

令和4年度
博士論文

AIを用いたパノラマX線画像からの
カルテ入力支援システムに関する研究

鳥井 浩平

徳島大学大学院先端技術科学教育部
システム創生工学専攻 博士後期課程3年

目次

第 1 章	序論	5
1.1	医療と人工知能	5
1.2	歯科診療におけるデジタルデータ	6
1.3	研究目的	11
第 2 章	関連研究	12
2.1	畳み込みニューラルネットワークの基本技術	12
2.2	EfficientNetV2	22
2.3	医用画像を用いた AI システム	24
2.4	CT 画像からのパノラマ再構成	26
第 3 章	カルテ入力支援システムの構築	28
3.1	パノラマ X 線画像データベース	28
3.2	カルテ入力支援システム	41
第 4 章	カルテ入力支援システムの評価	46
4.1	状態診断モデルの評価	46
4.2	テスト用パノラマ X 線画像に対するカルテ入力支援システムの評価	49
4.3	CT 画像から生成したパノラマ X 線画像に対するカルテ入力支援の例	58
第 5 章	考察	63
第 6 章	結論	66
	謝辞	67
	参考文献	68
	付録 A	73

表目次

1.1	カルテ用語	9
1.2	FDI 方式に基づく永久歯の歯式	10
1.3	FDI 方式に基づく乳歯の歯式	10
1.4	日本における一般診療所の電子カルテ普及率	10
3.1	DICOM タグと格納値の抜粋	29
3.2	状態ラベルの定義	30
3.3	アノテーションファイルの構成	36
3.4	学習用パノラマ X 線画像データベースにおける状態ラベルの内訳 (種類順)	39
3.5	学習用パノラマ X 線画像データベースにおける状態ラベルの内訳 (データ数順)	39
3.6	9 状態および 20 状態の定義	40
3.7	テスト用パノラマ X 線画像データベースにおける状態ラベルの内訳	40
3.8	各データセットにおける状態ラベルの内訳	45
4.1	状態診断モデルの交差検証による評価	47
4.2	評価値の計算例	48
4.3	カルテ入力支援システムの感度	50
4.4	歯の検出におけるカルテ入力支援システムの感度	51
4.5	歯の状態診断におけるカルテ入力支援システムの感度	52
A.1	状態診断モデルにおける Dataset1 の評価	74
A.2	状態診断モデルにおける Dataset2 の評価	75
A.3	状態診断モデルにおける Dataset3 の評価	76
A.4	状態診断モデルにおける Dataset4 の評価	77
A.5	状態診断モデルにおける Dataset5 の評価	78

目次

1.1	パノラマ X 線画像の例	6
1.2	パノラマ X 線撮影法における断層域	7
1.3	歯科用 CT 画像の例	8
2.1	5 層のニューラルネットワーク	13
2.2	損失関数の最小化のイメージ	15
2.3	畳み込み計算の例	17
2.4	最大プーリングと平均プーリングの例	19
2.5	残差接続	20
2.6	アップコンボリューション	21
2.7	MBCConv と Fused-MBCConv のネットワーク構造	22
2.8	SE モジュールのネットワーク構造	23
2.9	EfficientNetV2-S のネットワーク構造	23
2.10	CT 画像からのパノラマ再構成のフロー	26
3.1	Anotee	31
3.2	マルチ周波数処理における周波数強調のアルゴリズム	33
3.3	マルチ周波数処理におけるダイナミックレンジ圧縮のアルゴリズム	34
3.4	ダイナミックレンジ圧縮における濃度変換のイメージ	34
3.5	Anotee のラベルダイアログ	35
3.6	Mask R-CNN を用いた自動ポリゴン作成	37
3.7	カルテ入力支援システムの運用イメージ	41
3.8	カルテ入力支援システムの概略図	42
3.9	状態診断モデル	43
3.10	画像の前処理	44
4.1	カルテ入力支援システムの出力例 1	53
4.2	カルテ入力支援システムの出力例 2	54
4.3	カルテ入力支援システムの出力例 3	55
4.4	カルテ入力支援システムの出力例 4	56

4.5	カルテ入力支援システムの出力例 5	57
4.6	生成したパノラマ X 線画像の例	59
4.7	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 1	60
4.8	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 2	60
4.9	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 3	61
4.10	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 4	61
4.11	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 5	62
4.12	生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 6	62

第 1 章

序論

1.1 医療と人工知能

診療情報の電子化とインターネットの普及により大量の医療データを収集することが容易になった現代では、さまざまな医療ビッグデータの構築が進み、人工知能（AI）への活用が盛んに行われている。AI は医療ビッグデータから診療パターンや画像とカルテ情報の関係などを学習することで、実際の間診や検査結果などから自動的に可能性の高い疾患の診断病名を提示したり、画像から診断材料として重要な部位を強調表示したりすることができる。AI による診断支援システムは医療従事者の業務負担軽減やより正確な診断に役立っており、さまざまな医療関係の分野で需要が高まっている。たとえば、米 IBM 社の Watson for Genomics (WfG) [1] はがんの発症や進行に関与している遺伝子変異の候補、その根拠となるデータ、抗がん剤の候補を提示することができる。WfG は医学論文や薬の特許情報などで構成される大規模なデータベースを学習した AI システムであり、がんに関係する遺伝子の塩基配列の解析結果を入力することで、10 分以下の処理時間でがんの対処に有益な情報を得ることができる。がんに関する解析作業は熟練した専門医でも通常 1 週間から 2 週間ほどかかるが、WfG は専門医の約 1,000 倍以上の速さで解析を行うことができるため、がんに対してより迅速な対処が可能になり、結果として救命の一助となることもある。実際、WfG の診断支援によって特殊な白血病を患った 60 代女性患者の命が救われたという事例があり、このことは WfG の有用性、すなわち AI による診断支援の有用性を示した。

歯科分野においても同様に AI を用いた診断支援システムの開発が盛んに行われており、すでに実用化された診断支援システムも登場している。たとえば、ドイツの dentalXrai 社が提供している Explore dentalXrai Pro 3.0^{*1} は AI を用いてパノラマ X 線画像の自動診断を行い、う蝕と思われる箇所や歯の修復の状態などを可視化して表示する。AI は数百万件の高品質なデータを学習しており、さまざまな診断を高速かつ高精度で行うことができる。ユーザインターフェースも優れており、タブレットなどのさまざまな端末で診断結果の表示を手軽に行うことができるため、患者に対して疾患や治療内容の説明を行うために用いるコミュニケーションツールとしても有用である。

^{*1} <https://www.dentalxr.ai/en/technology/software-version-3-0/>



図 1.1 パノラマ X 線画像の例. 主に金属などは X 線不透過像として写り、この例ではメタルインレーや全部金属冠、根管充填などを確認できる。

また、診断結果から自動的にカルテを生成する機能があるため、診断結果の記録に要する時間を短縮し、より効率的な診断を支援することができる。このように、AI を用いた診断支援はさまざまなメリットがあり、AI を用いた歯科診断支援システムの開発が求められている。

1.2 歯科診療におけるデジタルデータ

1.2.1 パノラマ X 線画像

パノラマ X 線画像はパノラマ X 線撮影法によって得られる上下顎骨を一枚に収めた X 線画像であり、欠損歯や埋伏歯の確認、全顎に至る歯周病の診断などに用いられる。パノラマ X 線画像の例を図 1.1 に示す。パノラマ X 線撮影法は断層撮影の一種であり、単純撮影に比べて像の重複が少ない画像を得ることができる。パノラマ X 線撮影法では、まず撮影機が頭部の周りを回転しながらスリットを用いて短冊型の X 線画像を複数撮影する。短冊型の X 線画像は前歯部や下顎枝などの歯列または顎骨の一部を回転軸として断層撮影した X 線画像である。短冊型の X 線画像に像として現れる回転軸周辺の一定範囲を断層域といい、パノラマ X 線撮影法では図 1.2 のようになる。最後に、短冊型の X 線画像を平面上に再構成することでパノラマ X 線画像を得る。

1.2.2 歯科用 CT 画像

歯科用 CT 画像はコーンビーム CT (CBCT) を用いて撮影される頭部の 3D 画像である。CT 画像の例を図 1.3 に示す。CBCT では、円錐状の X 線ビーム (コーンビーム) を被写体に照射して回転撮影を行い、撮影した画像データをコンピュータを用いて再構成することで 3D 画像を得る。撮影条件にもよるが、CBCT による撮影は一般的な医科用 CT よりも被曝量が少なく、撮影

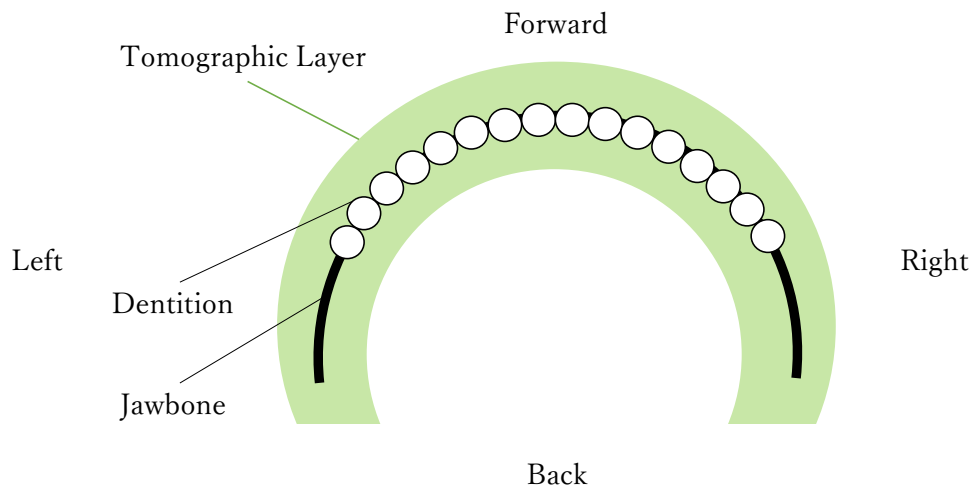


図 1.2 パノラマ X 線撮影法における断層域（緑の領域）. 俯瞰視点（軸位断面）における顎骨（Jawbone）と歯列（Dentition）を表した図. パノラマ X 線撮影法における断層域（Tomographic Layer）の形状は顎骨と歯列に沿うように湾曲している. 歯列と顎骨が断層域から外れている場合は正確なパノラマ X 線画像を得ることができない.

時間も短いというメリットがある. 一方で, 撮影範囲が狭いことや軟組織の診断が難しいことがデメリットとして挙げられる. 歯科用 CT 画像は歯や顎骨, 上顎洞, 下顎管などの立体的構造を平面上に表現できるため, インプラント治療や親知らずの抜歯, 顎関節症などの診断に用いられる.

1.2.3 歯科診療録（カルテ）

カルテには個人情報に加えて歯の番号（歯式）や傷病, 治療履歴などを記入する. 傷病名および処置名は表 1.1 に示すようなカルテ用語を用いて記載する. 歯式の付け方には, Zsigmondy-Palmer 方式 [3, 4], 国際歯科連盟 (Federation Dentaire Internationale, FDI) 方式*², Universal 方式 [5] があり, 日本, ヨーロッパ, オーストラリアでは Zsigmondy-Palmer 方式, アメリカでは Universal 方式が主に用いられる. FDI 方式は国際標準化機構 (International Organization for Standardization, ISO) で定められた歯式方式であり, 世界保健機関 (World Health Organization, WHO) や FDI の国際基準となっている. FDI 方式の定義を表 1.2 および表 1.3 に示す.

歯科におけるカルテの作成では, 歯科用カルテ用紙に手書きで記入する方法と, 歯科用レセプトコンピュータ (レセコン) を用いて入力する方法があり, カルテの保存形式にも紙媒体または電子データの 2 通りがある. カルテを電子データで作成および管理するシステムまたは記録を電子カルテという. 歯科医院における電子カルテの普及率は年々上昇している*³ (表 1.4).

*² ISO 3950:2016 Dentistry – Designation system for teeth and areas of the oral cavity

*³ 厚生労働省. 電子カルテシステム等の普及状況の推移. 医療施設調査, 2020.

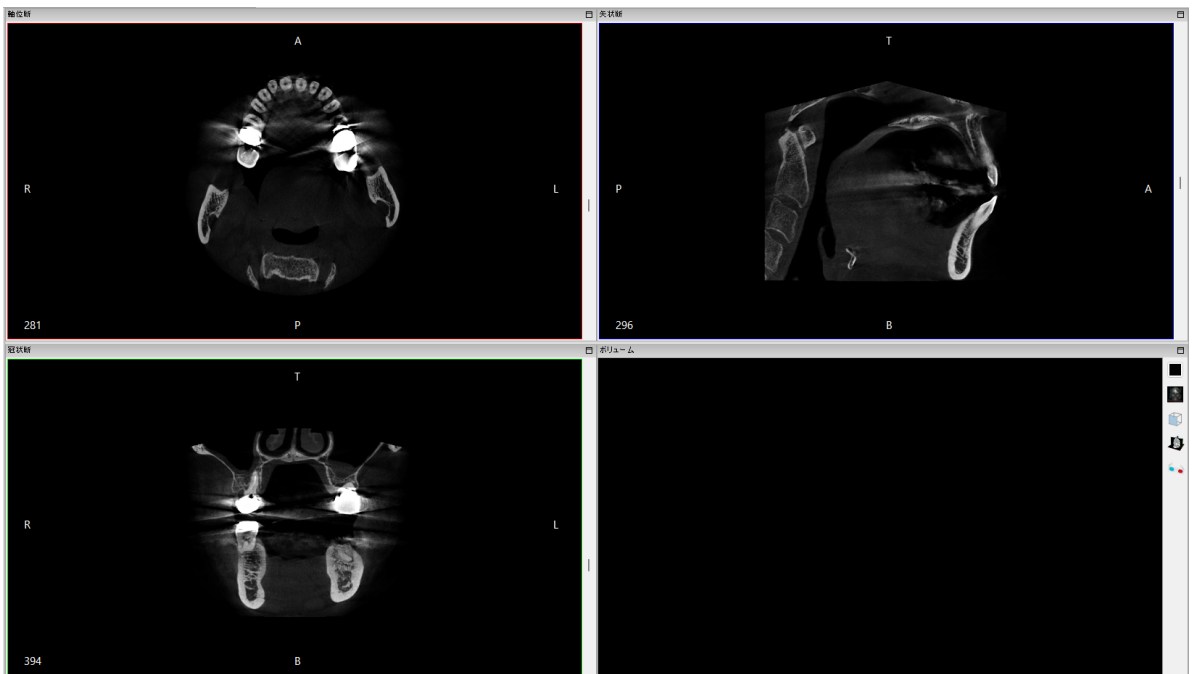


図 1.3 歯科用 CT 画像の例. CT 画像ビューア InVesalius[2] を用いて, 歯科用 CT 画像における軸位断面 (左上), 矢状断面 (右上), 冠状断面 (左下) を表示している.

表 1.1 カルテ用語. 略称 (Label) は共同研究先の歯科医院における定義に基づく. 本研究で扱った傷病および処置のカルテ用語を抜粋して記載している.

Name	Label	Discription
う蝕第 2 度	C2	う蝕が象牙質へ達した状態
う蝕第 3 度	C3	う蝕が歯髄まで達した状態
う蝕第 4 度	C4	う蝕が進行し歯髄のみが残存している状態
根尖性歯周炎	Per	歯根周辺に炎症が発生した状態
歯根嚢胞	WZ	歯根周辺に膿が溜まった状態
不適合	Futeki	詰め物や被せ物と歯の間に隙間がある状態
脱離	Daturi	詰め物や被せ物が外れた状態
破折	Frt	歯が欠けている状態
根分岐部病変	Bunki	歯根の分岐部部分まで組織破壊が進んだ状態
乳歯晩期残存	RDT	乳歯が萌出せずに残存している状態
過剰歯	SNT	通常の歯の本数よりも多く生えてきた歯
捻転歯	ROT	歯が捻れている状態
根管充填	RCF	根管内に樹脂や薬剤を詰めた状態
メタルインレー	In	金属製の詰め物
コンポジットレジン	CR	合成樹脂の詰め物
アマルガム充填	AF	アマルガムの詰め物
全部金属冠	FMC	金属製の被せ物
陶材焼付鑄造冠	MB	金属にセラミックを焼き付けた被せ物
硬質レジンジャケット冠	HJC	硬質レジン製の被せ物
レジンジャケット冠	RJC	レジン製の被せ物
レジン前装金属冠	VMC	金属に硬質レジンを焼き付けた被せ物
CAD/CAM 冠	CADCAM	コンピュータを用いて作成した被せ物
ジルコニアクラウン	Zi	ジルコニア製の被せ物
セラミッククラウン	Cera	セラミック製の被せ物
五分の四冠	4/5Cro	被せ物を被せる度合いを 5 分の 4 に調節した状態
四分の三冠	3/4Cro	被せ物を被せる度合いを 4 分の 3 に調節した状態
ブリッジ	Br	連結した人工歯を残存する永久歯に装着したもの
ポンティック	Pon	ブリッジにおいて欠損歯を補っている人口歯
インプラント	Implant	人口歯を装着するための人工物が埋まった状態
埋伏歯	RT	歯が歯肉に埋まった状態
半埋伏歯	HRT	歯冠の一部が見えている埋伏歯
完全埋伏歯	CRT	歯が完全に顎骨内にある埋伏歯
水平智歯	HET	智歯が水平に萌出した状態
水平埋伏智歯	HIT	智歯が顎骨内で水平に埋まった状態

表 1.2 FDI 方式に基づく永久歯の歯式. 中切歯を 1, 側切歯を 2, 犬歯を 3, 第一小白歯を 4, 第二小白歯を 5, 第一大臼歯を 6, 第二大臼歯を 7, 第三大臼歯を 8 とする. また, 上顎左側を 1, 上顎右側を 2, 下顎右側を 3, 下顎左側を 4 とする. 各歯を位置, 種類の順に番号を示す.

18	17	16	15	14	13	12	11	21	22	23	24	25	26	27	28
□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
48	47	46	45	44	43	42	41	31	32	33	34	35	36	37	38

表 1.3 FDI 方式に基づく乳歯の歯式. 乳中切歯を 1, 乳側切歯を 2, 乳犬歯を 3, 第一乳臼歯を 4, 第二乳臼歯を 5 とする. また, 上顎左側を 5, 上顎右側を 6, 下顎右側を 7, 下顎左側を 8 とする. 各歯を位置, 種類の順に番号を示す.

55	54	53	52	51	61	62	63	64	65
□	□	□	□	□	□	□	□	□	□
□	□	□	□	□	□	□	□	□	□
85	84	83	82	81	71	72	73	74	75

表 1.4 日本における一般診療所の電子カルテ普及率

Year	Penetration (%)	Breakdown
2008	14.7	14,602 / 99,083
2011	21.2	20,797 / 98,004
2014	35.0	35,178 / 100,461
2017	41.6	42,167 / 101,471
2020	49.9	51,199 / 102,612

1.3 研究目的

我々は従来の診断支援に関する技術をカルテ入力支援に応用するために、詳細なアノテーションを付与した歯科データベースの構築と、パノラマ X 線画像を対象としたカルテ入力支援システムの研究開発を行ってきた。カルテ入力支援システムは歯検出モデル、セグメンテーションモデル、状態診断モデルを用いて、パノラマ X 線画像から歯の検出、歯式推定、歯の輪郭推定および歯の状態診断を行うシステムであり、パノラマ X 線画像から簡易的な画像診断情報を作成することができる。カルテ入力支援システムとパノラマ X 線画像ビューアおよび電子カルテ（レセコン）を連携することでカルテの自動入力を可能にするため、歯科医のカルテ作成に必要な労力を軽減し、診断全体の効率化に貢献することができる。また、歯科医の画像診断に AI の画像診断が加わることで画像診断における見落としを防ぐことも期待できる。ただし、最終的なカルテの確定は歯科医が視診等を踏まえた診断のあとにカルテを修正して行うため、カルテ入力支援システムは確定診断のために用いるものではなく、画像から明確に判断できる情報を AI が事前に自動入力することを目的としている。

カルテ入力支援システムの有用性を検討するために、本論文では構築したパノラマ X 線画像データベースとデータ作成のプロセスについて述べ、カルテ入力支援システムに対する精度評価を行う。また、パノラマ X 線画像再構成システムを用いて CT 画像からパノラマ X 線画像を自動生成し、生成したパノラマ X 線画像に対するカルテ入力支援システムの応用可能性について述べる。なお、本研究は徳島大学大学院社会産業理工学研究部の倫理委員会から承認を得ている（承認番号:20002、課題名:AI を用いた歯科診断システムの研究開発）。

第 2 章

関連研究

2.1 畳み込みニューラルネットワークの基本技術

医用画像を対象にした診断支援システム開発で用いられる AI 技術に畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) がある。2014 年に発表された AlexNet[6] の登場以降, CNN による画像認識の精度は人間に匹敵するレベルを達成しており, 畳み込みニューラルネットワークはあらゆる分野における画像認識タスクの基本的なソリューションになりつつある。また, CNN の構造に関する研究も盛んに行われており, バッチ正規化 (Batch Normalization) [7] や残差接続 (Residual Connection) [8] など, CNN の学習安定化や精度改善などに有効な多くのテクニックが確立された。

2.1.1 ニューラルネットワーク

ニューラルネットワークは人間の神経回路網を模倣したアルゴリズムである。ニューラルネットワークを構成する最小の要素はユニットといい, ユニットは式 2.1 で得られた値 u をある関数 $z = f(u)$ に与えて値 z を出力する。

$$u = w_1x_1 + w_2x_2 + \dots + w_nx_n + b \quad (2.1)$$

ただし, $x_1 \dots x_n$ を入力とし, 各入力に対して掛ける値 $w_1 \dots w_n$ を重み, b をバイアスという。また, 関数 $f(u)$ は活性化関数と呼ばれる。式 2.1 を重回帰分析の視点で見れば, $x_1 \dots x_n$ を説明変数, $w_1 \dots w_n$ および b を偏回帰係数, u を目的変数ととらえることもできる。

最も簡単なニューラルネットワークは単純パーセプトロン [9] である。単純パーセプトロンはユニットを 1 つだけもつニューラルネットワークであり, 出力値 z があらかじめ設定したしきい値以上であれば 1, しきい値未満であれば 0 を出力する。単純パーセプトロンは線形回帰モデルであるため, 入出力間の非線形な関係を解くことができないが, 複数のユニットと層で構成されるニューラルネットワークは入出力間のより複雑な関係を解くことができる。たとえば, 5 層のニューラルネットワークは図 2.1 のように表される。第 1 層は入力層, 第 2-4 層は中間層, 第 5 層は出力層といい, 各層は任意の数のユニットをもつ。各ユニットの入出力は, 入力を $i = 1, \dots, I$, ユニット

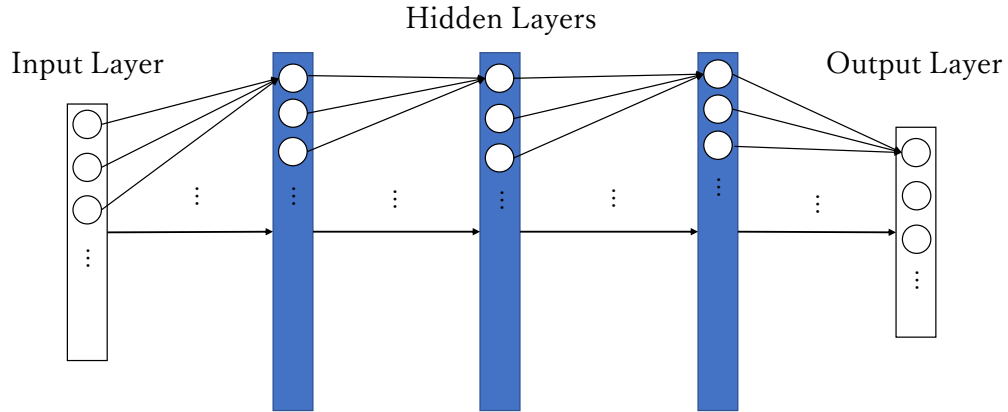


図 2.1 5層のニューラルネットワーク

を $j = 1, \dots, J$ とすると, 式 2.2 に従い計算される.

$$\begin{aligned}
 u_j &= \sum_{i=1}^I w_{ji}x_i + b_j \\
 z_j &= f(u_j)
 \end{aligned}
 \tag{2.2}$$

また, 各層を $1, \dots, L$ とし, 式 2.3 に従いベクトルと行列で表すと式 2.4 のようになる.

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_J \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_I \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_J \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_J \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & \cdots & w_{1I} \\ \vdots & \ddots & \vdots \\ w_{J1} & \cdots & w_{JI} \end{bmatrix}, \quad \mathbf{f}(\mathbf{u}) = \begin{bmatrix} f(u_1) \\ \vdots \\ f(u_J) \end{bmatrix}$$

$$\begin{aligned}
 \mathbf{u}^{(l+1)} &= \mathbf{W}^{(l+1)}\mathbf{z}^{(l)} + \mathbf{b}^{(l+1)} \\
 \mathbf{z}^{(l+1)} &= \mathbf{f}(\mathbf{u}^{(l+1)})
 \end{aligned}
 \tag{2.4}$$

ニューラルネットワークの出力は重み \mathbf{W} とバイアス \mathbf{b} によって決まる. これらをニューラルネットワークのパラメータといい, 入力に対して適切な出力を得られるようにパラメータを調節することをニューラルネットワークの学習という.

活性化関数には Rectified liner unit (ReLU) やロジスティックシグモイド関数などがある.

ReLU は式 2.5, ロジスティックシグモイド関数は式 2.6 で定義される.

$$f(u) = \max(u, 0) \quad (2.5)$$

$$f(u) = \frac{1}{1 + e^{-u}} \quad (2.6)$$

ReLU は勾配消失を防ぎ, 計算処理を高速化する効果をもつ. ロジスティックシグモイド関数は 0 から 1 の範囲で値を出力するため, 二値分類を行うニューラルネットワークの出力層の活性化関数などに用いられる.

2.1.2 損失関数

ニューラルネットワークの学習では入力と目標出力のペア (サンプル) を与え, 入力から計算した実際の出力と与えた目標出力との誤差を計算してニューラルネットワークに反映 (誤差逆伝播) することでパラメータを調節する. 誤差を計算するために用いる関数 $E(\mathbf{w})$ を損失関数という. 損失関数は解く問題によって適切に設計する必要がある. たとえば, 回帰の問題では式 2.7 のような二乗誤差に基づく損失関数が用いられる.

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{t}_n - \mathbf{y}(\mathbf{x}_n)\|^2 \quad (2.7)$$

ただし, \mathbf{w} はパラメータ, N はサンプルの個数, \mathbf{t}_n は目標出力, $\mathbf{y}(\mathbf{x}_n)$ は入力 \mathbf{x}_n に対するニューラルネットワークの出力とする.

二値分類では目標出力を 1 か 0 で表現し, 確率 (尤度) を出力するように出力層の設計を行う. たとえば, 入力画像に犬がいるか否かの分類を行う場合, 入力画像に犬がいれば 1, 犬がいなければ 0 として目標出力を与え, 出力が入力画像に犬がいる確率となるように設計する. 数式で表現すると, 式 2.8 のようになる.

$$\begin{aligned} \mathbf{t}_n &= \{0, 1\} \\ \mathbf{y}(\mathbf{x}_n) &= [0, 1] \end{aligned} \quad (2.8)$$

値域が $[0, 1]$ であるロジスティックシグモイド関数を出力層の活性化関数に選定し, 出力層のユニット数を 1 個とすると, 式 2.9 のような損失関数が用いられる.

$$E(\mathbf{w}) = - \sum_{n=1}^N [t_n \log y(\mathbf{x}_n) + (1 - t_n) \log \{1 - y(\mathbf{x}_n)\}] \quad (2.9)$$

式 2.9 は二値交差エントロピー誤差と呼ばれており, すべてのサンプルに対する \mathbf{w} の尤度の対数を取り符号を反転した式である. 尤度は式 2.10 で表される.

$$L(\mathbf{w}) = \prod_{n=1}^N \{y(\mathbf{x}_n)\}^{t_n} \{1 - y(\mathbf{x}_n)\}^{1-t_n} \quad (2.10)$$

二値交差エントロピー誤差は尤度の対数をとることでパラメータ更新に必要な微分計算を簡略化し, 符号を反転することで尤度の最大化問題を損失関数の最小化問題に変換している.

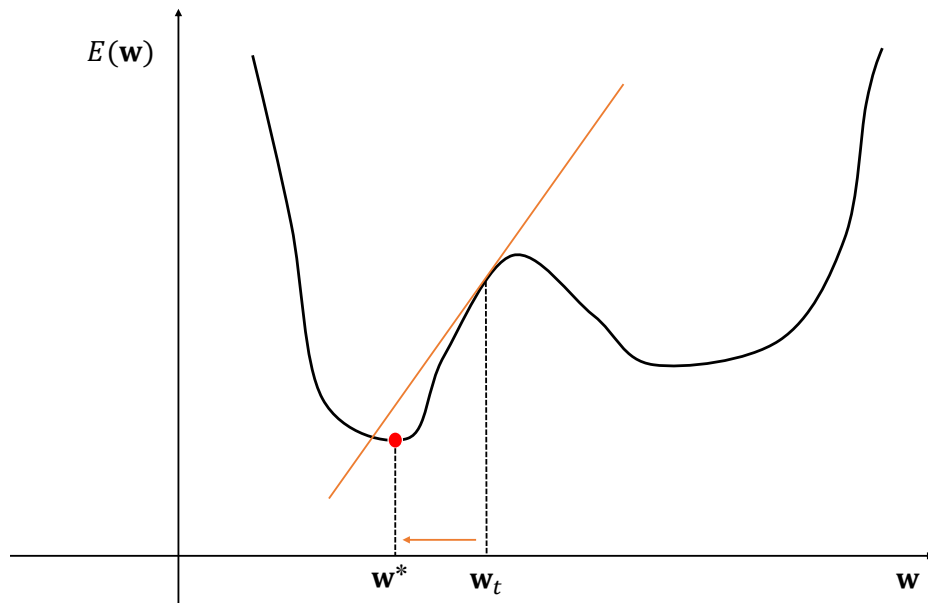


図 2.2 損失関数の最小化のイメージ. ある \mathbf{w}_t における勾配を計算し, 勾配とは逆方向に \mathbf{w}_t を繰り返し更新することで, 最終的に最適解 \mathbf{w}^* を得る.

マルチラベル分類は複数のラベルに対する予測を行うタスクである. たとえば, 入力画像に犬または猫がいるか否かの分類を行う場合, 入力画像に犬と猫がいれば $(1, 1)$, 犬のみがいる場合は $(1, 0)$, 猫のみがいる場合は $(0, 1)$, どちらもいない場合は $(0, 0)$ を目標出力とし, 各ラベルごとに確率を出力するように設計する. 設計は二値分類と似ているが, 複数のラベルに対応した出力層と損失関数を用いる. ラベルの総数を K とすると, 出力層のユニット数は K 個必要であり, 損失関数は式 2.11 のようになる.

$$E(\mathbf{w}) = - \sum_{n=1}^N \sum_{k=1}^K [t_{nk} \log y_k(\mathbf{x}_n) + (1 - t_{nk}) \log \{1 - y_k(\mathbf{x}_n)\}] \quad (2.11)$$

2.1.3 最適化の手法

損失関数を最小化するパラメータ \mathbf{w}^* を求めるために, さまざまな最適化手法が考案されている. ニューラルネットワークにおいて基本となる最適化手法に勾配降下法 (Gradient Descent Method) がある [10, 11]. 勾配降下法では損失関数の \mathbf{w} による偏微分によって得られるベクトル ∇E を用いてパラメータを更新する. ∇E は勾配と呼ばれており, 式 2.12 で定義される.

$$\nabla E \equiv \frac{\partial E}{\partial \mathbf{w}} = \left[\frac{\partial E}{\partial w_1} \frac{\partial E}{\partial w_2} \cdots \right]^\top \quad (2.12)$$

勾配はある時点のパラメータ \mathbf{w}_t における損失関数の接線の傾きを表す (図 2.2). 現在のパラメー

タを \mathbf{w}_t , 更新後のパラメータを \mathbf{w}_{t+1} とすると, パラメータの更新式は式 2.13 のようになる.

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \epsilon \nabla E \quad (2.13)$$

ただし, ϵ はパラメータの更新量を決める定数であり, 学習率と呼ばれる. 学習率の値は学習の成否に大きく影響を与えるため, 学習率は適切に設定する必要がある.

勾配降下法ではすべてのサンプルに対する損失関数を計算しパラメータの更新を行うが, ミニバッチと呼ばれるいくつかのサンプルをまとめた集合ごとに損失関数の計算とパラメータ更新を行う手法をミニバッチ学習といい, ミニバッチを用いた勾配降下法を確率的勾配降下法 (Stochastic Gradient Descent, SGD) という. ミニバッチが含むサンプルの数 (ミニバッチのサイズ) を N_t , 各ミニバッチを D_t とすると, 各ミニバッチに対する損失関数は式 2.14 のように計算される.

$$E_t(\mathbf{w}) = \frac{1}{N_t} \sum_{n \in D_t} E_n(\mathbf{w}) \quad (2.14)$$

SGD には改良型としてモメンタム [12] を取り入れたものがある. モメンタムはパラメータを更新する際に前回のパラメータ更新量を加味する手法であり, 学習の収束性を改善する. 勾配を $\mathbf{g}_t \equiv \nabla E_t(\mathbf{w}_t)$ とすると, モメンタム SGD のパラメータ更新は式 2.15 に従って行う.

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \epsilon \mathbf{g}_t + \mu(\mathbf{w}_t - \mathbf{w}_{t-1}) \quad (2.15)$$

ただし, μ は前回のパラメータ更新量の影響度を制御するための定数である.

ほかにも, SGD をさらに改良した手法として RMSProp[13] がある. RMSProp はパラメータの更新幅を勾配の二乗和の移動平均で除算することで, 勾配の更新量が大きい成分では更新幅を小さく, 勾配の更新量が小さい成分では更新幅を大きくして学習の収束性を改善する. RMSProp ではパラメータ更新を式 2.16 に従って行う.

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \epsilon \frac{\mathbf{g}_t}{\sqrt{\gamma \mathbf{g}_{t-1}^2 + (1 - \gamma) \mathbf{g}_t^2 + \epsilon}} \quad (2.16)$$

モメンタムと RMSProp を組み合わせた Adam (Adaptive Moment) という手法もある [14]. Adam によるパラメータ更新は式 2.21 に従って行う.

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t \quad (2.17)$$

$$\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^2 \quad (2.18)$$

$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{(1 - \beta_1^t)} \quad (2.19)$$

$$\hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{(1 - \beta_2^t)} \quad (2.20)$$

$$\mathbf{w}_t = \mathbf{w}_{t-1} - \epsilon \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t + \epsilon}} \quad (2.21)$$

ただし, 式 2.19 と式 2.20 では \mathbf{m}_t および \mathbf{v}_t における期待値の偏差を打ち消すための補正を行っている. Adam は学習率による影響を受けにくく, 安定した学習結果を得ることができる.

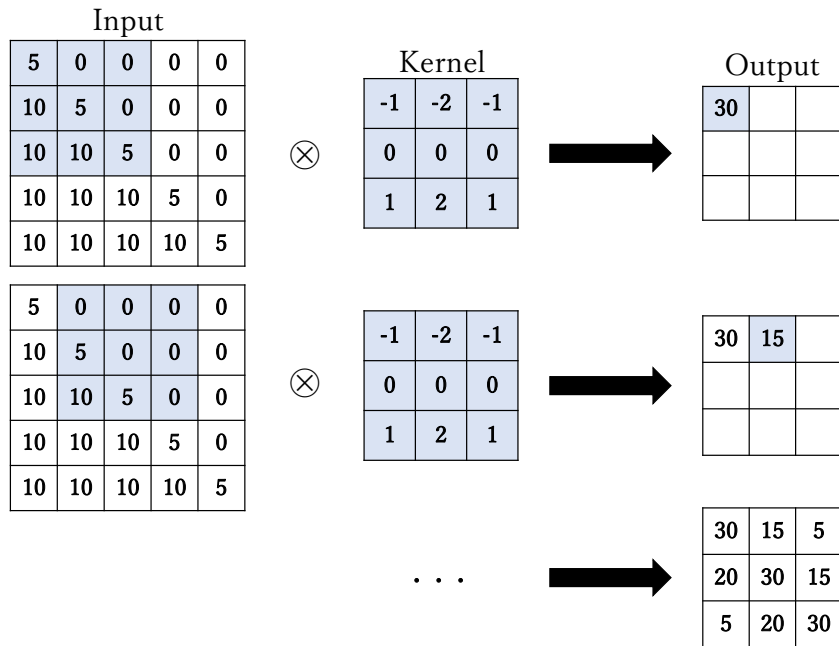


図 2.3 畳み込み計算の例. このカーネルは縦方向 Sobel フィルタ [15] と呼ばれており, 画像のエッジを強調する効果がある. この例では畳み込み計算の結果, 画素値の変化が大きい箇所で大い値を出力していることが確認できる.

2.1.4 畳み込み層

CNN はニューラルネットワークに畳み込みの構造を組み込んだものであり, この畳み込みの構造をもつ層を畳み込み層という. 画像処理における畳み込みはフィルタ (カーネル) と呼ばれる小さな画像を用いて図 2.3 のように行う. また, 畳み込みは入力サイズの $W \times H$, 各入力を $x_{ij} (i = 0, \dots, W - 1, j = 0, \dots, H - 1)$, カーネルのサイズを $W_f \times H_f$, カーネルの各値を $w_{pq} (p = 0, \dots, W_f - 1, q = 0, \dots, H_f - 1)$ とすると, 式 2.22 で定義される.

$$u_{ij} = \sum_{p=0}^{W_f-1} \sum_{q=0}^{H_f-1} x_{i+p, j+q} w_{pq} \quad (2.22)$$

図 2.3 では 1 画素ずつカーネルをずらして畳み込みを行っているが, 数画素ずつ行うことも可能である. このずらす間隔をストライド (stride) という. ストライドを s とすると, ストライドを考慮した畳み込み計算は式 2.23 のようになる.

$$u_{ij} = \sum_{p=0}^{W_f-1} \sum_{q=0}^{H_f-1} x_{si+p, sj+q} w_{pq} \quad (2.23)$$

このとき、出力サイズを $W_{out} \times H_{out}$ とすると、出力サイズは式 2.24 で計算できる。

$$\begin{aligned} W_{out} &= \lfloor \frac{W-1}{s} \rfloor + 1 \\ H_{out} &= \lfloor \frac{H-1}{s} \rfloor + 1 \end{aligned} \quad (2.24)$$

畳み込み層では 3 次元の入出力に対応した畳み込みの計算を行う。 $(l-1)$ 層から受け取った入力のサイズを $W \times H \times C$ 、 $(l-1)$ 層から受け取った各入力を $z_{ijc}^{(l-1)}$ ($c = 0, \dots, C-1$)、カーネルのサイズを $W_f \times H_f \times C$ 、カーネルの数を K 、カーネルの各値を w_{pqck} ($k = 0, \dots, K-1$) とすると、式 2.25 のように計算する。

$$u_{ijk} = \sum_{c=0}^{C-1} \sum_{p=0}^{W_f-1} \sum_{q=0}^{H_f-1} z_{i+p,j+q,c}^{(l-1)} w_{pqck} + b_k \quad (2.25)$$

ただし、 b_k はバイアスとする。 C および K はチャンネル数と呼ばれており、RGB 画像の入力であればチャンネル数は 3、グレースケール画像の入力であればチャンネル数は 1 となる。また、畳み込み層の出力チャンネル数は畳み込み層がもつカーネルの数と一致するため、畳み込み層の出力サイズは $W \times H \times K$ となる。こうして得られた u_{ijk} には活性化関数が適用され、畳み込み層の最終的な出力は $z_{ijk} = f(u_{ijk})$ となる。 w_{pqck} および b_k は CNN のパラメータであり、CNN の学習によって求められる。

畳み込み層では畳み込みのほかに、入力サイズと出力サイズを揃えるために特定の値で埋めたふちを追加するパディング処理を行う。0 埋めを行うゼロパディングが一般的に用いられている。

2.1.5 プーリング層

CNN の特徴のひとつに、入力画像内の特徴量に対する位置不変性がある。この位置不変性はプーリング層で行われるプーリング (Pooling) という処理に起因する。プーリングには、最大プーリング (Max Pooling) や平均プーリング (Average Pooling) などがある (図 2.4)。最大プーリングは領域内の最大値を取り出す処理であり、領域サイズを $W_p \times H_p$ 、領域 k に含まれる画素の集合を P_k ($k = 0, 1, \dots$) とすると、式 2.26 で表される。

$$u_{ijc} = \max_{(s,t) \in P_k} z_{stc} \quad (2.26)$$

また、平均プーリングは領域内の平均値を計算する処理であり、式 2.27 で表される。

$$u_{ijc} = \frac{1}{W_p H_p} \sum_{(s,t) \in P_k} z_{stc} \quad (2.27)$$

特に $W = W_p, H = H_p$ のときの平均プーリングを大域平均プーリング (Global Average Pooling, GAP) [16] といい、GAP を行う層を GAP 層という。GAP の出力サイズは $1 \times 1 \times C$ となるため、CNN から得た特徴量を 1 次元に変換するために用いられることが多い。

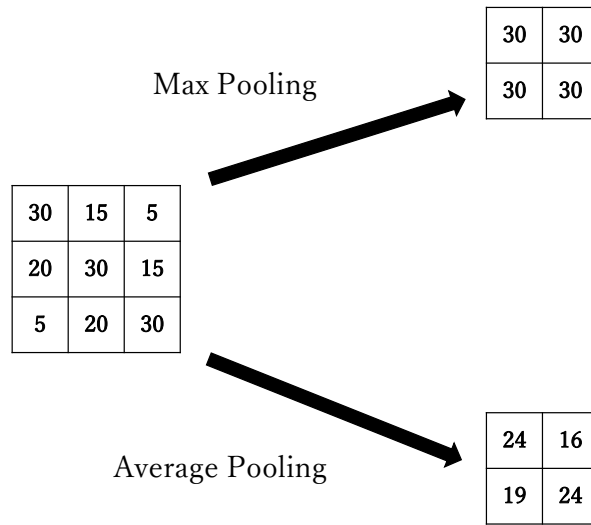


図 2.4 最大プーリングと平均プーリングの例. この例ではストライドを 1, 領域サイズを 2×2 としている.

2.1.6 バッチ正規化

バッチ正規化ではミニバッチ内の全サンプルにおけるチャンネルごとの平均と分散を用いて出力を正規化する. ミニバッチを B , ミニバッチ内のサンプルを $n (= 1, \dots, N)$, ミニバッチ内のサンプルのチャンネル c におけるユニットを $u_{ijc}^{(n)}$ とすると, バッチ正規化における計算は式 2.28 のように表すことができる.

$$\begin{aligned} \mu_{B,c} &= \frac{1}{NWH} \sum_{i,j,n} u_{ijc}^{(n)} \\ \sigma_{B,c}^2 &= \frac{1}{NWH} \sum_{i,j,n} (u_{ijc}^{(n)} - \mu_{B,c})^2 \\ \hat{u}_{ijc}^{(n)} &= \gamma_c \frac{u_{ijc}^{(n)} - \mu_{B,c}}{\sqrt{\sigma_{B,c}^2 + \varepsilon}} + \beta_c \end{aligned} \quad (2.28)$$

$\mu_{B,c}$ および $\sigma_{B,c}$ はそれぞれミニバッチ B のチャンネル c における全サンプルの平均および分散を表し, ε は分母が 0 になることを防ぐための極小の定数とする. γ_c および β_c は学習によって定める定数であり, これによりニューラルネットワークの非線形性を失わずに学習が可能となる.

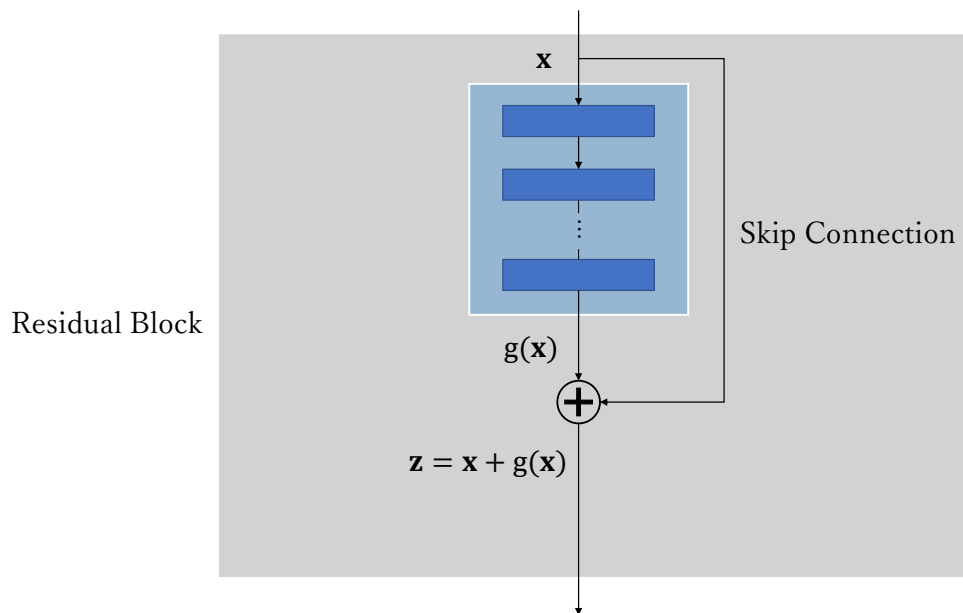


図 2.5 残差接続.

2.1.7 残差接続

多層のニューラルネットワークでは誤差逆伝播における勾配が重みの大小によって発散または消失 (0 になる) することがあり、学習がうまく進まなくなる現象、すなわち勾配消失問題が起こる。勾配消失を防ぐために、いくつかの層を迂回する道を設け、迂回せずに伝播してきた信号に加算する構造を残差接続という (図 2.5)。残差接続は勾配消失問題の対策として主要な手法のひとつであり、さまざまなニューラルネットワークの構造に取り入れられている。迂回路はスキップ接続 (Skip Connection) と呼ばれる。入力を x 、迂回される部分を $y = g(x)$ 、出力を z とすると、残差接続は式 2.29 のように表すことができる。

$$z = y + x = g(x) + x \quad (2.29)$$

$$g(x) = z - x \quad (2.30)$$

入力 x から出力 z までの構造を残差ブロック (Residual Block) といい、 $g(x)$ は残差ブロックの入出力間の残差を表現している。残差ブロックはスキップ接続により入力 x を次の層に伝えるため、勾配情報を失わずに信号を伝播することができ、より深い層のニューラルネットワークの学習が可能となる。

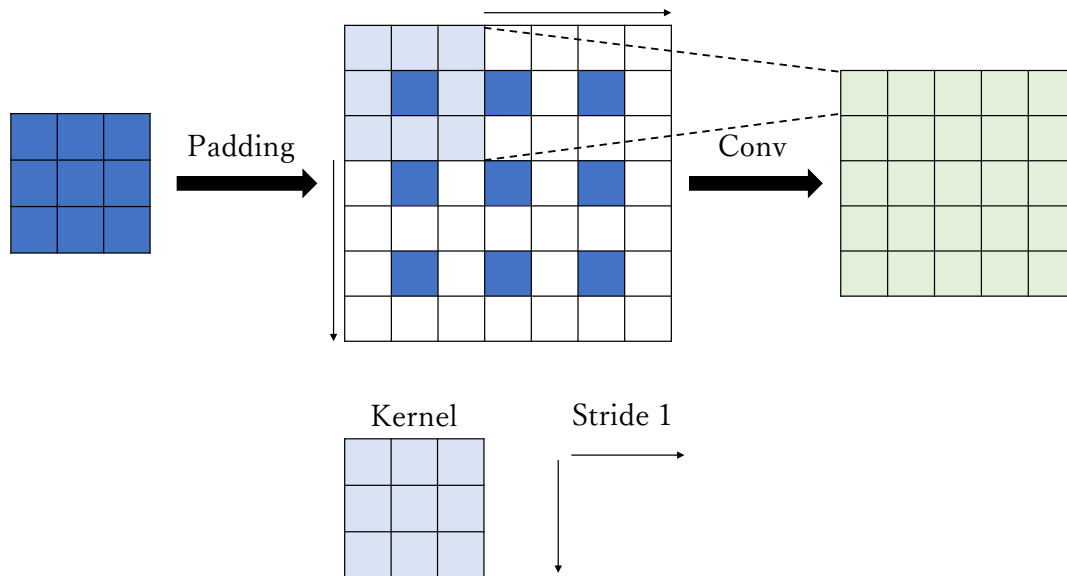


図 2.6 アップコンボリューション. この図では 3×3 のカーネルを用いて, 3×3 の特徴マップから 5×5 の特徴マップにアップコンボリューションを行っている.

2.1.8 アップコンボリューション

出力を画像とする問題（画像生成など）に対して用いられる手法にアップコンボリューションがある（図 2.6）. アップコンボリューションはアップサンプリングと畳み込みを順に行うことを指す. アップサンプリングを行うときの補完方法には最近傍補完やバイリニア補完などがある. アップコンボリューションを用いることで, 畳み込みとプーリングによって縮小した画像を再度拡大する構造を CNN に組み込むことができる.

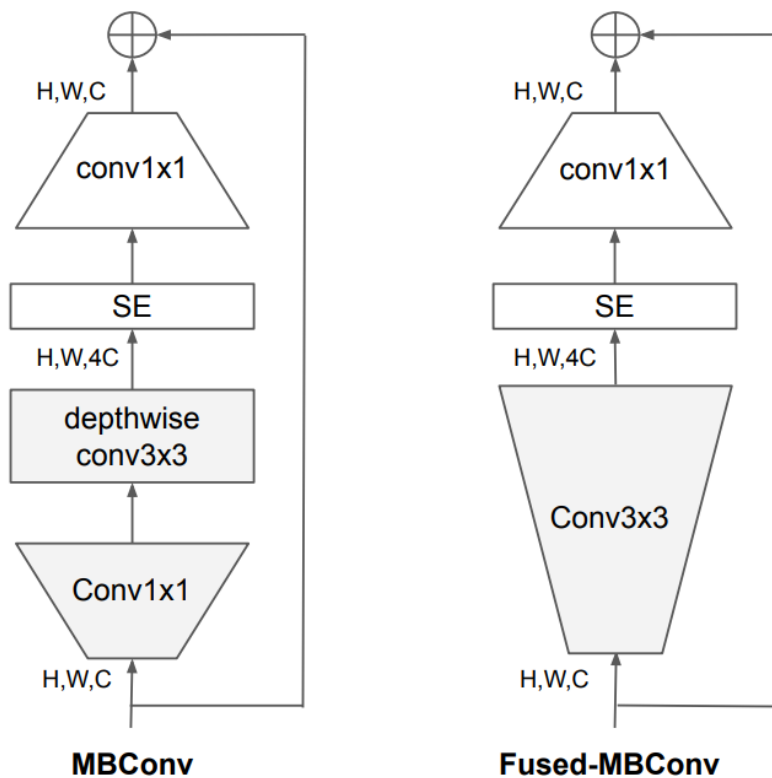


図 2.7 MBConv と Fused-MBConv のネットワーク構造 ([17] より引用). 台形はチャンネル数の縮小, 逆台形はチャンネル数の拡大を表す.

2.2 EfficientNetV2

EfficientNetV2[17] は学習パラメータの削減を目指した CNN であり, EfficientNet[18] の改良版である. 同時期に登場した Resnet-RS[19] のパラメータ数が約 1 億 6400 万, Vision Transformer[20] のパラメータ数が約 8600 万であるが, EfficientNetV2 はこれらの手法とほぼ同じ精度を保ちながら, パラメータ数を約 2400 万に抑えることに成功している.

2.2.1 ネットワーク

EfficientNetV2 は主に Mobile inverted Bottleneck Conv (MBConv) ブロック, Fused-MBConv ブロック [21] で構成される (図 2.7). MBConv ブロックは MobileNetV2[22] の Inverted Residual Block に SE モジュール [23] を追加したものである. SE モジュールでは, 入力の各チャンネルに対してチャンネルの重要度を計算し, 重み付けを行う (図 2.8). Fused-MBConv ブロックは MBConv ブロックにおける Pointwise 畳み込みと Depthwise 畳み込み [24] を 3×3 の畳み込みに置き換えたものである. MBConv ブロックによる計算の高速化, Fused-MBConv による高精度化, SE モジュールによる Attention[25] に似た特徴の重み付けなどが EfficientNetV2 の特徴で

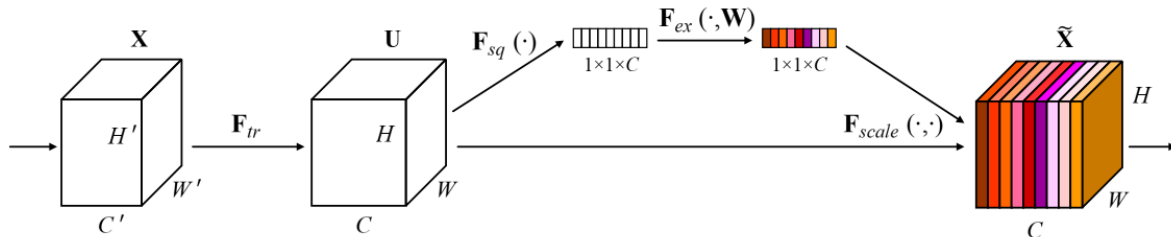


図 2.8 SE モジュールのネットワーク構造 ([23] より引用). $\mathbf{F}_{sq}(\cdot)$ は GAP に相当し, $\mathbf{F}_{ex}(\cdot, \mathbf{W})$ は次元削減を行う全結合および次元を元に戻す全結合を表す. 算出した各チャンネルの重要度を用いて, $\mathbf{F}_{scale}(\cdot, \cdot)$ で各チャンネルに重み付けを行う.

Stage	Operator	Stride	#Channels	#Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv4, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv6, k3x3, SE0.25	2	256	15
7	Conv1x1 & Pooling & FC	-	1280	1

図 2.9 EfficientNetV2-S のネットワーク構造 ([17] より引用). 活性化関数はすべて Swish 関数 [26] を用いている. MBConv x は MBConv ブロック内でチャンネル数を x 倍に拡大することを表し, SE x は SE モジュール内で x 倍の次元削減を行うことを表す.

ある.

EfficientNetV2 には S, M, L の 3 つのサイズのモデルがあり, S サイズのモデルは EfficientNetV2-S と呼ばれる (図 2.9). ネットワーク構造は Neural Architecture Search (NAS) によって探索することで決定している. EfficientNet と比較すると, 入力層近くは Fused-MBConv を採用し, 小さなカーネルおよび拡大率の MBConv ブロックおよび Fused-MBConv ブロックを用いて層の数を増やしている.

2.3 医用画像を用いた AI システム

2019 年 12 月初旬に発生した新型コロナウイルス感染症の拡大に伴い、AI を活用した新型コロナウイルス肺炎の診断支援システムの開発が活発に行われている。Erdaw ら [27] は Histogram of Gradient (HOG)[28] と Support Vector Machine (SVM)[29] を用いて胸部 X 線画像から新型コロナウイルス肺炎を検出することができる AI システムを開発した。従来手法よりも少ない計算量で高速かつ高精度な検出が可能であり、AI システムの評価では 97% を超える精度を達成している。AI の学習にはインターネット上の公開データベースから収集した胸部 X 線画像 1,100 枚を用いている。Zhao ら [30] は他分野のデータセットで事前学習した AI に胸部 CT 画像を用いて転移学習することで、胸部 CT 画像からの新型コロナウイルス肺炎の検出において 99.2% の精度を達成した。畳み込みニューラルネットワーク (CNN) を用いて特徴抽出と分類を同時に行うことで、エンドツーエンドな新型コロナウイルス肺炎の検出を実現した。

胸部の診断支援に関する研究開発では整備された大規模な医用画像データベースが診断支援システムの開発を支えている。胸部の医用画像データベースには National Institutes of Health (NIH) の NIH Chest X-ray Dataset[31] や China National Center for Bioinformation の China Consortium of Chest CT Image Investigation Dataset[32] などの公開データベースがあるため、AI の学習に必要なデータを比較的容易に収集することができる。しかし、パノラマ X 線画像のデータベースはインターネット上で公開されているものが少なく、特定のタスクを行うために作成されたデータベースが多いため、歯科用 AI に用いるパノラマ X 線画像のデータベースは目的に合わせて独自に作成しているケースが散見される。また、パノラマ X 線画像は多くの歯科所見を含んでいるにもかかわらず、これらの所見が付帯情報 (アノテーション) に反映されていない場合が往々にしてある。Panetta ら [33] はこの問題に対処すべく、2021 年 10 月に世界で初めてマルチモーダルなパノラマ X 線画像のデータベースである Tufts Dental Database (TDD) を公開した。TDD は 1,000 枚のパノラマ X 線画像と、各画像に対する歯と病変部のセグメンテーション情報、アノテーション作成者の関心領域およびアイトラッキングの情報、診断理由のテキストデータを含んでおり、歯科用 AI に応用できるデータベースとして有用である。TDD では各画像に対して 1 人の歯科放射線科医と 1 人の歯学部学生がアノテーションを作成している。

歯科データベースの研究開発と並行して、AI を用いた歯科画像診断支援システムの開発も活発に行われている。Chen ら [34] は 1,250 枚のアノテーション付き画像から学習した AI モデルと、歯科学に基づく独自の歯式修正アルゴリズムを用いて、デンタル X 線画像から歯の検出と歯式推定を行った。Tuzoff ら [35] は 1,574 枚のアノテーション付き画像から学習した 2 つの AI モデルを用いてパノラマ X 線画像から歯の検出と歯式推定を行った。いずれも専門医の診断精度に匹敵するほどの高い精度を達成し、歯科分野における AI の有効性を示した。AI を用いたう蝕の診断支援に関する研究開発も行われており、Lee ら [36] は小白歯と大白歯を対象に、3,000 枚のアノテーション付き画像から学習した AI モデルを用いて、デンタル X 線画像からう蝕か否かを診断するシステムを開発し、感度、特異度ともに 80% 以上を達成した。Zhu ら [37] はパノラマ X 線画像から

う蝕部位のセグメンテーションとう蝕の深度を診断する AI モデル CariesNet を開発した。1,159 枚のアノテーション付き画像から学習した CariesNet は 93% の精度を達成した。ほかにも、インプラントの検出やプラークの検出など、さまざまな画像診断支援システムの研究開発で AI が用いられている。

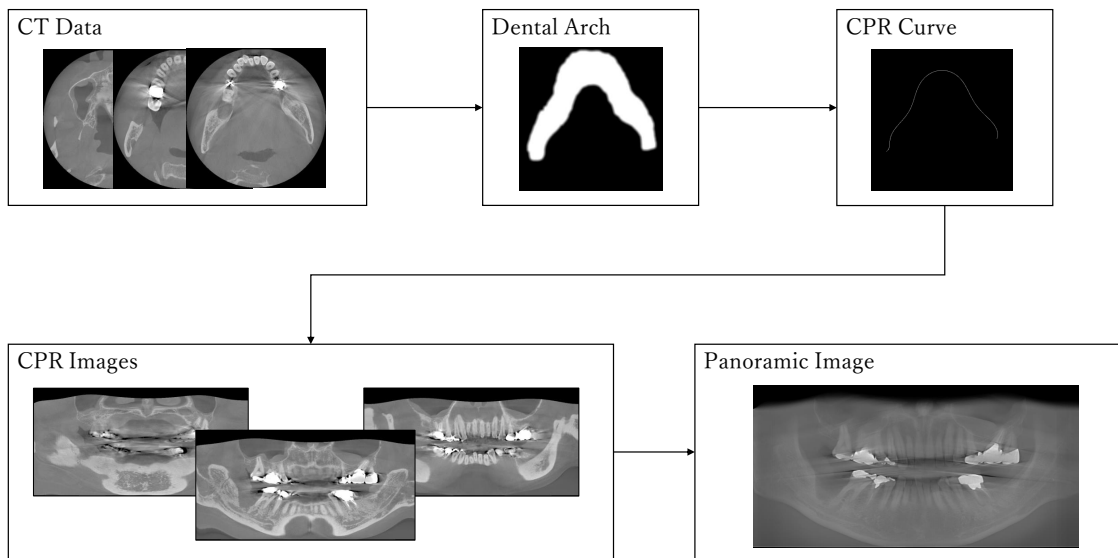


図 2.10 CT 画像からのパノラマ再構成のフロー．CT 画像から歯列弓を抽出し，CPR 曲線を作成と CPR 断面の生成および合成を行う．

2.4 CT 画像からのパノラマ再構成

CT 画像は 3D 画像であるが，顎骨と歯列の中心を通る曲線に沿った断面の画像を展開することで顎骨と歯列を平面上に再構成した 2D 画像を得ることができる．この曲線を Curved Planar Reconstruction 曲線（CPR 曲線）といい，得られる 2D 画像を CPR 画像という．CPR 曲線および CPR 曲線を中心とした並行な複数の曲線が形成する領域は，パノラマ撮影における断層領域として見做すことができる．したがって，生成した曲線に基づく CPR 画像を重ねるとパノラマ X 線画像と同様の 2D 画像が得られる．パノラマ再構成は主に CT 画像から歯列全体を観察するために用いられる．

CT 画像からのパノラマ再構成の基本的なフローを図 2.10 に示す．Luo ら [38] は CT 画像から最大値投影法（Maximum Intensity Projection, MIP）を用いて軸位断面の合成画像を生成し，合成画像に対して二値化と K-means++ クラスタリング [39, 40] を用いて歯列弓の自動セグメンテーションを行うことで，CT 画像からパノラマ X 線画像の自動生成を行った．CPR 曲線は歯列弓に対して 3 次スプライン曲線のカーブフィッティングを行うことで得ることができる [41] が，Luo らは最小二乗法を用いた放物線のカーブフィッティングを用いて，3 次スプライン曲線のカーブフィッティングに必要な制御点も自動的に決定している．Papakosta ら [42] は学習用 CT データセットから生成したテンプレート曲線と各スライス生成曲線に対して二乗平均平方根誤差を用いて比較し，もっとも誤差が小さいスライスを目的の歯列弓を含むスライスとしている．また，歯列

弓の抽出に膨張縮小処理と細線化処理を用いており、二値化のしきい値はヒストグラムのガウスフィッティングの結果に基づいて自動的に決定している。Yun ら [43] は冠状断面の MIP 合成画像を用いた歯列弓の選択や、ガウシアンフィルタを用いた細線化処理におけるひげの対策などを提案している。CPR 画像の合成には不必要な組織を取り除くためのパラメータを取り入れた非線形な関数を用いており、生成したパノラマ X 線画像の画質改善のためにガウシアンフィルタを用いた高周波強調処理も取り入れている。Amorim ら [44] は CPR 曲線の生成にベジェ曲線を用いており、平均値投影法 (Average Intensity Projection, AIP) [45] による CPR 画像の合成を行っている。

第 3 章

カルテ入力支援システムの構築

3.1 パノラマ X 線画像データベース

パノラマ X 線画像は機種によって画像特性が異なるため、さまざまな機種で撮影されたパノラマ X 画像をデータベースに用いることが望ましい。本研究では共同研究先の歯科クリニックから Panoura 15D と CareStream 9600 で撮影されたパノラマ X 線画像を収集し、共同研究先の医療機器メーカーからは Panoura A1 で撮影されたパノラマ X 線画像を収集した。ただし Panoura 15D と CareStream 9600 で撮影されたパノラマ X 線画像は AI モデルの学習用に使用し、これらの画像と各画像に付与したアノテーションのセットを学習用パノラマ X 線画像データベースと呼ぶ。また、Panoura A1 で収集されたパノラマ X 線画像はカルテ入力支援システムのテスト用に使用し、これらの画像と各画像に付与したアノテーションのセットをテスト用パノラマ X 線画像データベースと呼ぶ。すべてのパノラマ X 線画像は Digital Imaging and Communications in Medicine (DICOM) 画像*¹である。DICOM は医用画像の国際標準規格であり、医療機器間におけるデータ通信などにこの方式が用いられる。表 3.1 に DICOM 画像がもつ情報の抜粋を示す。DICOM 画像には患者の個人情報や撮影機器情報、画像の保存方式などの情報が含まれているが、収集したパノラマ X 線画像の個人情報はあらかじめ削除してある。

アノテーションは各歯に対する歯の輪郭、FDI 方式に基づく歯式、歯の状態で構成される。歯の輪郭は輪郭を形成する点の XY 座標集合である。歯の状態は状態ラベルの定義に基づいたラベルで表す。状態ラベルの定義を表 3.2 に示す。状態ラベルの付与では、各歯に対して大グループ (Major Group) のラベル付与を基本とし、パノラマ X 線画像から小グループ (Minor Group) まで診断可能な歯に対しては小グループのラベルも付与する。また、複数の状態をもつ歯に対しては複数の状態ラベルを付与する。たとえば、クラウンと根管充填処置が認められた歯に対しては、FMC と RCF を付与する必要がある。

アノテーションの付与は、共同研究先の歯科クリニックに所属する歯科医 3 名と、徳島大学歯科放射線科の専門医 1 名が行った。なお、3 名の歯科医は 5 年以上の臨床経験を有している。また、

*¹ <https://www.dicomstandard.org/>

表 3.1 DICOM タグと格納値の抜粋. DICOM タグは 16 進数表記の 4 バイトで表す. また, 最初の 2 バイトはグループ番号を表す.

Tag	Attribute Name	Discription
(0008, 0070)	Manufacturer	製造者
(0008, 1090)	Manufacturer's Model Name	撮影機種名
(0008, 2111)	Derivation Descripthon	画像圧縮に関する情報
(0010, 0010)	Patient's Name	患者名
(0010, 0020)	Patient's ID	患者 ID
(0010, 0030)	Patient's Birth Date	患者の生年月日
(0010, 0040)	Patient's Sex	患者の性別
(0010, 1010)	Patient's Age	患者の年齢
(0028, 0002)	Samples per Pixel	Gray Scale:1, RGB:3
(0028, 0004)	Photometric Interpretation	画素の解釈
(0028, 0010)	Rows	画像の横幅
(0028, 0011)	Columns	画像の高さ
(0028, 0030)	Pixel Spacing	画素間隔 (mm)
(0028, 0100)	Bits Allocated	割当ビット
(0028, 0101)	Bits Stored	格納ビット
(0028, 0102)	Hight Bit	高位ビット
(0028, 0103)	Pixel Representation	unsigned int:0, int:1
(0028, 1050)	Window Center	ウィンドウ中心
(0028, 1051)	Window Width	ウィンドウ幅
(0028, 1052)	Rescale Intercept	ピクセルデータ変換の切片
(0028, 1053)	Rescale Slope	ピクセルデータ変換の勾配
(7fe0, 0010)	Pixel Data	ピクセルデータ

本研究におけるデータとはパノラマ X 線画像とアノテーションのセットを指す.

表 3.2 状態ラベルの定義. 歯科医や放射線科の専門医の監修を受けて定義した.

Major Group	Label	Minor Group	Label
異常なし	N		
う蝕	C	う蝕第 2 度	C2
		う蝕第 3 度	C3
		う蝕第 4 度	C4
根尖病巣	Per		
不適合	Futeki		
脱離	Daturi		
破折	Frft		
分岐部病変	Bunki		
捻転歯	ROT		
エナメル系形質不全	EHP		
根管充填	RCF		
インプラント	Implant		
インレー	In	メタル	Metal
		複合レジン	CR
		ジルコニア	Zi
		セラミック	Cera
		アマルガムフィリング	AF
		ジーセム	gcem
クラウン	Cro	全部金属冠	FMC
		ポンティック	Pon
		五分の四冠	4/5Cro
		四分の三冠	3/4Cro
		レジン前装金属冠	VMC
		メタルボンド	MB
		ジルコニア	Zi
		セラミック	Cera
		CAD/CAM 冠	CAD
		レジンジャケット冠	RJC
		硬質レジンジャケット冠	HJC
		埋伏	IT
埋伏歯	RT		
完全埋伏歯	CRT		
水平埋伏智歯	HIT		
水平智歯	HET		

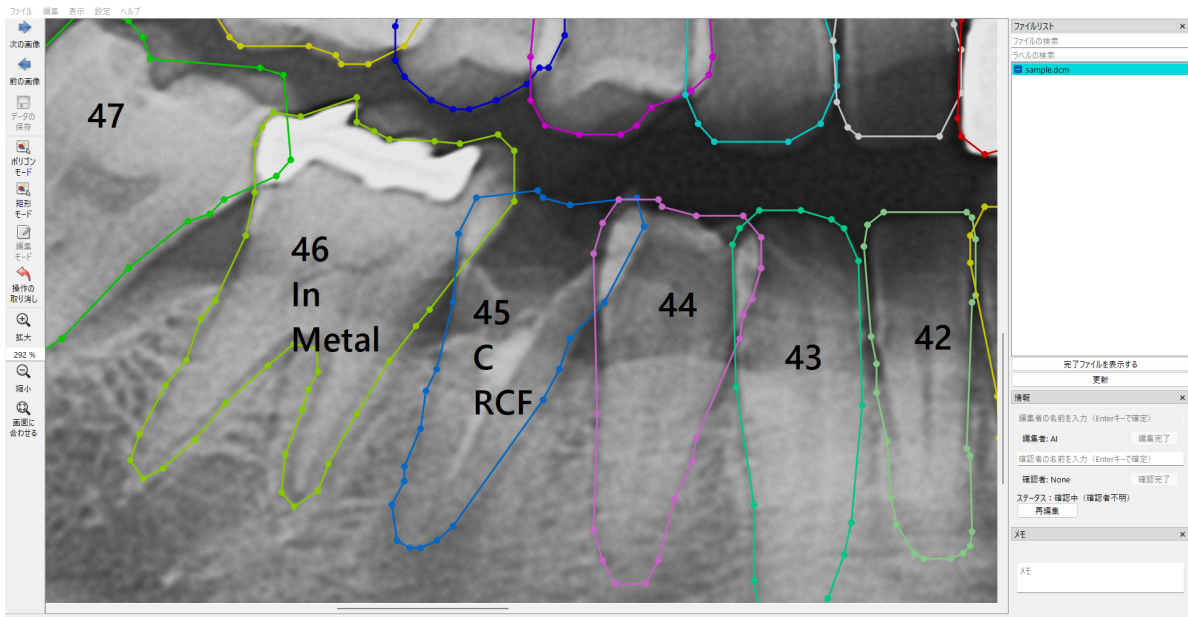


図 3.1 Anotee. 中央にパノラマ X 線画像とポリゴンの表示があり、各歯に対するアノテーションを一目で確認することができる。右側にはワンクリックでデータにアクセス可能なファイルリストと、編集者および確認者を登録するペイン、ノートペインがある。

3.1.1 アノテーションソフトウェア Anotee

効率的なアノテーション付与のために、我々は DICOM ファイルの読み込みや自動アノテーションなどに対応したアノテーションソフトウェア Anotee (Anotate Teeth, アノティー) を開発した (図 3.1)。本研究で用いるアノテーションファイルはすべて Anotee を用いて作成し、管理や確認修正作業も Anotee の機能を活用した。Anotee は Labelme[46] をベースに Python^{*2}と PyQt^{*3}を用いて開発した。共同研究先の歯科クリニックや医療機器メーカーが主な利用環境であるため、Anotee は Windows10 および Windows11 で動作可能な Executable (EXE) ファイルとして提供している。Anotee には AI を用いた機能があるため、Anotee の運用は 8GB 以上のランダムアクセスメモリ (RAM) を搭載した計算機を使用することを奨励している。

Anotee は DICOM 画像の読み込みに対応している。読み込みには pydicom^{*4}を用いており、圧縮画像から非圧縮画像への変換には GDCM^{*5}を用いている。適切な画像表示のために、DICOM タグの情報を用いた基本的な画像変換処理も行っている。Photometric Interpretation が MONOCHROME2 (最小ピクセル値が黒、最大ピクセル値が白を表す) の場合は、Rescale Slope

*2 <https://www.python.org/>

*3 <https://www.qt.io/ja-jp/>

*4 <https://pydicom.github.io/>

*5 <http://gdcm.sourceforge.net/>

および Rescale Intercept を用いて式 3.1 に従い Pixel Data の変換を行う。

$$Output = RescaleSlope \times PixelData + RescaleIntercept \quad (3.1)$$

また、Photometric Interpretation が MONOCHROME1 (最小ピクセル値が白, 最大ピクセル値が黒を表す) 場合は, 式 3.2 に従い白黒を反転させる。

$$Output = max(PixelData) - PixelData \quad (3.2)$$

ピクセルデータの最大値は DCIOM 画像によって異なるため, 通常 Window Center および Window Width を用いて DICOM ビューアに適したピクセルデータのマッピングを行い画像を描画するが, Anotee では画像のビット深度を 16 ビット (符号なし整数) に統一して表示するため, 式 3.3 に従ってピクセルデータをマッピングして表示する。

$$DisplayPixelData = \frac{PixelData}{max(PixelData)} \times 65535 \quad (3.3)$$

Anotee では上述のような画像処理に加えてマルチ周波数処理 [47] も行う。マルチ周波数処理は歯のエッジを強調したり, アノテーション付与に必要なコントラスト調節の手間を軽減できるため, 読影やアノテーションの補助に効果的である。マルチ周波数処理は周波数強調とダイナミックレンジ圧縮の 2 つの機能をもつ。周波数強調およびダイナミックレンジ圧縮のアルゴリズムを図 3.2 および図 3.3 に示す。周波数強調では, まず元画像に対して何度か平滑化を繰り返し, 平滑化前の画像と平滑化後の画像の差分画像をそれぞれ生成する。次に, 生成した差分画像に係数を乗算して和をとり, 強調成分画像を得る。最後に, 強調成分画像に係数を乗算し, 元の画像に加算することで周波数強調画像を得る。Anotee では平滑化の回数を 8 回とし, 平滑化に用いるバイラテラルフィルタのフィルターサイズを 9×9 , $\sigma = 75$ とした。また, 各差分画像に乗算する係数を 1, 強調成分画像に乗算する係数を 5 とした。ダイナミックレンジ圧縮では, まず周波数強調で得られた強調成分画像を元画像から減算し, エッジ保存画像を得る。次に, エッジ保存画像に平滑化と濃度変換を行いマスク画像を得る。最後に, マスク画像と元画像を加算することでダイナミックレンジ圧縮画像を得る。Anotee ではエッジ保存画像の平滑化にフィルターサイズ 5×5 の平均化フィルタを用いた。ピクセル値を x , ピクセル値の平均を \bar{x} , 変換後のピクセル値を y とすると, ダイナミックレンジ圧縮における濃度変換は式 3.4 に従う。ただし, 定数 $grad$ はハイパーパラメータであり, Anotee では $grad = -0.2$ をデフォルト値としている。式 3.4 による濃度変換のイメージは図 3.4 のようになる。マスク画像には平均ピクセル値から離れているピクセル値を抑制する効果があり, 画像全体の濃淡の変化を小さくする (圧縮する) ことができる。通常のコントラスト調節と違う点はエッジ保存画像を用いている点であり, エッジが鮮明な状態を維持したまま濃淡圧縮が可能である。Anotee では $grad$ の値を調節することでコントラスト調節を行う。

$$y = \begin{cases} grad \times (x - \bar{x}), & x \geq \bar{x} \\ -grad \times (\bar{x} - x), & x < \bar{x} \end{cases} \quad (3.4)$$

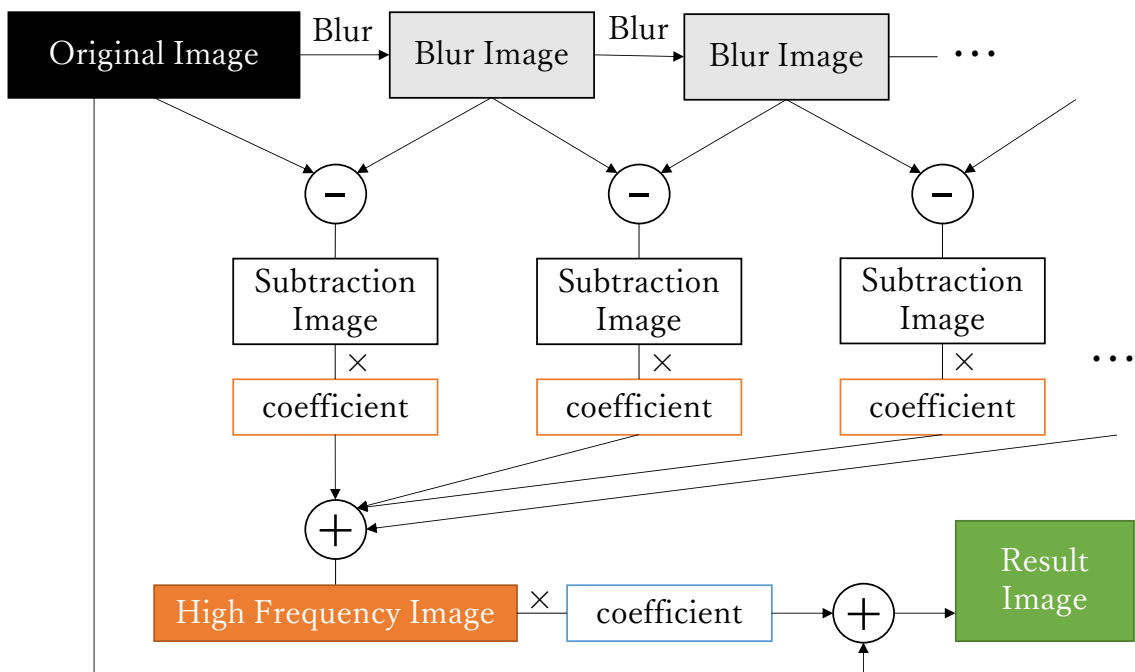


図 3.2 マルチ周波数処理における周波数強調のアルゴリズム. 平滑化画像 (Blur Image) の数と各係数 (coefficient) はハイパーパラメータである.

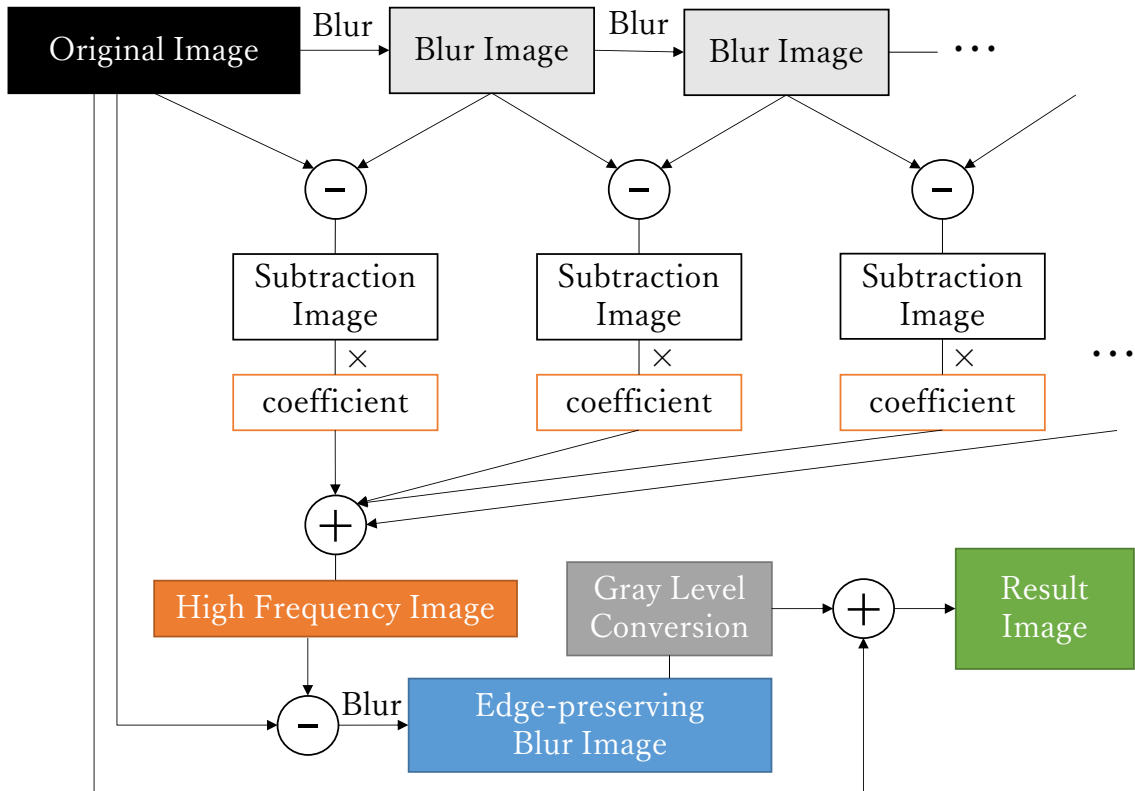


図 3.3 マルチ周波数処理におけるダイナミックレンジ圧縮のアルゴリズム．強調成分画像（High Frequency Image）の生成までは周波数強調と同様であるため，実際は周波数処理の結果を利用してダイナミックレンジ圧縮を行う．

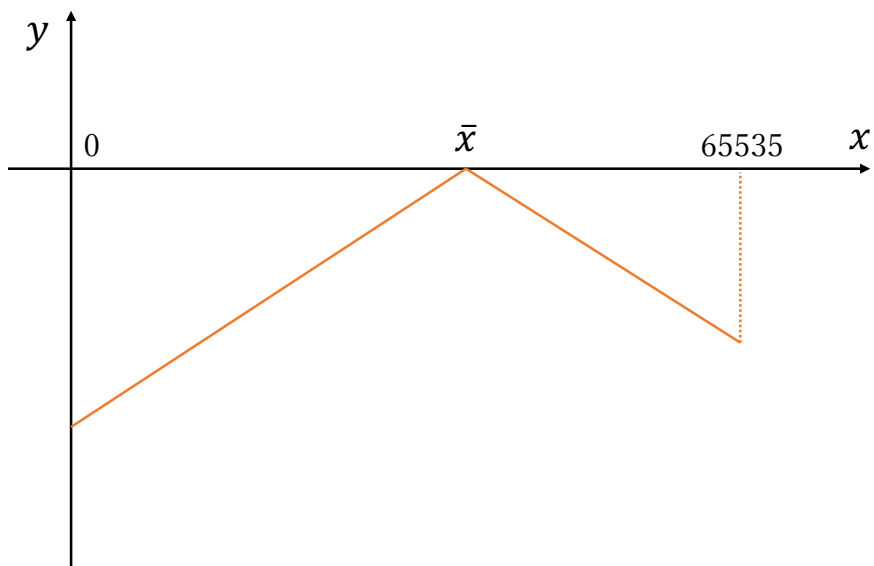


図 3.4 ダイナミックレンジ圧縮における濃度変換のイメージ

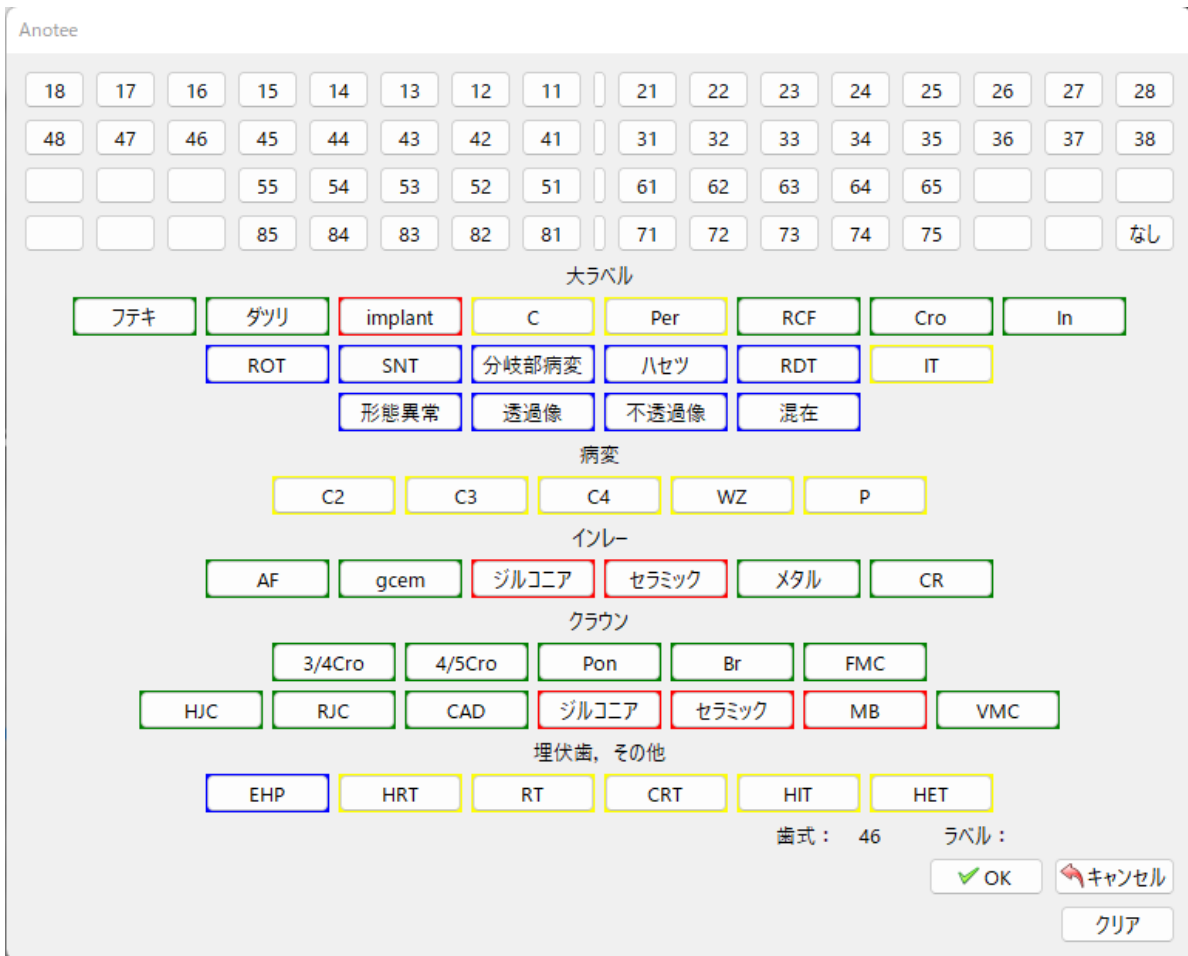


図 3.5 Anotee のラベルダイアログ。頻出する状態ラベルのボタンを右側に配置することでラベル付けの効率化を図っている。

アノテーション付与では歯や病変部位などに対してポリゴン（多角形）を作成し、ラベルを付与することができる。たとえば、上顎右側第一大臼歯（26 番）の歯にメタルインレーの処置が認められた場合、26 番の歯の輪郭に沿って作成したポリゴンに対して“26”，“In”，“Metal” のラベルを付与する。ポリゴン作成にはポリゴンモードと矩形モードがあり、ポリゴンを作成する必要がないときは矩形による領域作成も可能である。ポリゴンモードではマウスのクリックで点を打ち、始点と終点を繋ぐことで 1 つのポリゴンを作成する。作成したポリゴンに対するラベル付けは図 3.5 に示すようなラベルダイアログを開いて行う。ラベルダイアログにはあらかじめ用意した歯式と状態ラベルのボタンを用意しているため、ボタンを押すだけでラベル付けを行うことができる。また、病変部位（貯留嚢胞など）や歯式を確定できない歯（過剰歯など）に対して“歯式なし”のラベルを登録することも可能である。ラベルダイアログで登録できる内容以外に特筆すべき所見がある場合は、ノートペイン上で記述および保存することができる。

アノテーションは JavaScript Object Notation (JSON) 形式で保存する。アノテーションファイルの構成を表 3.3 に示す。アノテーションファイルには歯の情報に加えて、Anotee で用いる画

表 3.3 アノテーションファイルの構成. info にはアノテーションファイルを作成または確認した各機関の情報が格納される. また, shapes には各歯のラベル, 形状の種類, ポリゴンを形成する各点の XY 座標集合が格納される. 形状の種類は矩形 (Rect) またはポリゴン (Polygon) の 2 種類だが, 本研究ではポリゴンのみを扱う.

Key1	Key2	Value
version		Anotee Version
filename		Filename of Image
height		Height of Image
width		Width of Image
brightness		Value of Brightness
contrast		Value of Contrast
note		Plain Text Note
info	org	Organization
	state	File State
	editor	Editor Name
	checker	Checker Name
shapes	label	Tooth Number & Conditions
	shape_type	Type of Shape
	points	Coordinates of Contour

像のメタ情とファイル管理用データが格納される. ファイル管理用データには機関とファイル状態, 編集者と確認者の情報を記録する. ファイル管理情報の登録は Anotee の情報ペインで行うことができる.

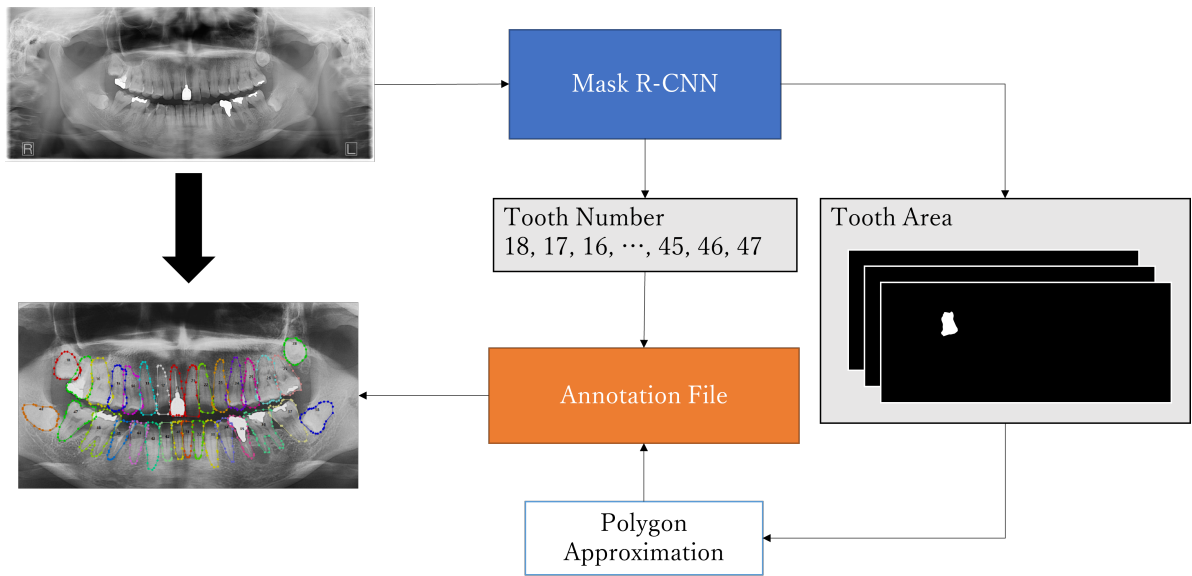


図 3.6 Mask R-CNN を用いた自動ポリゴン作成. Mask R-CNN モデルは各歯の矩形領域を構成する座標 (ymin, xmin, ymax, xmax) も出力するが, Anotee では用いない.

ポリゴン作成の負担を軽減するため, Mask R-CNN[48] を用いた簡易的な自動ポリゴン作成機能を搭載した. 自動ポリゴン作成のフローを図 3.6 に示す. この方法を用いて作成したアノテーションファイルでは編集者を“AI”として自動的に登録する. なお, GPU を搭載していない計算機でも利用することができる.

Mask R-CNN は物体検出とインスタンスセグメンテーションを同時に行うエンドツーエンドな深層ニューラルネットワークアーキテクチャである. Mask R-CNN の学習では Microsoft Common Objects in Context (MS COCO) データセット [49] を事前学習したモデルに対して, 収集したデータを用いて逐次的にファインチューニングを行うことで, 少しずつ精度を向上させた. ファインチューニングに関する学習テクニックに関して Hara ら [50] と同様の手法を用いた. ただし, 乳歯を含むデータは十分に収集できなかったため, 学習対象は永久歯のみの 32 種類に限定した. 画像の前処理では入力画像に対してヒストグラム平坦化と 3 チャネル 8 ビット画像への変換を行う. また, 入力画像のピクセル値を単精度浮動小数点数として扱い, 正規化処理として各ピクセル値からすべての入力画像の平均ピクセル値を減算する. さらに, 画像のランダム左右反転をデータ拡張として適用する. 実装は Abdulla の実装コード [51] に基づいており, ハイパーパラメータは出力クラス数を 33 クラス (歯式 32 種類と背景), デフォルトバウンディングボックスの縦横比設定を正方形と縦長の矩形のみに変更した. Mask R-CNN は各ピクセルに対する歯式の尤度を出力するため, しきい値を 0.5 とした二値化処理を行い, 8 ビット画像に変換する. その後, 各歯の輪郭に対してポリゴン近似を行い, 輪郭の端点をアノテーションファイルに保存する. ただし, 面積が 500 ピクセル未満の領域は歯の領域として扱わずに除外する. ポリゴン近似の手法は Suzuki ら [52] が提案したアルゴリズムに従い, 近似精度を決定するパラメータはデフォルトで 0.003 とした. なお, このパラメータは Anotee 上で調節することが可能である.

3.1.2 パノラマ X 線画像データベースの解析と整備

学習用パノラマ X 線画像データベースでは乳歯を含まない 1,781 枚のパノラマ X 線画像に対してアノテーションを付与した。AI モデルの学習に向けてデータの整備を行うため、学習用パノラマ X 線画像データベースに対して歯の状態に関する解析を行った。表 3.4 に状態ラベルの集計結果を示す。表 3.5 では集計結果に基づいて出現頻度の高い歯の状態から並べている。ただし、大グループに属する状態ラベル (C, In, Cro, IT) は小グループの状態ラベルの数も含んでいる。歯の状態の出現頻度には大きな偏りがあり、エナメル形質不全 (EHP) とジルコニアインレー (Zi(In)) に関しては 1 件も例がない。したがって、AI モデルの学習に十分な数がない、エナメル形質不全、ジルコニアインレー、G-CEM (gcem)、四分の三冠 (3/4Cro)、破折 (Frt)、捻転歯 (ROT)、五分の四冠 (4/5Cro)、分岐部病変 (Bunki)、セラミックインレー (Cera(In))、不適合 (Futeki) は本研究の診断対象から除外した。さらに、状態ラベルの選定について歯科医および放射線科の専門医と協議を重ね、パノラマ X 線画像では深度の判別が困難なう蝕第 2-4 度 (C2, C3, C4) はまとめてう蝕 (C) とし、診断が困難な脱離 (Daturi) は除外した。また、特徴が似ている CAD/CAM 冠 (CAD)、レジンジャケット冠 (RJC)、硬質レジンジャケット冠 (HJC) はまとめて CADCAM とした。最終的に、実際のカルテ入力を考慮した 20 状態と大ラベルによる分類を中心とした 9 状態の 2 つの分類パターンを定義した (表 3.6)。20 状態では補綴物の材質や形態などを考慮しているが、9 状態では補綴物の材質や形態、埋伏の深度を無視した状態ラベルを設定した。

また、テスト用パノラマ X 線画像データベースでは乳歯を含まない 215 枚のパノラマ X 線画像に対してアノテーションを付与した。テスト用パノラマ X 線画像データベースに対する状態ラベルの集計結果を表 3.7 に示す。ただし、カルテ入力支援システムの評価では 9 状態の診断が可能な状態診断モデルを用いてテストを行うため、9 状態に対する集計のみを行った。

表 3.4 学習用パノラマ X 線画像データベースにおける状態ラベルの内訳 (種類順)

Condition	Number
N	29,131
C	1,145
C2	567
C3	143
C4	421
Per	985
Futeki	82
Daturi	433
Frt	12
Bunki	33
ROT	28
EHP	0
RCF	6,197
Implant	349
In	8,285
Metal	5,316
CR	2,808
AF	181
Zi(In)	0
Cera(In)	34
gcem	1
Cro	6,872
FMC	3,640
3/4Cro	2
4/5Cro	32
MB	816
VMC	1,680
Zi(Cro)	198
Cera(Cro)	411
CAD	93
RJC	68
HJC	43
Pon	1,099
IT	2,889
HRT	167
RT	375
CRT	1,768
HIT	390
HET	191

表 3.5 学習用パノラマ X 線画像データベースにおける状態ラベルの内訳 (データ数順)

Condition	Number
N	29,131
In	8,285
Cro	6,872
RCF	6,197
Metal	5,316
FMC	3,640
IT	2,889
CR	2,808
CRT	1,768
VMC	1,680
C	1,145
Pon	1,099
MB	816
C2	567
Per	985
Daturi	433
C4	421
Cera(Cro)	411
HIT	390
RT	375
Implant	349
Zi(Cro)	198
HET	191
AF	181
HRT	167
C3	143
CAD	93
Futeki	82
RJC	68
HJC	43
Cera(In)	34
Bunki	33
4/5Cro	32
ROT	28
Frt	12
3/4Cro	2
gcem	1
EHP	0
Zi(In)	0

表 3.6 9 状態および 20 状態の定義. ポンティック (Pon) は大グループに該当しないが, クラウンの中でも歯冠部分のみを示す唯一の状態ラベルであるため, 大グループ相当の状態ラベルとして扱う. また, ポンティックの材質は問わないこととする.

9 Conditions	20 Conditions
N	N
C	C
Per	Per
RCF	RCF
Implant	Implant
In	Metal
	CR
	AF
Cro	FMC
	MB
	VMC
	CADCAM
	Zi
	Cera
Pon	Pon
IT	HRT
	RT
	CRT
	HIT
	HET

表 3.7 テスト用パノラマ X 線画像データベースにおける状態ラベルの内訳

Condition	Number
N	3,063
C	166
Per	230
RCF	939
Implant	167
In	900
Cro	1,245
Pon	196
IT	177

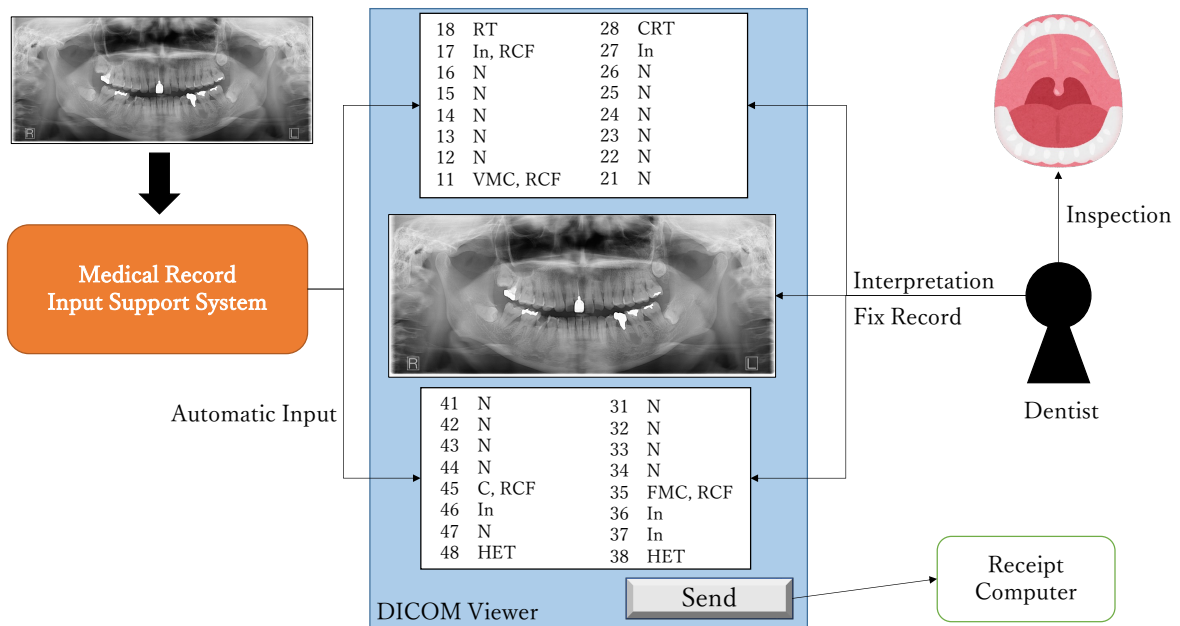


図 3.7 カルテ入力支援システムの運用イメージ. カルテ入力支援システムを既存の DICOM ビューアに統合し, カルテの自動入力を行う. 歯科医による読影 (Interpretation) と視診 (Inspection) を踏まえたカルテ修正の後, レセプトコンピュータへ送信する.

3.2 カルテ入力支援システム

カルテ入力支援システムの概要を図 3.7 および図 3.8 に示す. カルテ入力支援システムはパノラマ X 線画像から歯の検出, 歯式推定, 歯の輪郭推定, 歯の状態診断を行いカルテ入力の補助を行うシステムであり, 歯検出モデル, セグメンテーションモデル, 状態診断モデルの 3 つの AI モデルで構成される. まず, パノラマ X 線画像を入力とし, 歯検出モデルを用いて歯の検出と歯式推定を行う. 次に, 検出された各歯に対してセグメンテーションモデルを用いて歯の輪郭推定を行う. 最後に, 歯の輪郭推定の結果を利用して歯の領域を抽出し, 各歯に対して状態診断モデルを用いて歯の状態診断を行う. 各モデルの結果はファイルに保存し, DICOM ビューアなどで読み込むことでカルテに対する自動入力を行う.

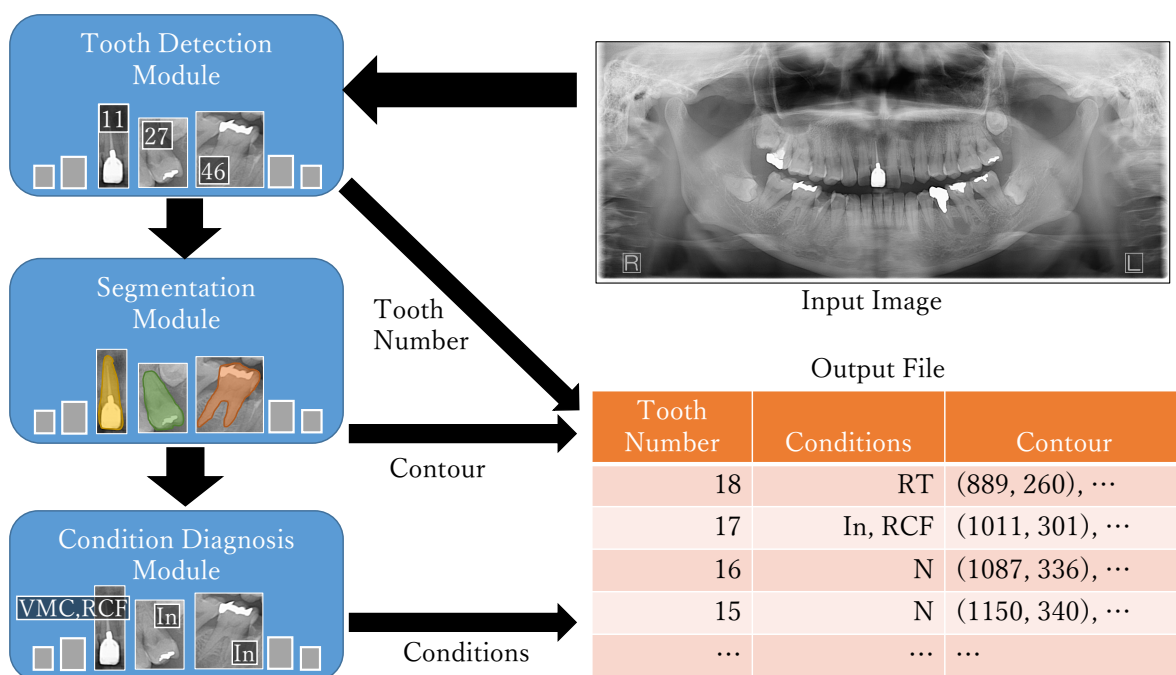


図 3.8 カルテ入力支援システムの概略図。カルテ入力支援システムの出力はアノテーションファイルと同様の形式で保存される。

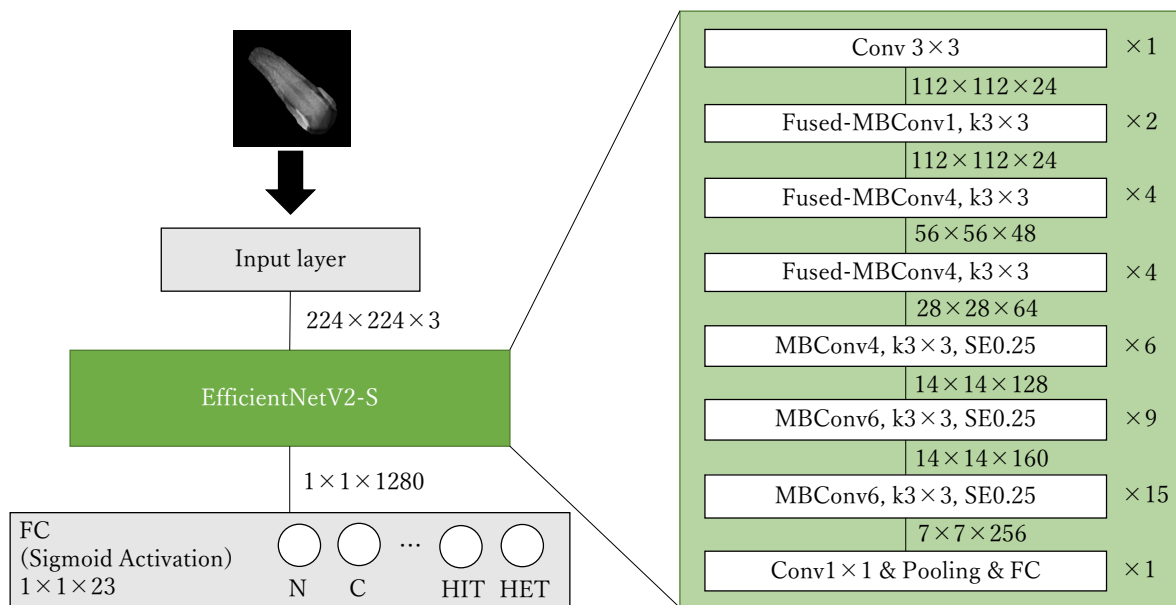


図 3.9 状態診断モデル. EfficientNetV2-S から出力される 1280 次元の特徴量を用いて分類器の作成を行う.

3.2.1 状態診断モデル

状態診断モデルでは EfficientNetV2-S[17] に基づいて構築した状態診断モデルを用いる. 状態診断モデルの概要を図 3.9 に示す. ImageNet-21K[53] を用いて学習した EfficientNetV2-S のモデルに対して, 学習用パノラマ X 線画像データベースを用いてファインチューニングを行う. 出力層のユニット数は学習する状態ラベルの種類数である 23 に設定し, マルチラベルの出力に対応するために, 出力層の活性化関数はシグモイド関数とする. 出力は各状態に対する尤度(確率)となるため, 出力の真値は対象の歯に該当する状態ラベルを 1, それ以外を 0 とした One-hot ベクトルとする. 損失関数は二値交差エントロピー誤差を使用する.

入力画像の前処理ではパノラマ X 線画像にヒストグラム平坦化の適用と 3 チャネル 8 ビット画像への変換を行い, 画像内に対象の歯以外の歯を含むことを避けるために, 付与されたアノテーションに基づいて歯の領域の切り抜きとマスキングを行った. 表 3.10 に前処理のフローを示す. まず, 歯の輪郭情報を用いて歯の矩形領域を切り抜き, 上下左右に 20 ピクセルのパディングを行う. 次に, 歯の輪郭に対してカーネルサイズが 20×20 ピクセルの膨張処理を行う. 最後に, 状態診断モデルの入力サイズである 224×224 ピクセルの大きさにリサイズする. データ拡張では上下左右のランダムフリップ, 0.9 倍から 1.1 倍までの拡大縮小, -20 度から 20 度までの回転, -4 度から 4 度までの水平方向シアー変換, -4 度から 4 度までの垂直方向シアー変換, 0.8 倍から 1.2 倍までの明るさ変化, ガウシアンぼかし, ガウシアンノイズ付加を適用した. ガウシアンぼかしでは標準偏差を 0 から 3 までの範囲でランダムに変動させた. ガウシアンノイズ付加における正規分布

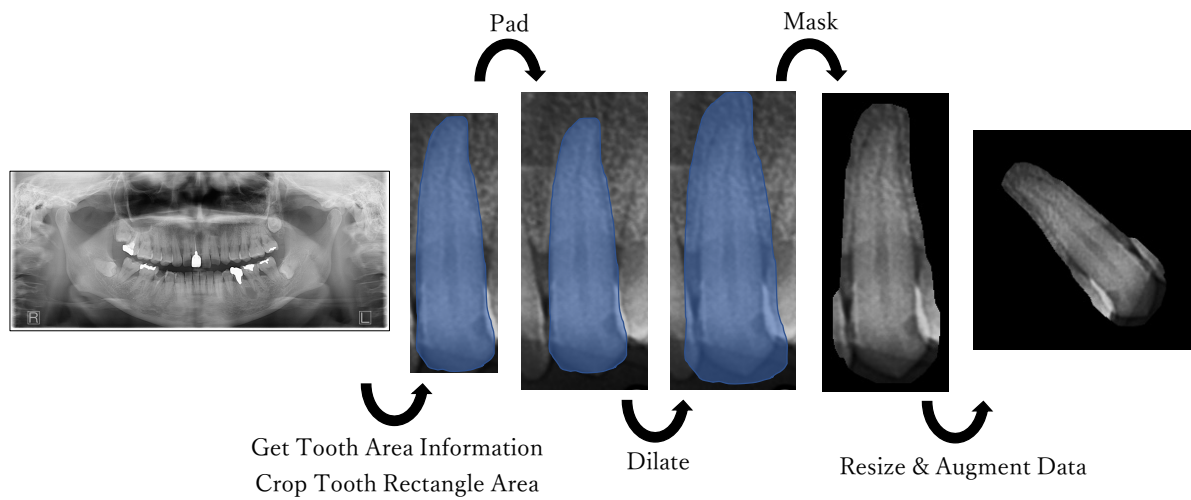


図 3.10 画像の前処理

の平均は 0 とし、標準偏差は 0 から 15 までの範囲でランダムに変動させた。

状態診断モデルの学習では入力画像のピクセル値を 0 から 1 に正規化し、単精度浮動小数点数として扱った。最適化は Adam[54] を使用し、初期学習率は 0.01、 β_1 は 0.9、 β_2 は 0.999 とした。また、バッチサイズは 32、学習回数は 50 回とし、ステップ数は学習データに含まれる歯の本数をバッチサイズで除算した値に設定した。交差検証を行うために、データセットは画像単位で 5 分割した学習用パノラマ X 線画像データベースを用いた。各データセットにおける状態ラベルの内訳は表 3.8 に示す。

状態診断モデルに関する実験は、Ubuntu 20.04、Intel Xeon W-2235 (3.80GHz)、NVIDIA RTX A4000、メモリ 16GB を搭載した計算機上で行った。コード実装は Python 3.8 を用いて行い、フレームワークは TensorFlow[55] 2.7.0 を用いた。

表 3.8 各データセットにおける状態ラベルの内訳. 学習データ (Training Data, train) および検証データ (Validation Data, val) に含まれる状態ラベルを集計した.

Condition	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
	train	val	train	val	train	val	train	val	train	val
N	23,549	5,582	22,932	6,199	23,316	5,815	23,349	5,782	23,378	5,753
C	867	278	947	198	905	240	951	194	910	235
Per	761	224	836	149	777	208	776	209	790	195
RCF	4,865	1,332	5,026	1,171	5,017	1,180	4,998	1,199	4,882	1,315
Implant	274	75	288	61	269	80	281	68	284	65
In	6,591	1,694	6,631	1,654	6,675	1,610	6,608	1,677	6,635	1,650
Metal	4,282	1,034	4,219	1,097	4,295	1,021	4,182	1,134	4,286	1,030
CR	2,191	617	2,277	531	2,229	579	2,293	515	2,242	566
AF	147	34	150	31	152	29	147	34	128	53
Cro	5,376	1,496	5,591	1,281	5,534	1,338	5,497	1,375	5,490	1,382
FMC	2,875	765	2,965	675	2,945	695	2,883	757	2,892	748
MB	609	207	674	142	667	149	665	151	649	167
VMC	1,306	374	1,372	308	1,342	338	1,357	323	1,343	337
CADCAM	162	42	165	39	161	43	157	47	171	33
Zi	169	29	160	38	147	51	158	40	158	40
Cera	323	88	335	76	326	85	318	93	342	69
Pon	861	238	896	203	876	223	884	215	879	220
IT	2,296	593	2,314	575	2,378	511	2,285	604	2,283	606
HRT	127	40	139	28	134	33	147	20	121	46
RT	302	73	300	75	306	69	296	79	296	79
CRT	1,412	356	1,425	343	1,458	310	1,373	395	1,404	364
HIT	307	83	299	91	328	62	315	75	311	79
HET	149	42	153	38	154	37	155	36	153	38

第 4 章

カルテ入力支援システムの評価

4.1 状態診断モデルの評価

状態診断モデルに対して交差検証を行い、適合率(Precision)、感度(Recall)、特異度(Specificity)を算出した結果を表 4.1 に示す。ただし、状態診断モデルの出力(尤度)に対するしきい値を 0.5 に設定し、しきい値以上の尤度をもつ状態ラベルを正(1)、しきい値未満の尤度をもつ状態ラベルを負(0)とした。なお、適合率、感度、特異度は、真陽性(True Positive, TP)、真陰性(True Negative, TN)、偽陽性(False Positive, FP)、偽陰性(False Negative, FN)を用いて、それぞれ式 4.1、式 4.2、式 4.3 で定義される。

$$Precision = \frac{TP}{TP + FP} \quad (4.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4.3)$$

適合率は予想したラベルのうち実際に正であったラベルの割合、感度は実際に正であるラベルのうち正と予測できたラベルの割合、特異度は実際に負であるラベルのうち負と予測できたラベルの割合を表す。真陽性、真陰性、偽陽性、偽陰性の計算は表 4.2 のように行った。

9 状態の診断では適合率 92.73%、感度 91.59%、特異度 98.95% を達成し、20 状態の診断では適合率 90.26%、感度 85.69%、特異度 99.44% を達成した。異常なし(N)、根管充填(RCF)、インプラント(Implant)、メタルインレー(Metal)、クラウン(Cro)、ポンティック(Pon)の感度は 90% 以上であるが、う蝕(C)、根尖病巣(Per)、複合レジン(CR)、メタルボンド(MB)、CAD/CAM 冠(CAD/CAM)、ジルコニア(Zi)、セラミック(Cera)、半埋伏歯(HRT)、埋伏歯(RT)、水平埋伏智歯(HET)の感度は 50% 以下であった。

表 4.1 状態診断モデルの交差検証による評価. 各指標の値は5つのデータセットにおける適合率, 感度, 特異度の平均値を表す.

Condition	Precision (%)	Recall (%)	Specificity (%)
N	94.33	96.63	91.79
C	72.74	38.42	99.62
Per	53.71	32.64	99.39
RCF	96.34	96.14	99.45
Implant	98.87	95.96	99.99
In	88.09	84.60	97.57
Metal	91.45	93.45	98.89
CR	71.90	46.93	98.85
AF	48.19	51.63	99.76
Cro	94.10	91.49	99.04
FMC	84.83	86.25	98.71
MB	69.89	18.31	99.79
VMC	71.79	65.69	99.01
CADCAM	61.90	6.14	99.98
Zi	84.19	40.65	99.95
Cera	67.06	16.62	99.93
Pon	97.09	96.79	99.93
IT	86.81	93.02	99.50
HRT	33.50	7.90	99.96
RT	27.85	15.53	99.80
CRT	67.90	65.98	99.63
HIT	70.16	58.67	99.79
HET	63.11	27.59	99.91
9 Conditions	92.73	91.59	98.95
20 Conditions	90.26	85.69	99.44

表 4.2 評価値の計算例. 9 状態の診断において, ある歯に対する予測 (Prediction) がインレーおよびクラウンであり, 正解 (Ground Truth) がインレーおよび根尖病巣であった場合の計算を示す.

Label	Prediction	Ground Truth	Result
N	0	0	TN
C	0	0	TN
Per	0	1	FN
RCF	0	0	TN
Implant	0	0	TN
In	1	1	TP
Cro	1	0	FP
Pon	0	0	TN
IT	0	0	TN

4.2 テスト用パノラマ X 線画像に対するカルテ入力支援システムの評価

テスト用パノラマ X 線画像に対するカルテ入力支援システムの結果について、感度を算出した結果を表 4.3、歯の検出に対する感度を表 4.4、歯の状態診断に対する感度を表 4.5 に示す。ただし、検出された歯と実際の歯の矩形領域に対して IoU を計算し、IoU が 0.8 以上の歯に対して歯式と状態ラベルの照合を行った。また、歯検出モデルで検出されなかった歯を欠損歯 (Missing Tooth, MT) とし、欠損歯も一つの歯の状態として扱った。

カルテ入力支援システムの感度は 85.31% を達成し、カルテ入力作業の効率化が可能であることが示された。う蝕と根尖病巣は感度が 50% を大きく下回っており、カルテ入力支援の対象として不適切であることが示唆された。

図 4.1, 4.2, 4.3, 4.4, 4.5 にテスト用パノラマ画像におけるカルテ入力支援システムの結果を例示する。ただし、上図はカルテ入力支援システムの結果を描画したパノラマ X 線画像、下図は正解のアノテーションを描画したパノラマ X 線画像である。

表 4.3 カルテ入力支援システムの感度. 行方向に各歯における感度, 列方向に各状態における感度を示す. 空白はテスト用パノラマ X 線画像データベースに該当する歯式と状態の組み合わせが存在しないことを表す.

	N	C	Per	RCF	Implant	In	Cro	Pon	IT	MT	ALL
11	91.89	0.00	0.00	94.44	50.00	85.71	100.00	78.57		100.00	85.58
12	94.49	0.00	0.00	88.89	75.00	72.22	93.48	87.50		100.00	87.04
13	94.89	11.11	0.00	76.00	100.00	61.54	81.40	66.67	100.00	87.50	84.19
14	96.74	33.33	0.00	65.52	88.89	91.11	78.38	91.67		78.57	84.19
15	96.10	0.00	0.00	76.74	83.33	94.74	84.38	83.33		100.00	82.79
16	98.18	25.00	9.09	76.19	92.31	88.89	75.00	90.91		94.74	80.47
17	91.07	0.00	0.00	86.11	100.00	91.49	80.36	83.33	100.00	94.74	79.07
18	81.82	100.00	0.00	100.00		100.00	100.00		79.41	99.37	94.42
21	98.21	0.00	0.00	90.48	100.00	76.92	93.75	92.31		100.00	84.19
22	94.35	0.00	0.00	89.47	100.00	80.00	93.33	100.00		77.78	84.19
23	94.44	12.50	0.00	73.08	87.50	61.54	96.88	100.00		100.00	83.26
24	97.06	0.00	0.00	80.00	83.33	84.09	75.00	83.33		100.00	84.65
25	93.26	0.00	0.00	83.33	100.00	74.42	80.70	100.00		100.00	83.26
26	97.78	25.00	0.00	91.67	75.00	87.50	90.00	100.00		95.24	80.93
27	90.74	60.00	0.00	90.91	100.00	86.54	87.30	100.00	100.00	96.97	79.07
28	91.67	40.00				50.00	100.00		86.21	98.79	94.42
31	99.44	0.00	0.00	75.00	50.00	50.00	75.00	66.67		100.00	92.56
32	97.27	25.00	0.00	87.50	0.00	20.00	83.33	75.00		100.00	90.70
33	99.43	0.00	0.00	88.89	100.00	50.00	90.00	100.00		33.33	89.77
34	99.24	33.33	0.00	76.92	100.00	81.25	90.32	100.00		100.00	88.84
35	95.18	0.00	0.00	65.62	100.00	79.49	81.63	90.00	50.00	95.65	81.86
36	100.00	33.33	22.22	86.67	91.67	85.71	75.38	90.00		100.00	76.15
37	86.96	70.00	15.38	90.00	84.62	87.76	72.13	100.00	0.00	96.97	77.67
38	100.00	33.33		50.00		75.00	100.00		90.20	97.18	92.56
41	98.91	0.00	0.00	88.89	0.00	0.00	76.92	40.00		100.00	90.70
42	98.92	0.00	0.00	100.00	0.00	80.00	91.67	66.67		100.00	93.02
43	99.43	0.00	0.00	90.00	100.00	36.36	86.96			50.00	92.13
44	94.29	0.00	0.00	84.85		75.00	79.41			81.82	83.72
45	95.56	20.00	7.69	87.50	44.44	92.31	84.21	100.00	0.00	92.31	82.33
46	95.00	33.33	29.17	86.96	61.54	93.65	80.60	80.00		91.67	78.90
47	88.00	50.00	33.33	92.31	66.67	87.50	80.00	75.00	0.00	88.24	79.53
48	66.67	66.67	0.00	55.56		85.71	46.15		83.64	98.36	87.91
ALL	96.15	21.69	9.13	84.35	80.24	83.33	83.53	87.24	83.62	96.43	85.31

表 4.4 歯の検出におけるカルテ入力支援システムの感度

	N	C	Per	RCF	Implant	In	Cro	Pon	IT	MT	ALL
11	100.00	100.00	100.00	97.22	100.00	100.00	100.00	85.71		100.00	98.60
12	100.00	66.67	33.33	93.33	75.00	88.89	95.65	87.50		100.00	96.30
13	97.81	100.00	100.00	84.00	100.00	100.00	90.70	66.67	100.00	87.50	95.81
14	98.91	83.33	83.33	93.10	100.00	100.00	94.59	91.67		78.57	96.28
15	100.00	66.67	83.33	97.67	83.33	97.37	93.75	83.33		100.00	96.28
16	100.00	87.50	90.91	95.24	92.31	98.15	95.31	90.91		94.74	96.28
17	100.00	77.78	100.00	97.22	100.00	100.00	96.43	83.33	100.00	94.74	96.74
18	100.00	100.00	100.00	100.00		100.00	100.00		94.12	99.37	98.60
21	100.00	100.00	100.00	95.24	100.00	100.00	95.83	92.31		100.00	98.14
22	99.19	100.00	100.00	94.74	100.00	100.00	97.78	100.00		77.78	97.67
23	99.31	100.00	100.00	92.31	87.50	100.00	100.00	100.00		100.00	98.14
24	98.04	33.33	80.00	85.71	83.33	93.18	92.50	83.33		100.00	93.49
25	95.51	100.00	100.00	90.48	100.00	93.02	92.98	100.00		100.00	94.42
26	97.78	75.00	88.89	95.00	100.00	100.00	98.57	100.00		95.24	97.21
27	98.15	80.00	95.00	95.45	100.00	100.00	96.83	100.00	100.00	96.97	97.67
28	100.00	60.00				100.00	100.00		89.66	98.79	96.74
31	99.44	100.00	100.00	100.00	100.00	100.00	100.00	66.67		100.00	99.07
32	98.91	100.00	100.00	100.00	100.00	100.00	100.00	75.00		100.00	98.60
33	100.00	100.00	100.00	94.44	100.00	100.00	95.00	100.00		33.33	98.60
34	99.24	100.00	83.33	84.62	100.00	100.00	96.77	100.00		100.00	97.67
35	100.00	100.00	87.50	87.50	100.00	100.00	91.84	90.00	50.00	95.65	96.74
36	100.00	83.33	85.19	91.67	91.67	100.00	89.23	90.00		100.00	95.41
37	100.00	100.00	92.31	92.50	92.31	100.00	90.16	100.00	100.00	96.97	96.74
38	100.00	100.00		50.00		100.00	100.00		98.04	97.18	97.21
41	98.91	100.00	100.00	100.00	100.00	100.00	100.00	60.00		100.00	98.14
42	98.92	100.00	100.00	100.00	100.00	80.00	100.00	66.67		100.00	98.14
43	100.00	100.00	66.67	90.00	100.00	90.91	91.30			50.00	97.69
44	99.29	83.33	83.33	87.88		100.00	88.24			81.82	96.74
45	98.89	40.00	84.62	93.75	66.67	100.00	92.98	100.00	100.00	92.31	94.88
46	100.00	83.33	83.33	91.30	61.54	100.00	89.55	80.00		91.67	92.20
47	100.00	66.67	91.67	96.15	77.78	93.75	94.00	75.00	100.00	88.24	93.02
48	100.00	66.67	16.67	66.67		100.00	46.15		98.18	98.36	94.88
ALL	99.22	86.75	86.96	92.65	89.22	98.22	93.90	88.27	95.48	96.43	96.69

表 4.5 歯の状態診断におけるカルテ入力支援システムの感度

	N	C	Per	RCF	Implant	In	Cro	Pon	IT	MT	ALL
11	91.89	0.00	0.00	94.44	50.00	85.71	100.00	78.57		100.00	85.58
12	94.49	0.00	0.00	91.11	100.00	83.33	95.65	87.50		100.00	87.96
13	94.89	11.11	0.00	80.00	100.00	61.54	83.72	100.00	100.00	87.50	85.12
14	96.74	33.33	0.00	65.52	88.89	91.11	78.38	100.00		100.00	86.05
15	96.10	0.00	0.00	79.07	91.67	97.37	89.06	100.00		100.00	85.12
16	98.18	25.00	9.09	76.19	100.00	90.74	78.12	90.91		100.00	81.86
17	91.07	11.11	0.00	88.89	100.00	91.49	82.14	83.33	100.00	100.00	80.93
18	81.82	100.00	0.00	100.00		100.00	100.00		79.41	99.37	94.42
21	98.21	0.00	0.00	95.24	100.00	76.92	95.83	92.31		100.00	84.65
22	94.35	0.00	0.00	89.47	100.00	80.00	93.33	100.00		100.00	85.12
23	94.44	12.50	0.00	73.38	100.00	61.54	96.88	100.00		100.00	83.72
24	98.04	0.00	0.00	80.00	83.33	84.09	75.00	83.33		100.00	85.12
25	93.26	0.00	0.00	85.71	100.00	74.42	82.46	100.00		100.00	83.72
26	97.78	25.00	0.00	93.33	75.00	87.50	90.00	100.00		100.00	81.86
27	90.74	60.00	0.00	90.91	100.00	86.54	87.30	100.00	100.00	100.00	79.53
28	91.67	60.00				50.00	100.00		86.21	100.00	95.81
31	99.44	0.00	0.00	75.00	50.00	50.00	75.00	66.67		100.00	92.56
32	98.36	25.00	0.00	87.50	0.00	20.00	83.33	75.00		100.00	91.63
33	99.43	0.00	0.00	88.89	100.00	50.00	90.00	100.00		100.00	90.70
34	99.24	33.33	0.00	76.92	100.00	81.25	90.32	100.00		100.00	88.84
35	95.18	0.00	0.00	65.62	100.00	79.49	81.63	100.00	50.00	100.00	82.79
36	100.00	33.33	22.22	90.00	91.67	85.71	80.00	95.00		100.00	77.06
37	86.96	70.00	15.38	92.50	84.62	87.76	77.05	100.00	0.00	100.00	79.07
38	100.00	33.33		50.00		75.00	100.00		88.24	97.89	93.02
41	98.91	0.00	0.00	88.89	0.00	0.00	76.92	40.00		100.00	90.70
42	100.00	0.00	0.00	100.00	0.00	80.00	91.67	66.67		100.00	93.95
43	99.43	0.00	0.00	100.00	100.00	36.36	95.65			100.00	93.52
44	94.29	0.00	0.00	87.88		75.00	82.35			100.00	84.65
45	95.56	0.00	7.69	89.58	55.56	92.31	89.47	100.00	0.00	100.00	84.19
46	95.00	33.33	29.17	89.13	84.62	93.65	85.07	86.67		100.00	82.11
47	88.00	50.00	33.33	92.31	88.89	93.75	82.00	87.50	0.00	100.00	84.65
48	66.67	66.67	0.00	77.78		85.71	92.31		85.45	99.18	89.30
ALL	96.31	22.89	9.13	86.26	86.23	84.22	86.1	90.82	84.18	99.40	86.41

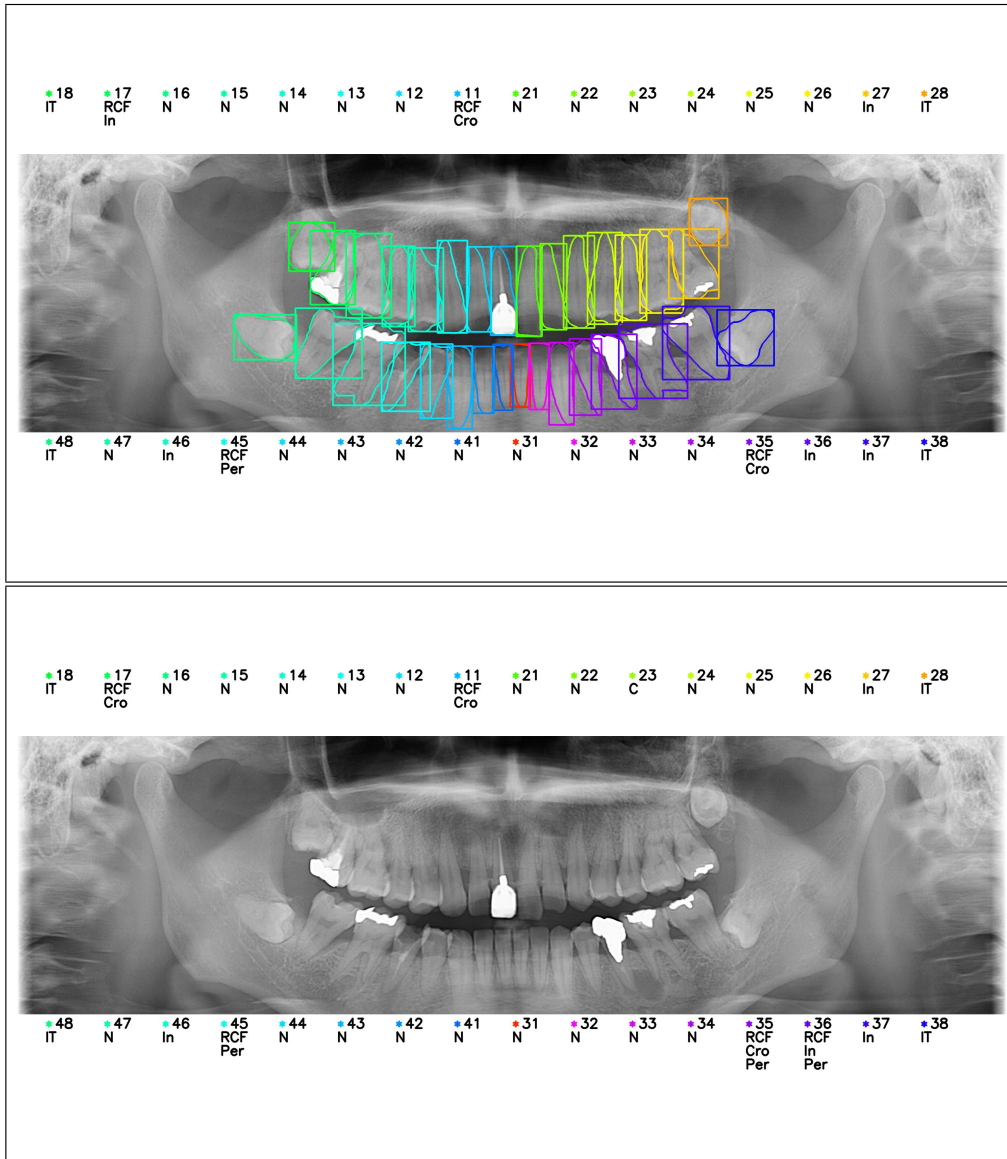


図 4.1 カルテ入力支援システムの出力例 1

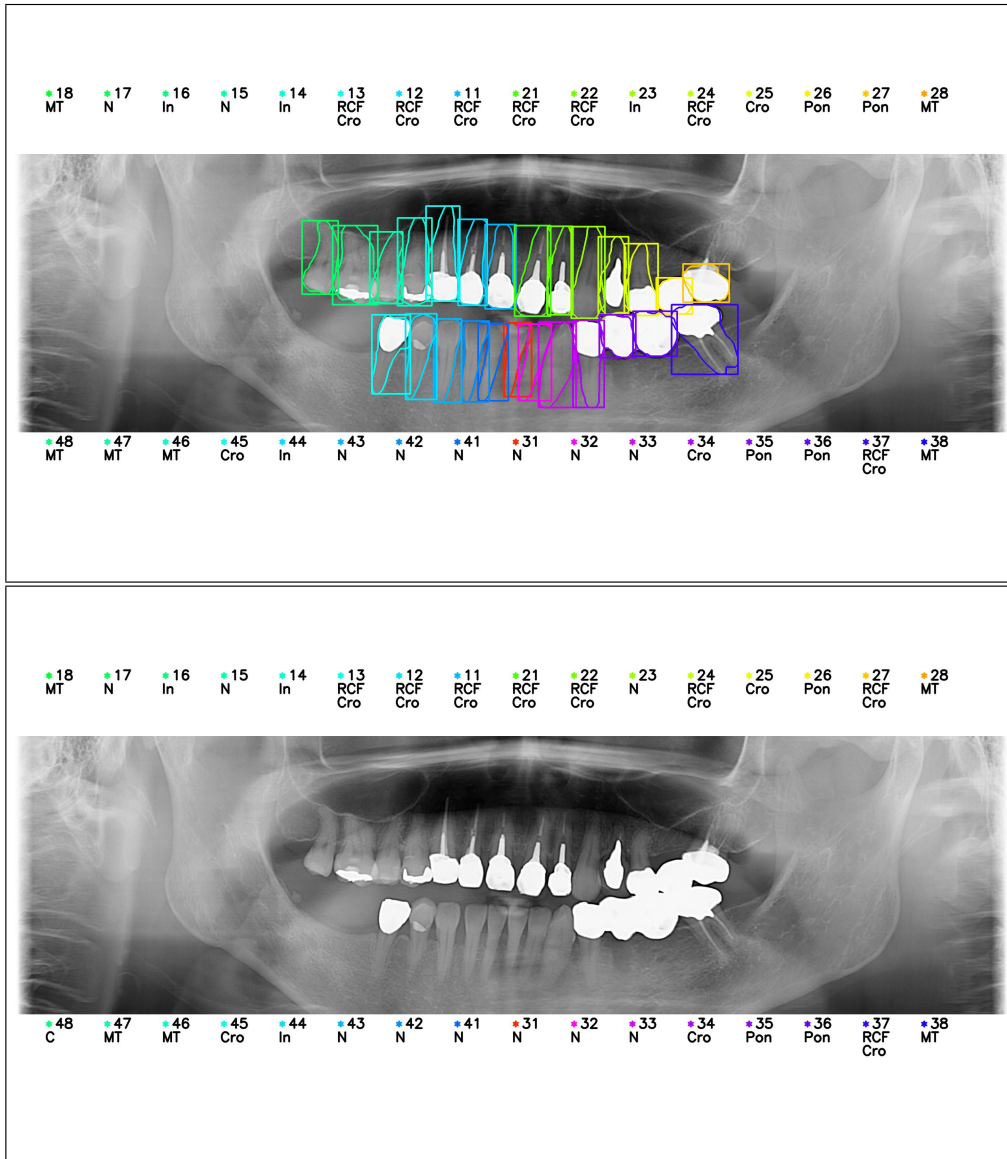


図 4.2 カルテ入力支援システムの出力例 2

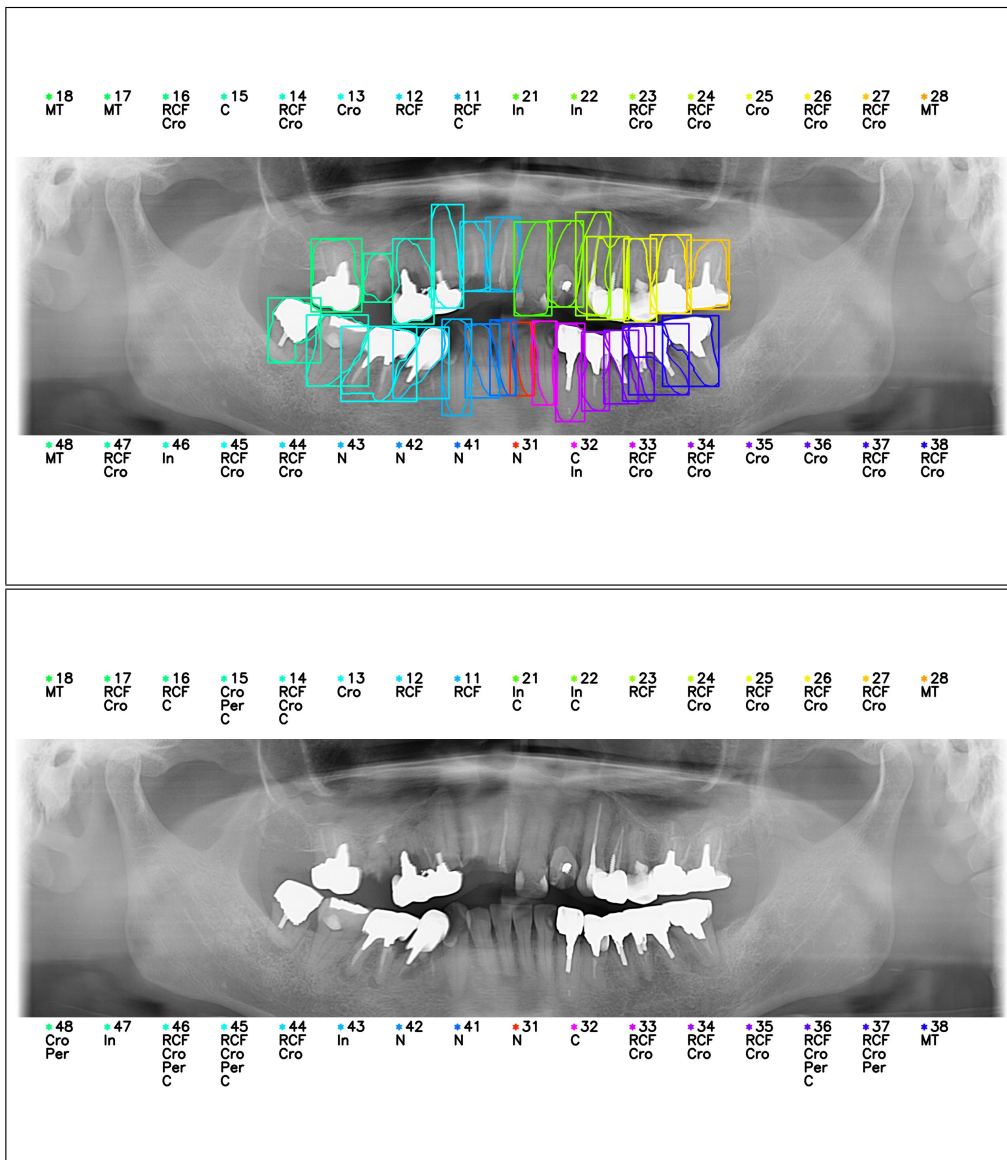


図 4.3 カルテ入力支援システムの出力例 3

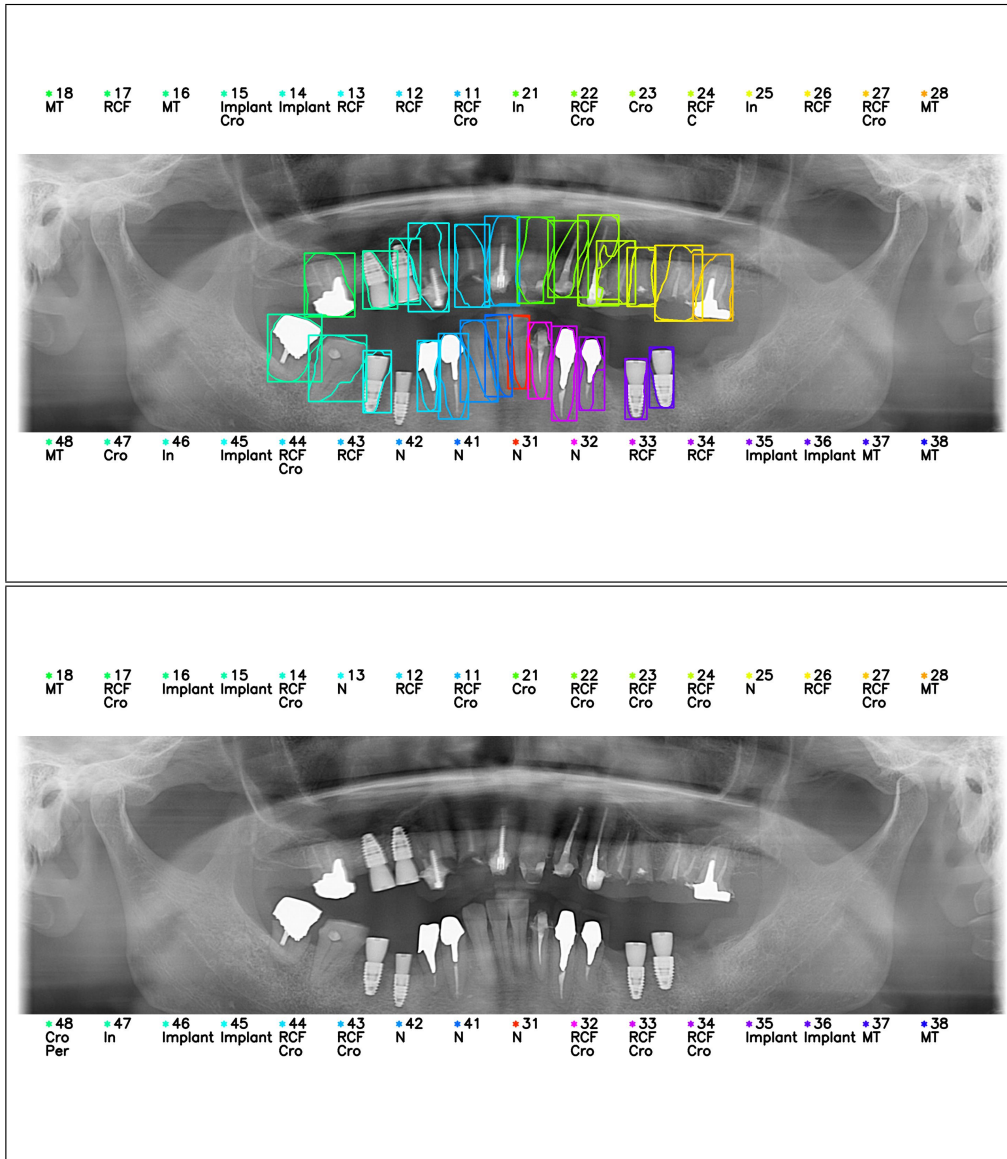


図 4.4 カルテ入力支援システムの出力例 4

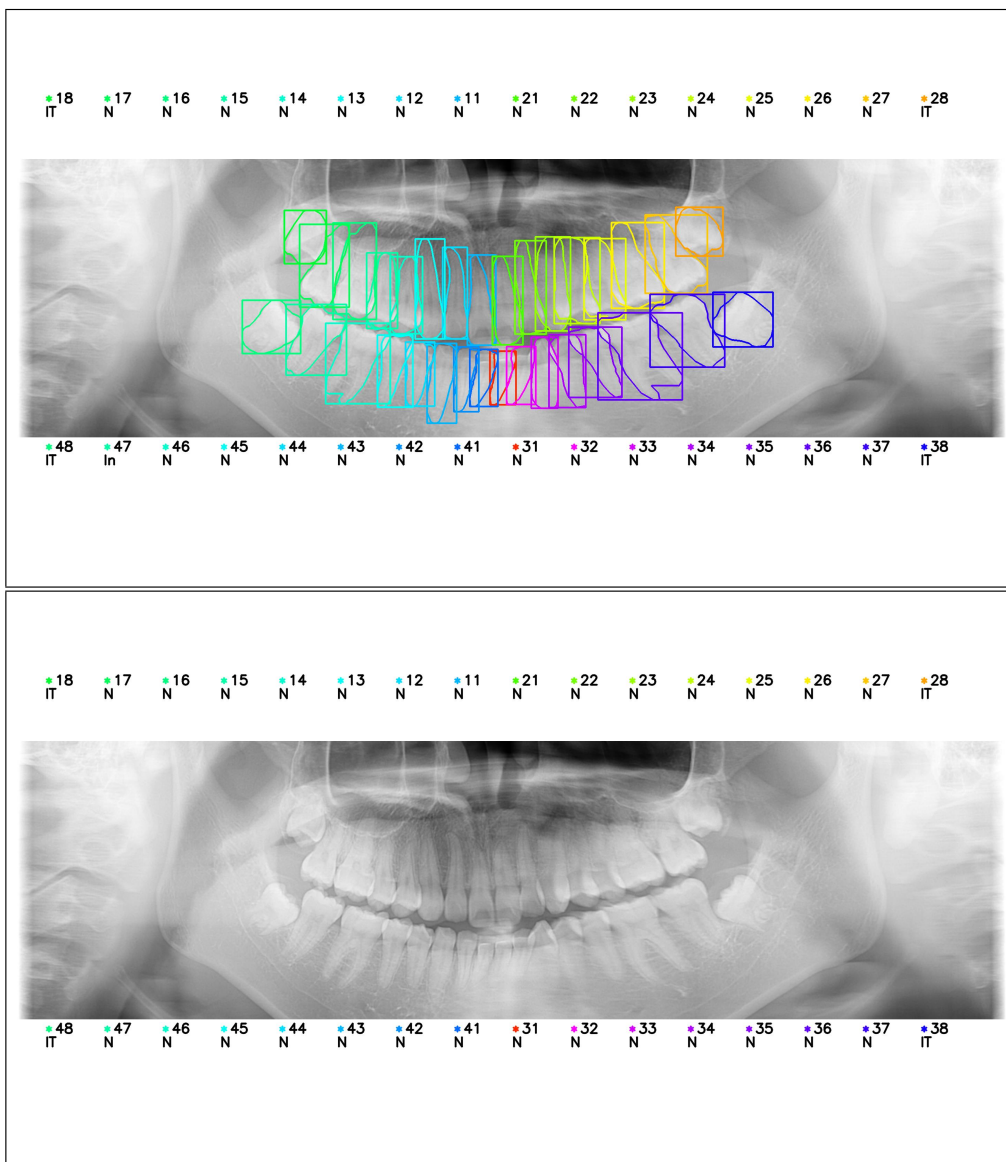


図 4.5 カルテ入力支援システムの出力例 5

4.3 CT 画像から生成したパノラマ X 線画像に対するカルテ入力支援の例

図 4.6 に生成したパノラマ X 線画像を例示する。また、図 4.7, 4.8, 4.9, 4.10, 4.11, 4.12 に生成したパノラマ X 線画像に対するカルテ入力支援システムの診断結果を示す。ただし、対象の歯に 0.5 以上のスコアをもつ状態ラベルが存在しない場合は、最大のスコアをもつ状態ラベルを括弧付きの状態ラベルとして表示している。

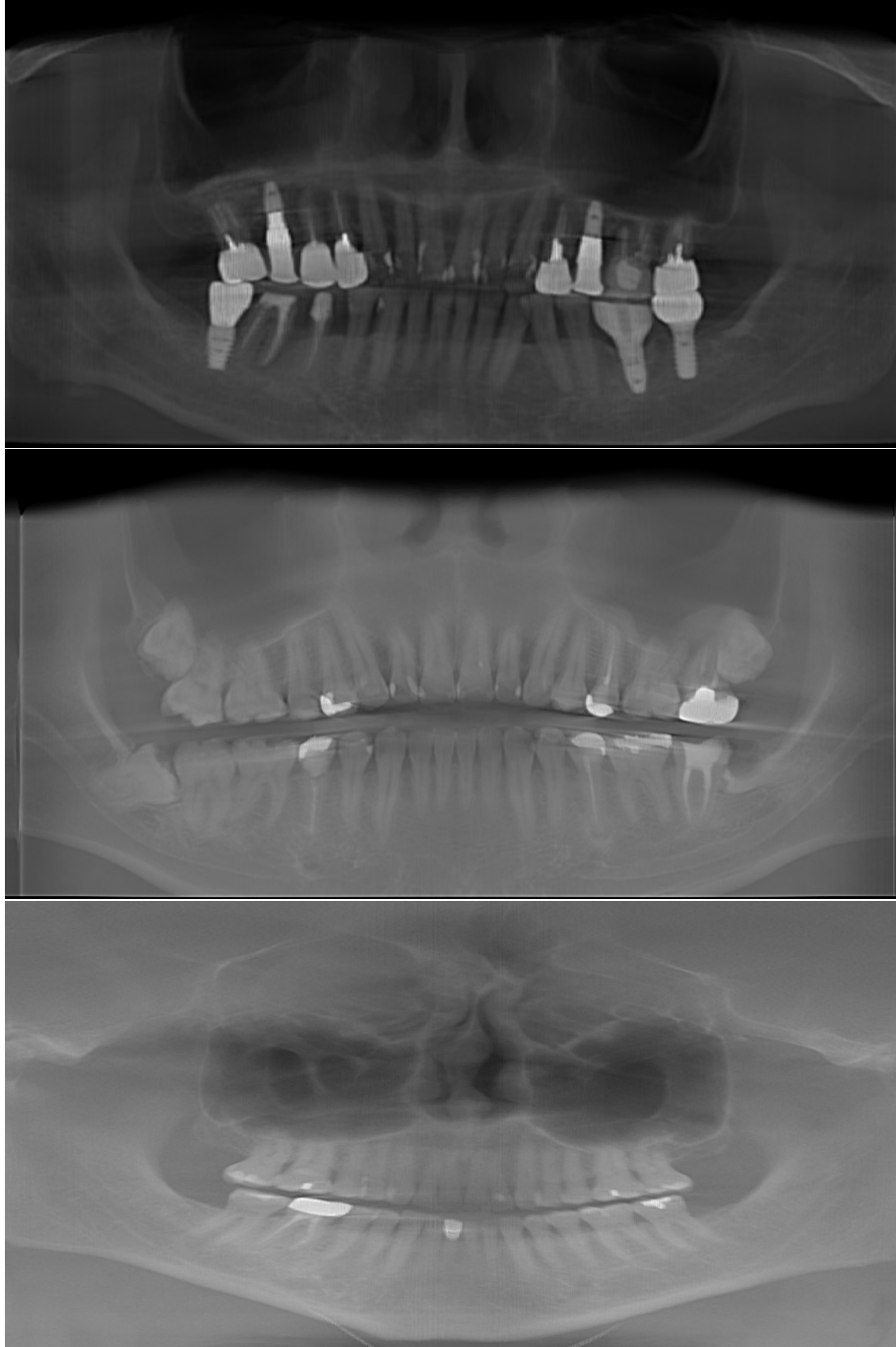


図 4.6 生成したパノラマ X 線画像の例. 上から順に CareStream 9600, Panoura A1, Prexion3D Exploler で撮影された CT 画像からパノラマ再構成システムを用いて自動生成した.

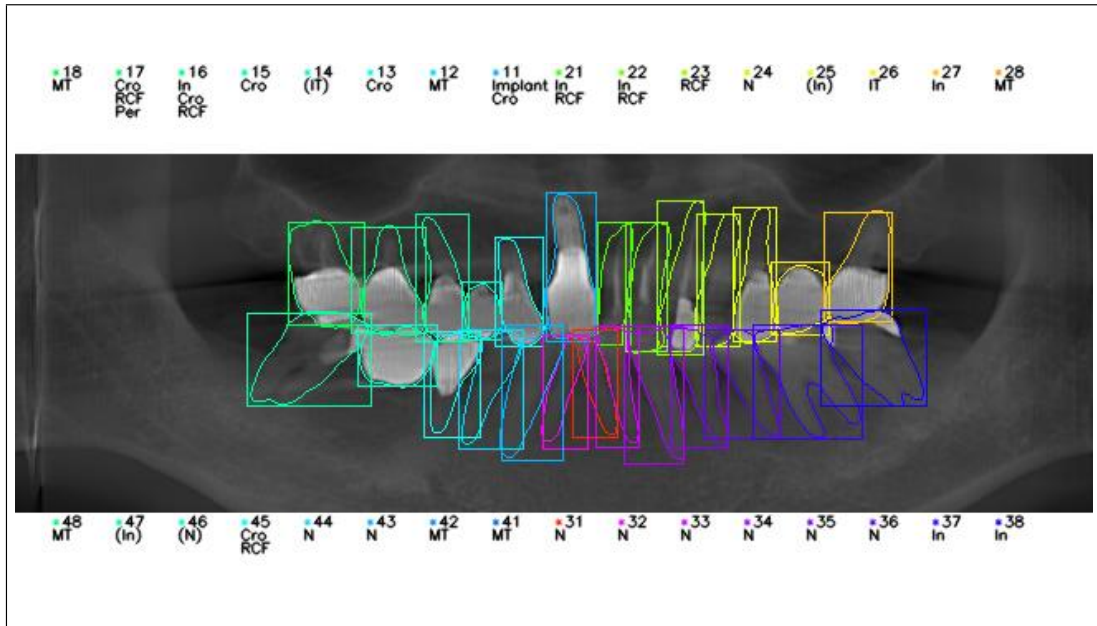


図 4.7 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 1

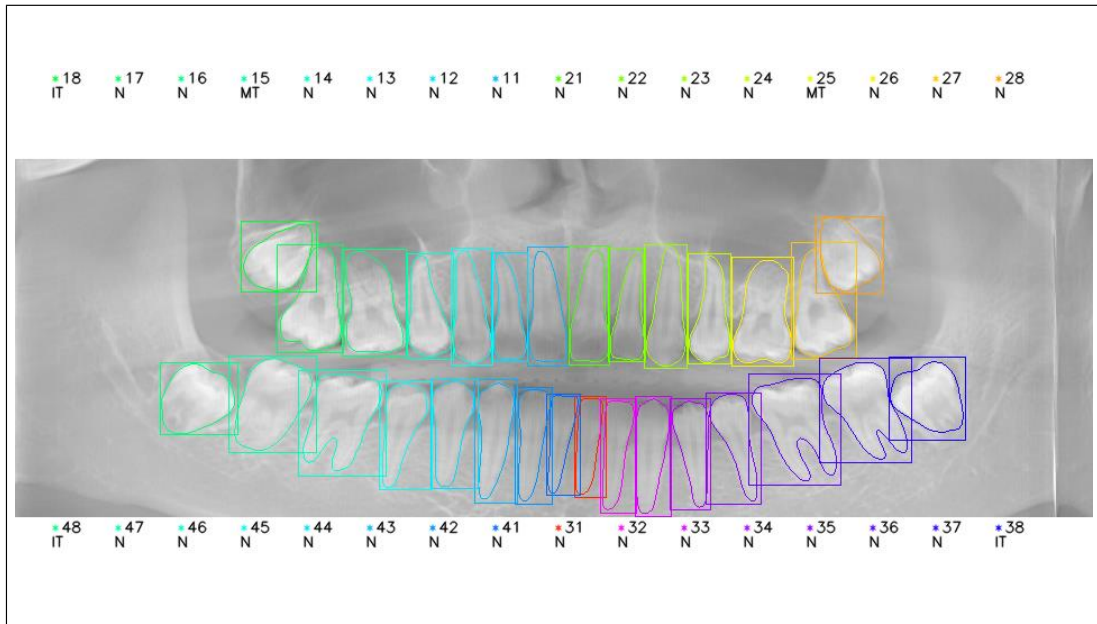


図 4.8 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 2

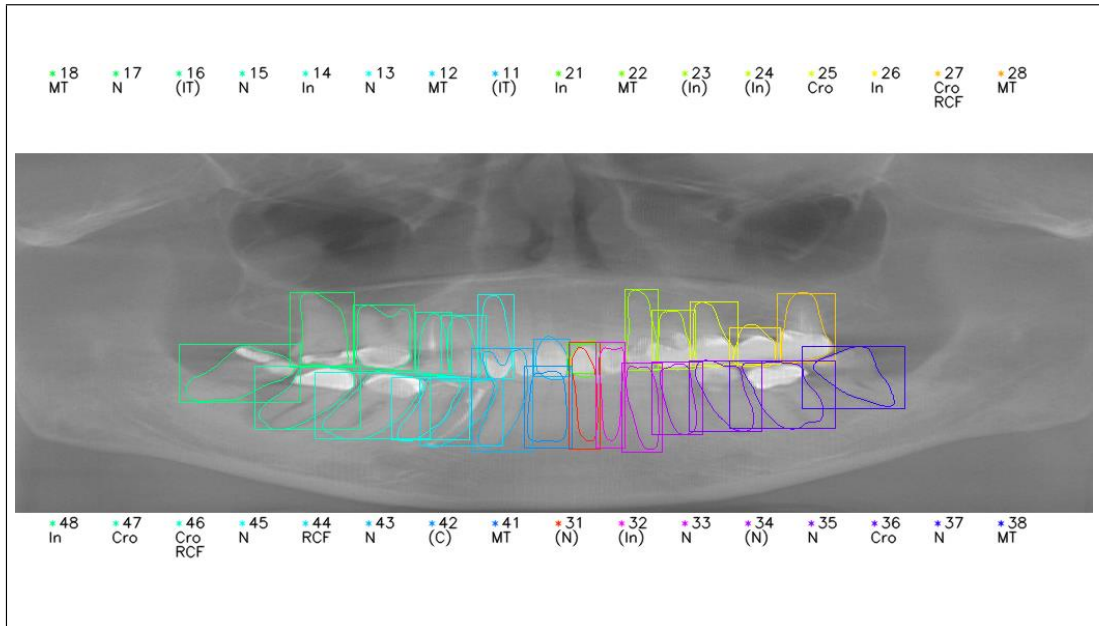


図 4.9 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 3

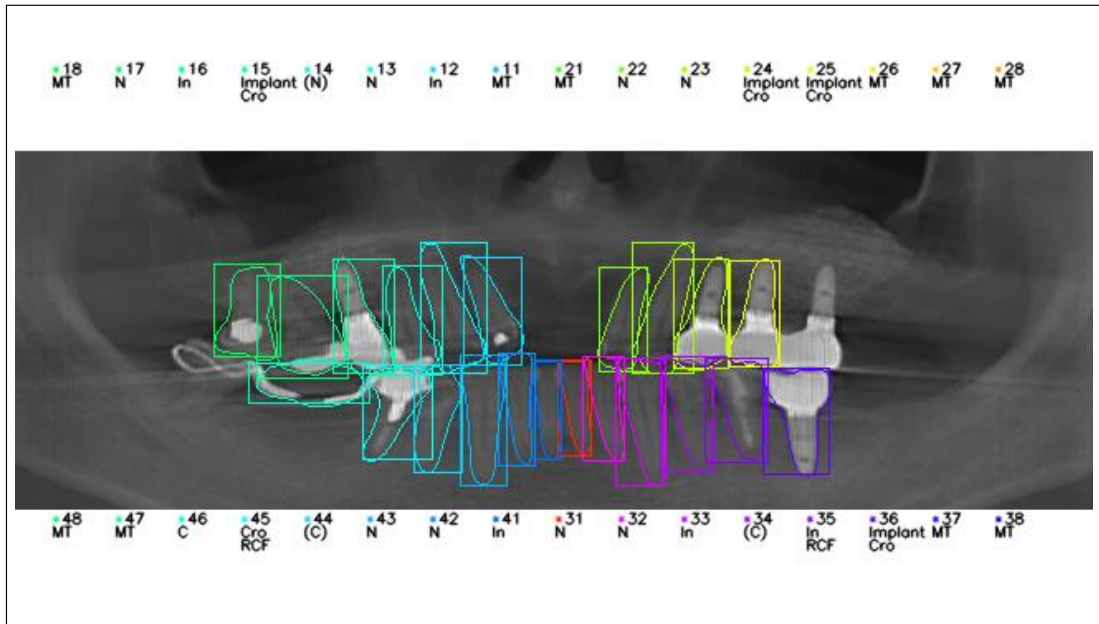


図 4.10 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 4

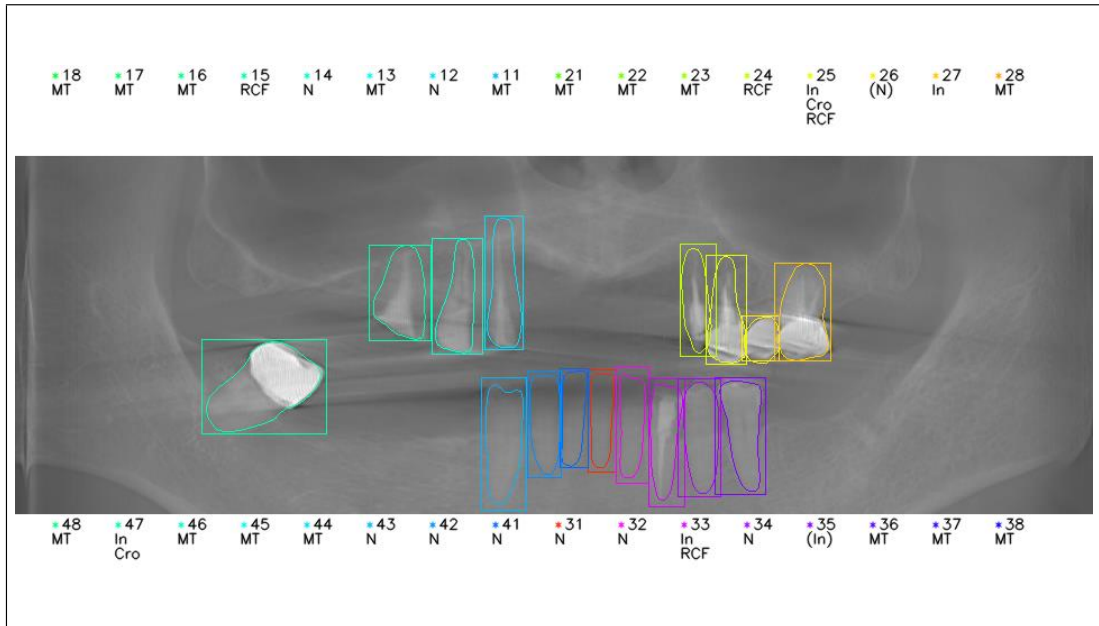


図 4.11 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 5

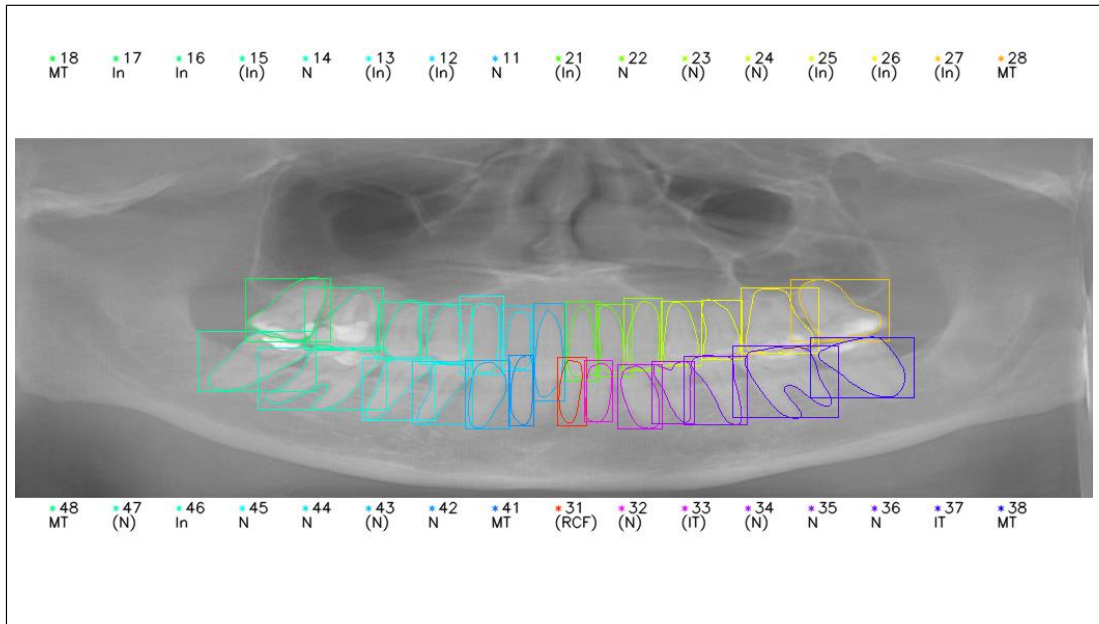


図 4.12 生成したパノラマ X 線画像に対するカルテ入力支援システムの結果 6

第 5 章

考察

明確な X 線不透過像を呈す歯根充填，インレー，クラウン，ポンティック，インプラントにおいて高い感度と適合率を達成した．特に，歯根充填，ポンティック，インプラントはカルテ入力支援システムとしてすでに実用的な精度であり，カルテ入力支援システムの有用性を示している．しかし，パノラマ X 線画像のみでは診断が難しい根尖病巣やう蝕，複合レジン，クラウンの材質などの適合率と感度は低い結果となり，テスト用パノラマ X 画像においても，う蝕の診断はできなかった．カルテ入力支援システムで X 線透過度の微妙な差異を考慮した診断は困難であるため，X 線透過度を重視した画像の前処理や診断アルゴリズムの開発が必要である．う蝕や根尖病巣に関しては状態診断モデルの入力サイズや学習したパノラマ X 線画像の画質などの影響を特に受けやすいと考えられるため，入力サイズの最適化をはじめとした状態診断モデルに関する設定の改善や高品質なパノラマ X 線画像の収集が必要である．また，根尖病巣は歯の周辺の情報的重要であり，歯の輪郭推定の結果や入力画像のマスキング処理などの影響も大きいいため，状態診断モデルの診断対象から除外することも検討しなければならない．

本研究におけるパノラマ X 線画像データベースの作成では複合レジン透過性と不透過性に分けてアノテーションを付与しなかった．透過性の複合レジンう蝕と誤診する可能性が高く，状態診断モデルの学習にも悪影響を与えたと考えられる．今後，複合レジン透過性と不透過性に分けてデータベースを再作成し，本研究では不透過性の複合レジンのみを扱うようにデータセットと状態診断モデルの再構築を行うことで，複合レジンう蝕の診断精度向上が期待できる．不透過性の複合レジンのみを扱う場合，メタルインレー，アマルガムフィリング，不透過性の複合レジン判別では形態を考慮することが有効であると考えられるため，形態に着目した診断モデルを開発する必要がある．

本研究は歯科医 3 名と歯科放射線の専門医 1 名がアノテーションを付与したが，アノテーションは人によって差異が生じることがわかっている．特に，う蝕や根尖病巣は視診を踏まえて観察することが前提であるため，パノラマ画像から得られる情報から判断できることに限界があり，アノテーション付与者の経験則でアノテーション付与を行う必要が生じる．結果として，アノテーションに個人差が生まれ，AI モデルの学習に悪影響を与えることが危惧される．本研究では，専門医 1 名の最終診断を正解として AI モデルの学習を行ったが，本来は複数名の専門医による診断

を考慮すべきである。今後のパノラマ X 線画像データベースの構築で複数名の専門医がアノテーション付与に関わる場合は、パノラマ X 線画像データベース自体の精度、すなわち人による診断の精度を算出し、カルテ入力支援システムの診断結果と比較する必要があると考えられる。また、人によるアノテーション付与は前述のような問題があるため、歯科医院が所有するカルテ情報を活用することも検討している。今後、Anotee とカルテ情報の連携などが可能となれば、より充実したパノラマ X 線画像データベースの構築が期待できる。

本研究で構築したパノラマ X 線画像データベースが含む状態ラベルの数は不均衡であり、状態診断モデルの学習に大きな影響を与えている。不均衡データの対処として、Class-balanced Loss[56] や Focal Loss[57] を状態診断モデルの損失関数に適用することで、状態診断の精度改善が期待できる。また、マルチラベル分類を行うのではなく、1つの状態を診断する診断モデルを複数構築することも精度改善に有効であると考えられる。

埋伏 (IT) の適合率と感度が高く、歯が埋伏しているか否かの診断はカルテ入力支援システムで可能であると考えられるが、半埋伏歯と埋伏歯の適合率と感度、水平智歯の感度は著しく低い。このことから埋入の深度を状態診断モデルで診断することは難しく、埋入の深度を測るために別途アルゴリズムの開発が必要である。埋入の深度の計算は PellGregory 分類 [58] と Winter 分類 [59] にしたがって定量的に行うことが望ましいが、Sukegawa ら [60] は CNN を用いた手法で Pell-Gregory 分類と Winter 分類を行っており、各分類で約 80-90% の精度を達成している。カルテ入力支援システムに対して埋伏の診断に特化した AI 診断モデルと歯科データベースを構築および追加することで、埋入の深度の診断にも対応可能であると考えられる。ただし、完全埋伏歯に関しては明確な特徴があり深度を計測する必要がないことから、今後データを追加し状態診断モデルを追加学習することで改善が期待できる。

本研究で開発したカルテ入力支援システムは歯と歯の周辺情報のみを対象に状態診断を行う。将来的に上顎洞炎や嚢胞などの歯の周辺だけでは診断が難しい状態も検出することを検討しているが、現状のシステムでは実現することができない。歯以外の部位や異常所見を検出できるように、既存データに対するアノテーションの追加とパノラマ X 線画像全体を対象とした検出アルゴリズムが必要である。また、カルテ入力支援システムは永久歯のみを対象としており、混合歯列および乳歯に対応していないが、今後 AI モデルの学習に十分な数の混合歯列のデータを収集し追加学習することで対応可能になると考えられる。混合歯列のデータ追加により乳歯晩期残存や智歯以外の歯の埋伏などの状態診断が可能となり、より多様な状態診断が可能になることが期待できる。

カルテ入力支援システムは CNN ベースの転移学習モデルに基づき構築したが、パノラマ X 線画像データベースの充実に伴い、アーキテクチャが異なる AI モデルの採用を検討することも今後の重要な課題である。たとえば、Vision Transformer (ViT) [20] は CNN よりも大域的な特徴を学習することができ、より人間の知覚に近い画像認識が可能な AI モデルとして注目を集めており、さまざまなタスクへの応用やアーキテクチャの改良が進んでいる。カルテ入力支援システムに ViT を取り入れることで、より多様な病変の発見や診断の精度改善が期待できる。また、本研究でデータ数の関係で扱わなかった破折などの例が少ない状態は、Auto Encoder[61] を用いた異常検知技術に基づく手法で対応することも検討している。

パノラマ再構成システムを用いて CT 画像から生成したパノラマ X 線画像は読影に有用な品質であり、生成したパノラマ X 線画像に対するカルテ入力支援システムの結果は、今後応用が期待できる精度であった。しかし、歯の検出および輪郭推定の結果は比較的良好であるのに対し、状態診断はモダリティの変化や CT 画像特有のノイズの影響を強く受けており、明確に判断できるクラウンやインレー、インプラント、ポンティックなどの誤診が目立った。特に、生成したパノラマ X 線画像にはアーティファクトが多分に含まれているため、アーティファクトを低減する処理を追加することで、状態診断の改善が期待できる。また、CT 画像から生成したパノラマ X 線画像と実際のパノラマ X 線画像は解像度やアスペクト比が大きく異なるため、生成したパノラマ X 線画像にアノテーションを付与し、カルテ入力支援システムの各 AI モデルを再学習を行うことで、よりモダリティの差異に頑健なカルテ入力支援システムを構築できると考えられる。

第 6 章

結論

本研究では歯科医のパノラマ X 線画像の読影およびカルテ入力における負担軽減を目指し、カルテ入力支援システムとパノラマ X 線画像再構成システムについて研究開発を行った。カルテ入力支援システムは高精度を達成し、実際にカルテ入力の作業量を大きく低減する効果があることを示した。また、パノラマ X 線画像再構成システムは CT 画像から読影に足る品質のパノラマ X 線画像を自動生成できることが示され、カルテ入力支援システムを CT 画像にも応用可能であることが示唆された。今後はパノラマ X 線画像データベースのさらなる充実とともに、カルテ入力支援システムおよびパノラマ X 線画像再構成システムの改良を進める。また、社会実装に向けてさまざまなパノラマ撮影機で撮影されたパノラマ X 線画像によるカルテ入力支援システムのテストが必要であるため、より多様な種類のパノラマ X 線画像の収集も引き続き行う。

謝辞

本研究を進めるにあたり，絶えず熱心に御指導，ご教授いただいた徳島大学 大学院 社会産業理工学研究部 北研二 教授，松本和幸 准教授，吉田稔 講師，西村良太 講師に心から感謝，御礼申し上げます。また，本論文の作成について主査，副査をご担当いただき，ご指導をいただいた徳島大学 大学院 社会産業理工学県研究部 獅々堀正幹 教授，泓田正雄 教授に心から御礼申し上げます。

本研究開発にご協力いただいたこばやし歯科クリニック 小林健一郎 氏，新井広幸 氏をはじめとした歯科医の方々，株式会社吉田製作所 友江剛 氏，高橋信生 氏，鶴木クリニック医科・歯科 鶴木次郎 氏，株式会社 Prexion 今田泰豊 氏，富井和仁 氏，古月昇 氏に心から感謝，御礼申し上げます。最後に，研究にご協力いただいた徳島大学 知能情報工学科 A-2 グループ北研究室 博士前期課程諸氏に感謝し，御礼申し上げます。

参考文献

- [1] D. Ferrucci. Introduction to “this is watson”. *IBM Journal of Research and Development*, Vol. 56, No. 3.4, p. 1, 2012.
- [2] P. Amorim, T. Moraes, J. Silva, and H. Pedrini. Invesalius: an interactive rendering framework for health care support. In *International symposium on visual computing*, pp. 45–54, 2015.
- [3] A. Zsigmondy. Grundzuge einer praktischen methode zur raschen und genauen vormerkung der zahnarztlichen beobachtungen und operationen. *Dtsch Vjschr Zahnhk*, Vol. 1, pp. 209–211, 1861.
- [4] C. Palmer. American dental association. *Dental Cosmos*, Vol. 12, pp. 522–523, 1870.
- [5] S. Peck and L. Peck. A time for change of tooth numbering systems. *Journal of dental education*, Vol. 57, No. 8, pp. 643–647, 1993.
- [6] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, Vol. 60, No. 6, pp. 84–90, 2017.
- [7] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pp. 448–456, 2015.
- [8] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [9] F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, Vol. 65, No. 6, p. 386, 1958.
- [10] S. Amari. Backpropagation and stochastic gradient descent method. *Neurocomputing*, Vol. 5, No. 4-5, pp. 185–196, 1993.
- [11] L. Bottou. Stochastic gradient learning in neural networks. *Proceedings of Neuro-Nimes*, Vol. 91, No. 8, p. 12, 1991.
- [12] B. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, Vol. 4, No. 5, pp. 1–17, 1964.
- [13] T. Tieleman and G. Hinton. Lecture 6.5-rmsprop: Divide the gradient by a running

- average of its recent magnitude. *COURSERA: Neural networks for machine learning*, Vol. 4, No. 2, pp. 26–31, 2012.
- [14] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [15] I. Sobel and G. Feldman. A 3×3 isotropic gradient operator for image processing. *Pattern Classification and Scene Analysis*, pp. 271–272, 1973.
- [16] M. Lin, Q. Chen, and S. Yan. Network in network. *arXiv preprint arXiv:1312.4400*, 2013.
- [17] M. Tan and Q. Le. Efficientnetv2: Smaller models and faster training. In *International Conference on Machine Learning*, pp. 10096–10106, 2021.
- [18] M. Tan and Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pp. 6105–6114, 2019.
- [19] I. Bello, W. Fedus, X. Du, E. Cubuk, A. Srinivas, T. Lin, J. Shlens, and B. Zoph. Revisiting resnets: Improved training and scaling strategies. *Advances in Neural Information Processing Systems*, Vol. 34, pp. 22614–22627, 2021.
- [20] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [21] S. Gupta and M. Tan. Efficientnet-edgetpu: Creating accelerator-optimized neural networks with automl. *Google AI Blog*, Vol. 2, p. 1, 2019.
- [22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018.
- [23] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, 2018.
- [24] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.
- [25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, Vol. 30, , 2017.
- [26] P. Ramachandran, B. Zoph, and Q. Le. Searching for activation functions. *arXiv preprint arXiv:1710.05941*, 2017.
- [27] Y. Erdaw and E. Tachbele. Machine learning model applied on chest x-ray images enables automatic detection of covid-19 cases with high accuracy. *International Journal of General Medicine*, Vol. 14, p. 4923, 2021.
- [28] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In

- 2005 *IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Vol. 1, pp. 886–893, 2005.
- [29] M. Hearst, S. Dumais, E. Osuna, J. Platt, and B. Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, Vol. 13, No. 4, pp. 18–28, 1998.
- [30] W. Zhao, W. Jiang, and X. Qiu. Deep learning for covid-19 detection based on ct images. *Scientific Reports*, Vol. 11, No. 1, pp. 1–12, 2021.
- [31] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2097–2106, 2017.
- [32] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang, L. Ye, M. Gao, Z. Zhou, L. Li, J. Wang, Z. Yang, H. Cai, J. Xu, L. Yang, W. Cai, W. Xu, S. Wu, W. Zhang, S. Jiang, L. Zheng, X. Zhang, L. Wang, L. Lu, J. Li, H. Yin, W. Wang, O. Li, C. Zhang, L. Liang, T. Wu, R. Deng, K. Wei, Y. Zhou, T. Chen, J. Lau, M. Fok, J. He, T. Lin, W. Li, and G. Wang. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. *Cell*, Vol. 181, No. 6, pp. 1423–1433, 2020.
- [33] K. Panetta, R. Rajendran, A. Ramesh, S. Rao, and S. Agaian. Tufts dental database: A multimodal panoramic x-ray dataset for benchmarking diagnostic systems. *IEEE Journal of Biomedical and Health Informatics*, Vol. 26, No. 4, pp. 1650–1659, 2021.
- [34] H. Chen, K. Zhang, P. Lyu, H. Li, L. Zhang, J. Wu, and C. Lee. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Scientific reports*, Vol. 9, No. 1, pp. 1–11, 2019.
- [35] D. Tuzoff, L. Tuzova, M. Bornstein, A. Krasnov, M. Kharchenko, S. Nikolenko, M. Sveshnikov, and G. Bednenko. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology*, Vol. 48, No. 4, p. 20180051, 2019.
- [36] J. Lee, D. Kim, S. Jeong, and S. Choi. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of dentistry*, Vol. 77, pp. 106–111, 2018.
- [37] H. Zhu, Z. Cao, L. Lian, G. Ye, H. Gao, and J. Wu. Cariesnet: a deep learning approach for segmentation of multi-stage caries lesion from oral panoramic x-ray image. *Neural Computing and Applications*, pp. 1–9, 2022.
- [38] T. Luo, C. Shi, X. Zhao, Y. Zhao, and J. Xu. Automatic synthesis of panoramic radiographs from dental cone beam computed tomography data. *PloS one*, Vol. 11, No. 6, p. e0156976, 2016.
- [39] J. Hartigan and M. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal*

- of the royal statistical society. series c (applied statistics)*, Vol. 28, No. 1, pp. 100–108, 1979.
- [40] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. Technical report, Stanford, 2006.
- [41] S. AlHarbi, E. Alkofide, and A. AlMadi. Mathematical analyses of dental arch curvature in normal occlusion. *The Angle Orthodontist*, Vol. 78, No. 2, pp. 281–287, 2008.
- [42] T. Papakosta, A. Savva, T. Economopoulos, G. Matsopoulos, and H. Gröhdal. An automatic panoramic image reconstruction scheme from dental computed tomography images. *Dentomaxillofacial Radiology*, Vol. 46, No. 4, p. 20160225, 2017.
- [43] Z. Yun, S. Yang, E. Huang, L. Zhao, W. Yang, and Q. Feng. Automatic reconstruction method for high-contrast panoramic image from dental cone-beam ct data. *Computer methods and programs in biomedicine*, Vol. 175, pp. 205–214, 2019.
- [44] P. Amorim, T. Moraes, J. Silva, H. Pedrini, and R. Ruben. Reconstruction of panoramic dental images through bézier function optimization. *Frontiers in Bioengineering and Biotechnology*, Vol. 8, p. 794, 2020.
- [45] J. Da Re, L. Cereser, G. Como, R. Girometti, and C. Zuiani. Average intensity projection (aip) reconstructions from mdct: a useful complement to the interpretation of chest x-ray examinations. In *European Congress of Radiology*, p. C1374, 2019.
- [46] K. Wada. Labelme: Image polygonal annotation with python. *GitHub Repository*, 2016.
- [47] 志村一男. 理想の x 線画像を目指して: マルチ周波数処理について. *日本放射線技術学会雑誌*, Vol. 57, No. 7, pp. 796–802, 2001.
- [48] K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.
- [49] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pp. 740–755, 2014.
- [50] 原和樹. Mask r-cnn を用いたパントモ画像からの歯牙検出及びセグメンテーション. *電気・電子・情報関係学会四国支部連合大会講演論文集*, No. 15, p. 9, 2020.
- [51] W. Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. *GitHub Repository*, 2017.
- [52] S. Suzuki. Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, Vol. 30, No. 1, pp. 32–46, 1985.
- [53] T. Ridnik, E. Ben-Baruch, A. Noy, and L. Zelnik-Manor. Imagenet-21k pretraining for the masses. *arXiv preprint arXiv:2104.10972*, 2021.
- [54] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [55] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat,

- G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng. Tensorflow: a system for large-scale machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pp. 265–283, 2016.
- [56] Y. Cui, M. Jia, T. Lin, Y. Song, and S. Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9268–9277, 2019.
- [57] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.
- [58] G. Pell and G. Gregory. Impacted mandibular third molars: Classification and modified technique for removal. *Dental Digest*, Vol. 39, pp. 330–338, 1933.
- [59] G. Winter. *Principles of Exodontia as Applied to the Impacted Mandibular Tird Molar: A Complete Treatise on the Operative Technic with Clinical Diagnoses and Radiographic Interpretations*. 1926.
- [60] S. Sukegawa, T. Matsuyama, F. Tanaka, T. Hara, K. Yoshii, K. Yamashita, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, and Y. Furuki. Evaluation of multi-task learning in deep learning-based positioning classification of mandibular third molars. *Scientific Reports*, Vol. 12, No. 1, 2022.
- [61] G. Hinton and R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, Vol. 313, No. 5786, pp. 504–507, 2006.

付録 A

表 A.1 状態診断モデルにおける Dataset1 の評価

	TP	TN	FP	FN	Precision	Recall	Specificity
N	5374	3993	393	230	93.19	95.90	91.04
C	113	9660	23	194	83.09	36.81	99.76
Per	92	9673	75	150	55.09	38.02	99.23
RCF	1350	8554	36	50	97.40	96.43	99.58
Implant	80	9906	0	4	100.00	95.24	100.00
In	1484	7967	242	297	85.98	83.32	97.05
Metal	999	8813	96	82	91.23	92.41	98.92
CR	342	9208	127	313	72.92	52.21	98.64
AF	24	9929	26	11	48.00	68.57	99.74
Cro	1439	8365	72	114	95.23	92.66	99.15
FMC	708	9059	127	96	84.79	88.06	98.62
MB	31	9764	18	177	63.27	14.90	99.82
VMC	265	9496	98	131	73.00	66.92	98.98
CADCAM	3	9940	4	43	42.86	6.52	99.96
Zi	20	9947	8	15	71.43	57.14	99.92
Cera	20	9893	9	68	68.97	22.73	99.91
Pon	251	9725	8	6	96.91	97.67	99.92
IT	323	9600	34	33	90.48	90.73	99.65
HRT	2	9948	2	38	50.00	5.00	99.98
RT	25	9865	56	44	30.86	36.23	99.44
CRT	61	9843	29	57	67.78	51.69	99.71
HIT	46	9897	6	41	88.46	52.87	99.94
HET	20	9920	27	23	42.55	46.51	99.73
9 Conditions	10506	77443	883	1078	92.25	90.69	98.87
20 Conditions	9826	187033	1168	1773	89.38	84.71	99.38

表 A.2 状態診断モデルにおける Dataset2 の評価

	TP	TN	FP	FN	Precision	Recall	Specificity
N	6098	3595	331	177	94.85	97.18	91.57
C	73	9971	22	135	76.84	35.10	99.78
Per	57	9959	86	99	39.86	36.54	99.14
RCF	1213	8915	40	33	96.81	97.35	99.55
Implant	61	10139	0	1	100.00	98.39	100.00
In	1449	8279	198	275	87.98	84.05	97.66
Metal	1082	8945	110	64	90.77	94.42	98.79
CR	219	9556	92	334	70.42	39.60	99.05
AF	11	10155	15	20	42.31	35.48	99.85
Cro	1223	8779	87	112	93.36	91.61	99.02
FMC	642	9338	141	80	81.99	88.92	98.51
MB	42	10012	42	105	50.00	28.57	99.58
VMC	186	9813	63	139	74.70	57.23	99.36
CADCAM	2	10157	0	42	100.00	4.55	100.00
Zi	10	10165	0	26	100.00	27.78	100.00
Cera	5	10115	3	78	62.50	6.02	99.97
Pon	201	9992	4	4	98.05	98.05	99.96
IT	315	9815	57	14	84.68	95.74	99.42
HRT	6	10163	10	22	37.50	21.43	99.90
RT	6	10120	12	63	33.33	8.70	99.88
CRT	60	10085	21	35	74.07	63.16	99.79
HIT	39	10087	20	55	66.10	41.49	99.80
HET	15	10146	12	28	55.56	34.88	99.88
9 Conditions	10690	79444	825	850	92.84	92.63	98.97
20 Conditions	10028	191428	1024	1540	90.73	86.69	99.47

表 A.3 状態診断モデルにおける Dataset3 の評価

	TP	TN	FP	FN	Precision	Recall	Specificity
N	5613	3745	299	209	94.94	96.41	92.61
C	95	9590	26	155	78.51	38.00	99.73
Per	57	9609	38	162	60.00	26.03	99.61
RCF	1188	8585	52	41	95.81	96.66	99.40
Implant	80	9779	0	7	100.00	91.95	100.00
In	1437	7987	188	254	88.43	84.98	97.70
Metal	1010	8703	83	70	92.41	93.52	99.06
CR	301	9146	120	299	71.50	50.17	98.70
AF	11	9828	7	20	61.11	35.48	99.93
Cro	1275	8395	83	113	93.89	91.86	99.02
FMC	610	9056	87	113	87.52	84.37	99.05
MB	24	9712	5	125	82.76	16.11	99.95
VMC	278	9420	96	72	74.33	79.43	98.99
CADCAM	8	9814	4	40	66.67	16.67	99.96
Zi	22	9793	7	44	75.86	33.33	99.93
Cera	10	9777	4	75	71.43	11.76	99.96
Pon	220	9630	3	13	98.65	94.42	99.97
IT	326	9458	73	9	81.70	97.31	99.23
HRT	3	9823	7	33	30.00	8.33	99.93
RT	7	9799	12	48	36.84	12.73	99.88
CRT	121	9666	56	23	68.36	84.03	99.42
HIT	44	9776	27	19	61.97	69.84	99.72
HET	8	9828	1	29	88.89	21.62	99.99
9 Conditions	10291	76778	762	963	93.11	91.44	99.02
20 Conditions	9710	185079	934	1597	91.23	85.88	99.50

表 A.4 状態診断モデルにおける Dataset4 の評価

	TP	TN	FP	FN	Precision	Recall	Specificity
N	5674	3817	326	173	94.57	97.04	92.13
C	85	9712	70	123	54.84	40.87	99.28
Per	67	9731	41	151	62.04	30.73	99.58
RCF	1234	8645	70	41	94.63	96.78	99.20
Implant	65	9921	0	4	100.00	94.20	100.00
In	1528	8029	169	264	90.04	85.27	97.94
Metal	1118	8691	78	103	93.48	91.56	99.11
CR	259	9340	111	280	70.00	48.05	98.83
AF	21	9917	35	17	37.50	55.26	99.65
Cro	1295	8475	69	151	94.94	89.56	99.19
FMC	664	9075	109	142	85.90	82.38	98.81
MB	47	9791	41	111	53.41	29.75	99.58
VMC	193	9583	60	154	76.28	55.62	99.38
CADCAM	0	9937	2	51	0.00	0.00	99.98
Zi	6	9950	0	34	100.00	15.00	100.00
Cera	22	9887	10	71	68.75	23.66	99.90
Pon	217	9758	6	9	97.31	96.02	99.94
IT	283	9633	42	32	87.08	89.84	99.57
HRT	1	9968	1	20	50.00	4.76	99.99
RT	13	9904	21	52	38.24	20.00	99.79
CRT	71	9848	30	41	70.30	63.39	99.70
HIT	48	9887	23	32	67.61	60.00	99.77
HET	4	9949	3	34	57.14	10.53	99.97
9 Conditions	10448	77721	793	948	92.95	91.68	98.99
20 Conditions	9809	187311	1037	1643	90.44	85.65	99.45

表 A.5 状態診断モデルにおける Dataset5 の評価

	TP	TN	FP	FN	Precision	Recall	Specificity
N	5602	3829	351	195	94.10	96.64	91.60
C	100	9693	42	142	70.42	41.32	99.57
Per	65	9712	61	139	51.59	31.86	99.38
RCF	1312	8533	40	92	97.04	93.45	99.53
Implant	67	9906	4	0	94.37	100.00	99.96
In	1507	8007	205	258	88.03	85.38	97.50
Metal	1066	8732	127	52	89.35	95.35	98.57
CR	262	9301	89	325	74.64	44.63	99.05
AF	38	9882	35	22	52.05	63.33	99.65
Cro	1329	8430	99	119	93.07	91.78	98.84
FMC	686	9062	131	98	83.97	87.50	98.58
MB	4	9795	0	178	100.00	2.20	100.00
VMC	248	9458	161	110	60.64	69.27	98.33
CADCAM	1	9943	0	33	100.00	2.94	100.00
Zi	28	9927	10	12	73.68	70.00	99.90
Cera	14	9895	8	60	63.64	18.92	99.92
Pon	224	9735	13	5	94.51	97.82	99.87
IT	310	9604	34	29	90.12	91.45	99.65
HRT	0	9932	0	45	0.00	0.00	100.00
RT	0	9907	0	70	0.00	0.00	100.00
CRT	69	9827	48	33	58.97	67.65	99.51
HIT	56	9868	28	25	66.67	69.14	99.72
HET	10	9932	4	31	71.43	24.39	99.96
9 Conditions	10516	77449	849	979	92.53	91.48	98.92
20 Conditions	9852	186869	1152	1667	89.53	85.53	99.39