ekberg, Shaheen Hamdy, Virginie Woisard, Anita Wuttge-Hannig, Primitivo Ortega “Social and Psychological Burden of Dysphagia; Its Impact on Diagnosis and Treatment” Dysphagia, Vol. 17, No. 2, pp. 139-146, 2002
- [6] 厚生労働省 「令和元年 (2019) 人口動態統計月報年計 (概数) の概況」 (<https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai19/dl/gaikyouR1.pdf>) Accessed December 2th 2020
- [7] 棚橋汀路，吉田義一「嚥下障害の X 線透視検査記録の試案」耳鼻と臨床，Vol. 34，No. 1，pp. 121-125，1988
- [8] 日本摂食嚥下リハビリテーション学会医療検討委員会 「嚥下造影の検査法 (詳細版)」 (<https://www.jsdr.or.jp/wp-content/uploads/file/doc/VF18-2-p166-186.pdf>) Accessed December 20th 2020
- [9] 日本耳鼻咽喉科学会「嚥下障害診療ガイドライン (2018 年版)」 ([http://www.jibika.or.jp/members/guidelines/enge\\_shougai\\_2018.pdf](http://www.jibika.or.jp/members/guidelines/enge_shougai_2018.pdf)), pp. 20-23, Accessed December 20th 2020
- [10] 日本医学放射線学会 「診断参考レベル (2020 年版)」 ([http://www.radher.jp/J-RIME/report/JapanDRL2020\\_jp.pdf](http://www.radher.jp/J-RIME/report/JapanDRL2020_jp.pdf)) Accessed December 2th 2020
- [11] 中瀬猛「腰椎椎間板の加齢変化に関する電顕的研究」日本医科大学雑誌，Vol. 52，No. 4，pp. 408-417，1985
- [12] 中寺尚志，岡田清，村田武穂「MRI による頸椎椎間板形態の検討」整形外科と災害外科，Vol. 37，No. 2，pp. 656-661，1988

- [13] 目片幸二郎, 滝沢穂高, 松林潤, 瀧川朋亨, 戸田一潔, 伊藤康夫, 工藤博幸 「嚙下時 X 線サジタル透視動画における頸椎前彎角に基づく頸椎疾患の識別」 日本コンピュータ外科学会誌, Vol. 21, No. 1, pp. 12-17, 2019
- [14] 根尾昌志 「Magerl 法の考え方と実際」 脊髄外科, Vol. 27, No. 1, pp. 22-27, 2013
- [15] 藤縄光留, 伊藤裕之, 小泉千秋, 前田淳一 「頸髄損傷者にみられた嚙下障害の予後に影響する因子」 第 28 回日本嚙下医学会, 耳鼻と臨床, Vol. 52, No. 1Supplement, pp. S66-S70, 2006
- [16] Masanori Izeki, Masashi Neo, Mitsuru Takemoto, Shunsuke Fujibayashi, Hiromu Ito, Koutatsu Nagai, Shuichi Matsuda “The O-C2 angle established at occipito-cervical fusion dictates the patient ’ s destiny in terms of postoperative dyspnea andor dysphagia” European Spine Journal, Vol. 23, No. 2, pp. 328-336, 2014
- [17] 森田颯馬, 松廣幹雄, 鈴木秀宣, 河田佳樹, 仁木登, 松元祐司, 土田敬明, 楠本昌彦, 金子昌弘 「胸部 CT 画像を用いた皮質骨および海綿骨の解析」 電子情報通信学会信学技報 (MI2020-37), Vol. 120, No. 220, pp. 21-23, 2020
- [18] 保田竜也, 滝沢穂高, 奥村俊昭, 工藤博幸, 岡田俊之 「腹部 X 線 CT 画像からの脊柱, 肋骨, 椎間板, 脊椎の段階的認識」 日本コンピュータ外科学会誌, Vol. 19, No. 3, pp. 131-138, 2017
- [19] Zhenyu Tang, Josef Pauli “Fully automatic extraction of human spine curve from MR images using methods of efficient intervertebral disk extraction and vertebra registration” International Journal of Computer Assisted Radiology and Surgery, Vol. 6, No. 1, pp. 21-33, 2011
- [20] Claudia Chevretil, Farida Chériet, Guy Grimard, Carl-Eric Aubin “Watershed Segmentation of Intervertebral Disk and Spinal Canal from MRI Images” International Conference Image Analysis and Recognition, pp. 1017-1027, 2007
- [21] Yiebin Kim, Dongsung Kim “A fully automatic vertebra segmentation method using 3D deformable fences” Computerized Medical Imaging and Graphics, Vol. 33, No. 5, pp. 343-352, 2009
- [22] Sovira Tan, Jianhua Yao, Lawrence Yao, Michael M. Ward “High precision semiautomated computed tomography measurement of lumbar disk and vertebral heights” Medical physics, Vol. 40, No. 1, pp. 011905, 2013
- [23] Robert Korez, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec “Parametric modeling of the intervertebral disc space in 3D: Application to CT images of the lumbar spine” Computerized Medical Imaging and Graphics, Vol. 38, No. 7, pp. 596-605, 2014

- [24] Xing Ji, Guoyan Zheng, Daniel Belavy, Dong Ni “Automated Intervertebral Disc Segmentation Using Deep Convolutional Neural Networks” International Workshop on Computational Methods and Clinical Applications for Spine Imaging, pp. 38-48, 2016
- [25] Xiaomeng Li, Qi Dou, Hao Chen, Chi-Wing Fu, Xiaojuan Qi, Daniel L. Belavý, Gabriele Ambrecht, Dieter Felsenberg, Guoyan Zheng, Pheng-Ann Heng “3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images” Medical image analysis, Vol. 45, pp. 41-54, 2018
- [26] Ruhan Sa, William Owens, Raymond Wiegand, Mark Studin, Donald Capoferri, Kenneth Barooha, Alexander Greaux, Robert Rattray, Adam Hutton, John Cintineo, Vipin Chaudhary “Intervertebral disc detection in X-ray images using faster R-CNN” 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, pp. 564-567, 2017
- [27] Yuko Shigeta, Takumi Ogawa, Eriko Ando, Glenn T. Clark, Reyes Enciso “Influence of tonguemandible volume ratio on oropharyngeal airway in Japanese male patients with obstructive sleep apnea” Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, Vol. 112, No. 2, pp. 239-243, 2011
- [28] Sampaio Rafael De Assuncao, Marcel Parolin Jackowski “Vocal tract morphology using real-time magnetic resonance imaging” 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images, IEEE, pp. 359-366, 2017
- [29] Samuel Silva, António Teixeira “Unsupervised segmentation of the vocal tract from real-time MRI sequences” Computer Speech & Language, Vol. 33, No. 1, pp. 25-46, 2015
- [30] Palmer Jeffrey B., Eiichi Tanaka, Erik Ensrud “Motions of the posterior pharyngeal wall in human swallowing: a quantitative videofluorographic study” Archives of physical medicine and rehabilitation, Vol. 81, No. 11, pp. 1520-1526, 2000
- [31] Jong Taek Lee, Eunhee Park, Tae-Du Jung “Automatic detection of the pharyngeal phase in raw videos for the videofluoroscopic swallowing study using efficient data collection and 3d convolutional networks” Sensors, Vol. 19, No. 18, 3873, 14 pages, 2019
- [32] Jung Chan Lee, Han Gil Seo, Woo Hyung Lee, Hee Chan Kim, Tai Ryoan Han, Byung-Mo Oh “Computer-assisted detection of swallowing difficulty” Computer Methods and Programs in Biomedicine, Vol. 134, pp. 79-88, 2016
- [33] Seong Jae Lee, Joo Young Ko, Hyun Il Kim, Sang-II Choi “Automatic Detection of Airway Invasion from Videofluoroscopy via Deep Learning Technology” Applied Sciences, Vol. 10, No. 18, 6179, 16 pages, 2020

- [34] 八板駿平, 滝沢穂高, 目片幸二郎, 工藤博幸 「Support vector machine を用いた嚙下時 X 線サジタル透視動画における舌骨体の追跡」 *Medical Imaging Technology*, Vol. 36, No. 5, pp. 209-216, 2018
- [35] Zhenwei Zhang, James L. Coyle, Ervin Sejdić “Automatic hyoid bone detection in fluoroscopic images using deep learning” *Scientific reports*, Vol. 8, No. 1, pp. 1-9, 2018
- [36] Youngkook Kim, Geun-Young Park, Yu Jung Seo, Sun Im “Effect of anterior cervical osteophyte in poststroke dysphagia: a case-control study” *Archives of Physical Medicine and Rehabilitation*, Vol. 96, No. 7, pp. 1269-1276, 2015
- [37] Rianne Reinartz, Bram Platel, Toon Boselie, Henk van Mameren, Henk van Santbrink, Bart ter Haar Romeny “Cervical vertebrae tracking in video-fluoroscopy using the normalized gradient field” *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 524-531, 2009
- [38] 齋藤祐季, 目片幸二郎, 滝沢穂高, 工藤博幸 「SVM を用いた X 線透視画像における椎間板抽出の基礎的検討」 第 36 回日本医用画像工学会予稿集, Vol. 1, No. 1, pp. 292-294, 2017
- [39] Connor Shorten, Taghi M. Khoshgoftaar “A survey on Image Data Augmentation for Deep Learning” *Journal of Big Data*, Vol. 6, No. 1, 48 pages, 2019
- [40] Martin A. Fischler, Robert C. Bolles “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography” *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, 1981
- [41] Yann Lecun, Léon Bottou, Yoshua Bengio, Patrick Haffner “Gradient-based learning applied to document recognition” *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, 1998
- [42] Github NVIDIA DIGITS (<https://github.com/NVIDIA/DIGITS>) Accessed December 21th 2020
- [43] 村上伸一 「画像処理工学」 東京電機大学出版局, pp. 28-33, 1996
- [44] 浅野晃, 浅野千恵, 木森義隆, 棟安実治, 延原肇, 藤尾光彦 「非線形画像・信号処理 モルフォロジーの基礎と応用」 丸善株式会社, pp. 43-78, 2010
- [45] William H. Press, S. A. Teukolsky, W.T. Vetterling, B.P. Flannery 「ニューメリカルレシピ・イン・シー C 言語による数値計算のレシピ」 *Numerical Recipes in C 日本語版*. 技術評論社, pp.330-338, 1993
- [46] 田村秀行 「コンピュータ画像処理」 オーム社, 2002

- [47] 平井有三 「はじめてのパターン認識」 森北出版, 2012
- [48] 手塚太郎 「しくみがわかる深層学習」 朝倉書店, 2018
- [49] 増井敏克 「プログラマのためのディープラーニングのしくみがわかる数学入門」 ソシム, 2018
- [50] 斎藤康毅 「ゼロから作る Deep Learning -Python で学ぶディープラーニングの理論と実装」 オライリージャパン, 2016
- [51] 久保幹雄, 戸川隼人, 杉原厚吉, 中嶋正之, 野寺隆志 「組合せ最適化とアルゴリズム (インターネット時代の数学シリーズ 8)」 共立出版, 2000
- [52] 穴井宏和, 斉藤努 「今日から使える! 組合せ最適化 離散問題ガイドブック」 講談社, 2015
- [53] 中井悦司 「TensorFlow で学ぶディープラーニング入門 畳み込みニューラルネットワーク徹底解説」 マイナビ出版, 2016
- [54] 下田倫大, 杉義宏, 太田満久, 久保隆宏, 有山圭二 「TensorFlow 活用ガイド」 技術評論社, 2018
- [55] M. Mohsin Jadoon, Qianni Zhang, Ihsan Ul Haq, Sharjeel Butt, Adeel Jadoon “Three-class mammogram classification based on descriptive CNN features” BioMed research international, pp. 1-11, 2017
- [56] Yun Jiang, Zhanhuai Li, Longbo Zhang, Peng Sun “An Improved SVM Classifier for Medical Image Classification” International Conference on Rough Sets and Intelligent Systems Paradigms, pp. 764-773, 2007
- [57] Loris Nanni, Alessandra Lumini, Sheryl Brahmam “Local binary patterns variants as texture descriptors for medical image analysis” Artificial Intelligence in Medicine, Vol. 49, No. 2, pp. 117-125, 2010
- [58] R. J. Ramteke, Khachane Monali Y. “Automatic medical image classification and abnormality detection using k-nearest neighbour” International Journal of Advanced Computer Research, Vol. 2, No. 4, pp. 190-196, 2012
- [59] Chi-Hoon Lee, Mark Schmidt, Albert Murtha, Aalo Bistritz, Jörg Sander, Russell Greiner “Segmenting Brain Tumors with Conditional Random Fields and Support Vector Machines” International Workshop on Computer Vision for Biomedical Image Applications, pp. 469-478, 2005
- [60] Yejin Jeon, Eunbyul Cho, Sehwa Moon, Seung-Hoon Chae, Hae Young Jo, Tae Oh Kim, Chang Mo Moon, Jang-Hwan Choi “Deep convolutional neural network-based automated

- lesion detection in wireless capsule endoscopy” International Forum on Medical Imaging in Asia 2019, Vol. 11050, pp. 1-5, International Society for Optics and Photonics, 2019
- [61] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei “Imagenet: A large-scale hierarchical image database” 2009 IEEE conference on computer vision and pattern recognition, IEEE, pp. 248-255, 2009
- [62] Wei Shen, Mu Zhou, Feng Yang, Caiyun Yang, Jie Tian “Multi-scale convolutional neural networks for lung nodule classification” International Conference on Information Processing in Medical Imaging, Springer, Cham, pp. 588-599, 2015
- [63] Gao Huang, Danlu Chen, Tianhong Li, Felix Wu, Laurens van der Maaten, Kilian Q. Weinberger “Multi-scale dense networks for resource efficient image classification” arXiv preprint arXiv:1703.09844, 2017
- [64] 平野銀次, 根本充貴, 木村裕一, 永岡隆 「深層学習を用いたメラノーマの自動診断システムの開発 -病変部位強調画像の採用-」 電子情報通信学会信学技報 (MI2020-46), Vol. 120, No. 220, pp. 62-64, 2020
- [65] Morteza Heidari, Seyedehnafiseh Mirniaharikandehei, Abolfazl Zargari Khuzani, Gopichandh Danala, Yuchen Qiu, Bin Zheng “Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms” International Journal of Medical Informatics, Vol. 144, 104284, 9 pages, 2020
- [66] 数尾昂, 工藤博幸, 川又 健太郎 「深層学習による圧縮センシング CT 画像再構成画質改善の新技术」 第 39 回日本医用画像工学会大会予稿集, pp. 368-372, 2020
- [67] 松山江里, 李鎔範, 高橋規之, 蔡篤儀 「ウェーブレット係数を用いた CNN による CT 画像の肺がんの組織型分類」 医用画像情報学会雑誌, Vol. 36, No. 2, pp. 64-71, 2019
- [68] Girish Chandrashekar, Ferat Sahin “A survey on feature selection methods” Computers & Electrical Engineering, Vol. 40, No. 1, pp. 16-28, 2014
- [69] Miron B. Kursa, Witold R. Rudnicki “Feature selection with the Boruta package” Journal of Statistical Software, Vol. 36, No. 11, pp. 1-13, 2010
- [70] Pin Zhang “A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model” Applied Soft Computing, Vol. 85, 105859, 12 pages, 2019
- [71] Labiba Gillani Fahad , Syed Fahad Tahir, Waseem Shahzad, Mehdi Hassan, Hani Alquhayz, Rabia Hassan “Ant Colony Optimization-Based Streaming Feature Selection: An Application to the Medical Image Diagnosis” Scientific Programming, Vol. 2020, 1064934, 10 pages, 2020

- [72] 根本充貴, 清水昭伸, 萩原義裕, 小畑秀文, 縄野繁 「多数の特徴量からの特徴選択による乳房 X 線像上の腫瘍影判別精度の改善と高速な特徴選択法の提案」 電子情報通信学会論文誌 D, Vol. J88-D2, No. 2, pp. 416-426, 2005
- [73] Jothi G, Hannah Inbarani H. “Hybrid Tolerance Rough Set–Firefly based supervised feature selection for MRI brain tumor image classification” *Applied Soft Computing*, Vol. 46, pp. 639-651, 2016
- [74] Muhammad Attique Khan, Imran Ashraf, Majed Alhaisoni, Robertas Damaševičius, Rafal Scherer, Amjad Rehman, Syed Ahmad Chan Bukhari “Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists” *Diagnostics*, Vol. 10, No. 8, 565, 19 pages, 2020
- [75] M. Toğaçar, B. Ergen, Z. Cömert, F. Özyurt “A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models” *Irbm*, Vol. 41, No. 4, pp. 212-222, 2020
- [76] Javad Salimi Sartakhti, Mohammad Hossein Zangooei, Kourosh Mozafari “Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA)” *Computer Methods and Programs in Biomedicine*, Vol. 108, No. 2, pp. 570-579, 2012
- [77] A. Asuntha, A. Brindha, S. Indirani, Andy Srinivasan “Lung cancer detection using SVM algorithm and optimization techniques” *Journal of Chemical and Pharmaceutical Sciences*, Vol. 9, No. 4, pp. 3198-3203, 2016
- [78] 奥富正敏 「デジタル画像処理」 CG-ARTS 協会, pp.88-126, 2004
- [79] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton “Imagenet classification with deep convolutional neural networks” *Advances in neural information processing systems*, pp. 1097-1105, 2012
- [80] Olaf Ronneberger, Philipp Fischer, Thomas Brox “U-net: Convolutional networks for biomedical image segmentation” *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, pp. 234-241, 2015
- [81] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie “Feature pyramid networks for object detection” *Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 2117-2125, 2017
- [82] Abhishek Chaurasia, Eugenio Culurciello “Linknet: Exploiting encoder representations for efficient semantic segmentation” *2017 IEEE Visual Communications and Image Processing*, IEEE, pp. 1-4, 2017

- [83] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia “Pyramid scene parsing network” Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2881-2890, 2017
- [84] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun “Deep residual learning for image recognition” Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016
- [85] Tensorflow (<https://www.tensorflow.org/>) Accessed December 21th 2020
- [86] Keras (<https://keras.io/>) Accessed December 21th 2020
- [87] Yuhang Liu, Fandong Zhang, Qianyi Zhang, Siwen Wang, Yizhou Wang, Yizhou Yu “Cross-View Correspondence Reasoning Based on Bipartite Graph Convolutional Network for Mammogram Mass Detection” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 3812-3822, 2020
- [88] Han Yang, Xingjian Zhen, Ying Chi, Lei Zhang, Xian-Sheng Hua “CPR-GCN: Conditional Partial-Residual Graph Convolutional Network in Automated Anatomical Labeling of Coronary Arteries” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 3803-3811, 2020
- [89] Victor Kulikov, Victor Lempitsky “Instance Segmentation of Biological Images Using Harmonic Embeddings” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 3843-3851, 2020
- [90] Noriaki Hashimoto, Daisuke Fukushima, Ryoichi Koga, Yusuke Takagi, Kaho Ko, Kei Kohno, Masato Nakaguro, Shigeo Nakamura, Hidekata Hontani, Ichiro Takeuchi “Multi-scale Domain-adversarial Multiple-instance CNN for Cancer Subtype Classification with Unannotated Histopathological Images” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 3852-3861, 2020
- [91] Sam Maksoud, Kun Zhao, Peter Hobson, Anthony Jennings, Brian C. Lovell “SOS: Selective Objective Switch for Rapid Immunofluorescence Whole Slide Image Classification” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 3862-3871, 2020
- [92] Jiahua Dong, Yang Cong, Gan Sun, Bineng Zhong, Xiaowei Xu “What Can Be Transferred: Unsupervised Domain Adaptation for Endoscopic Lesions Segmentation” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 4023-4032, 2020
- [93] Qier Meng, Satoh Shin’ichi “ADINet: Attribute Driven Incremental Network for Retinal Image Classification” Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, pp. 4033-4042, 2020



- [94] Qihang Yu, Dong Yang, Holger Roth, Yutong Bai, Yixiao Zhang, Alan L. Yuille, Daguang Xu “C2FNAS: Coarse-to-Fine Neural Architecture Search for 3D Medical Image Segmentation” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4126-4135, 2020
- [95] Yu Zhao, Fan Yang, Yuqi Fang, Hailing Liu, Niyun Zhou, Jun Zhang, Jiarui Sun, Sen Yang, Bjoern Menze, Xinjuan Fan, Jianhua Yao “Predicting Lymph Node Metastasis Using Histopathological Images Based on Multiple Instance Learning With Deep Graph Convolution” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4837-4846, 2020
- [96] Hong Joo Lee, Jung Uk Kim, Sangmin Lee, Hak Gu Kim, Yong Man Ro “Structure Boundary Preserving Segmentation for Medical Image With Ambiguous Boundary” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4817-4826, 2020
- [97] Hanchao Yu, Shanhui Sun, Haichao Yu, Xiao Chen, Honghui Shi, Thomas S. Huang, Terrence Chen “FOAL: Fast Online Adaptive Learning for Cardiac Motion Estimation” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4313-4323, 2020
- [98] Adithya Balaji, Alexander Allen “Benchmarking automatic machine learning frameworks” arXiv preprint arXiv:1808.06492, 9 pages, 2018
- [99] Hayder Hasan, Helmi Z.M. Shafri, Mohammed Habshi “A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Hyperspectral Image Classification” IOP Conference Series: Earth and Environmental Science, Vol. 357, No. 1, 11 pages, 2019
- [100] Bin Zhang, Changqin Quan, Fuji Ren “Study on CNN in the recognition of emotion in audio and images” 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pp. 1-5, 2016
- [101] Luis Perez, Jason Wang “The effectiveness of data augmentation in image classification using deep learning” arXiv preprint arXiv:1712.04621, 8 pages, 2017
- [102] Pranav Rajpurkar, Allison Park, Jeremy Irvin, Chris Chute, Michael Bereket, Domenico Mastrodicasa, Curtis P. Langlotz, Matthew P. Lungren, Andrew Y. Ng, Bhavik N. Patel “AppendixNet: Deep Learning for Diagnosis of Appendicitis from A Small Dataset of CT Exams Using Video Pretraining” Scientific reports, Vol. 10, No. 1, pp. 1-7, 2020
- [103] Misgina Tsighe Hagos, Shri Kant “Transfer Learning based Detection of Diabetic Retinopathy from Small Dataset” arXiv preprint arXiv:1905.07203, 9 pages, 2019

- [104] Ye Liu, Shuohong Wang, Hao Gao, Baoyun Wang “Detecting dense crowds of microbes from microscope images in a global optimization framework” *Optik*, Vol. 127, No. 1, pp. 76-80, 2016
- [105] Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, Yutaka Satoh “Pre-training without Natural Images” *Proceedings of the Asian Conference on Computer Vision*, 17 pages, 2020
- [106] Jun Shi, Shichong Zhou, Xiao Liu, Qi Zhang, Minhua Lu, Tianfu Wang “Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset” *Neurocomputing*, Vol. 194, pp. 87-94, 2016
- [107] 文部科学省, 厚生労働省 「人を対象とする医学系研究に関する倫理指針」 (<https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000153339.pdf>) Accessed February 12th 2021