

RNN モデルを用いた歯科用器具の 経時的存在パターン検出に基づく処置内容の分類

Recurrent neural network model for classifying dental treatment

岡 真太郎¹ 野崎 一徳² 林 美加子¹

Oka Shintaro¹, Nozaki Kazunori², Hayashi Mikako¹

¹ 大阪大学大学院歯学研究科 口腔分子感染制御学講座 (歯科保存学教室)

¹Department of Restorative Dentistry and Endodontology,
Osaka University Graduate School of Dentistry

² 大阪大学歯学部附属病院 医療情報室

² Division for Medical Information, Osaka University Dental Hospital

Abstract: In general dental clinic, dentists need to treat many patients safely within a restricted time. However, the excessive of dental routine work may lead to occur medical incident. Thus, we need a realistic solution to keep safety and quality of dental routine work. Taking it into account for the trend found in general society, we try to develop a general dental AI which can be in charge of recognition of dental treatment and suggestion of optimal method based on statistical analyses. To achieve the development of the general dental AI, at first, we need to gather as much as primary clinical sensing data such as video, sound and signal data. In our future dental treatment work, we aim at free from writing the medical records and incidents with AI. For the first step of the general dental AI, we developed the object detection using deep learning YOLO v3, with small IoT camera device, which stored large movie data via wireless network. Stored movie data was analyzed by YOLO v3 which recognized several dental instruments at high accuracy, which enables to record the sequence of dental instruments used in dental operations for each patient. We got 49 of dental instruments series data in the clinic. The sequence of dental instruments can be used to estimate the variety of treatment contents by LSTM. Constructing LSTM led to classify all series data into the largest number of label, which meant the failure of learning. Testing the LSTM with small number of simple samples, the accuracy was lower than large number of samples. Therefore, it was found that a larger number of samples is necessary to estimate the dental treatment contents using the series information of dental instruments.

1 背景

歯科医師法第23条において歯科医師は患者を診療したら遅滞なく経過を記録することが義務付けられている。現在歯科医師は、診療途中もしくは診療終了後に手動でこれを記録しているが、診療をありのまま記録しているわけではない。

コンピュータで術者を支援する仕組みとしては、手術室での器具管理の自動化などが試みられている^[1]。また耳鼻科においては診療行為、すなわち用いた器具と部位等から診療録を自動的に作成する技術の開発が以前より行われている^[2,3]が、歯科において

はまだ行われていない。診療内容を記録する際には処置中の口腔内を撮影することがまず思いつくが、診療台に対する口腔の位置は患者の体格などにより変化し、それに対応するためには口腔周囲、すなわち顔面・顔貌も同時に録画することになる。しかし多くの患者は顔貌が録画されることを拒否する傾向にあるため、岡ら(2018)は、歯科診療で用いる診療台上の器具トレイ(バット)を撮影する手法を提案した^[4]。歯科診療は外科的処置がメインであり、その特徴として診療目的に応じて複数の処置を組み合わせ、その処置のために複数の器具をバットから取り出して使うということがあげられる。すなわち用

いた器具とその順列から診療内容を推定することができると考え、診療中の器具の使用回数から、歯周処置 (P)、う蝕処置 (CR)、根管貼薬処置 (RCT)、根管充填処置 (RCF) の 4 種類の診療内容推定を行い、その結果 overall accuracy が 61.2% となることを報告した^[4]。

本研究では器具の使用順序から診療内容の推定を行い、器具使用回数での推定との比較をすることを目的とする。

2 実験方法

2.1 診療中のバットの撮影

大阪大学大学院歯学研究科・歯学部及び歯学部附属病院倫理審査委員会承認 (H29-E23-1) に基づき、同意を得られた患者への処置中のバットの様子を録画した (図 1)。

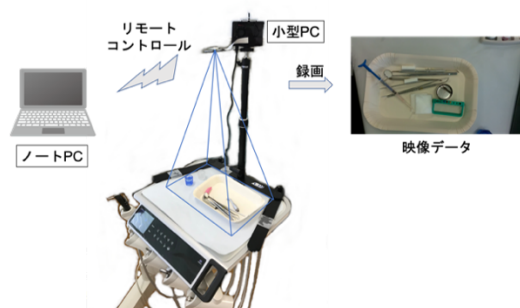


図 1. 診療中の録画方法の模式図

バットを置く診療台にデジカメ撮影スタンド (ハクバ写真産業製) を固定し、その上に Raspberry Pi 3 Model B と専用カメラモジュールを固定した。これをノート PC でリモート操作し、Raspberry Pi に接続した USB メモリに Full HD 30 fps でバットの様子を診療中の一次情報として一診療単位で録画し、ファイルとして保存した。

2.2 診療一次情報の処理

上記の方法で得られた映像データからバット上の器具の有無情報を得るために、映像に対して歯科用器具の画像認識を行った。その際画像認識ソフトウェアとして YOLO v3^[5] を用いた。YOLO v3 はシングルショット系の一般物体検出アルゴリズムで、検出物体の位置とクラスを同時に推定するものである。これを用いて、保存科外来診療室でよく用いられる 22 種類の器具と術者の手およびガーゼの学習を 1,522 枚の画像で行い、得られた判別器を用いて録画した映像に対して画像認識を行い、検出される器具の種類と数の情報を得た。

2.3 検出器具についての情報処理

動画ファイルである映像データのフレーム数を N とし、 $j (= 0, \dots, N - 1)$ に含まれるバット内の器具 i の個数を $x \in \mathbb{Z}$ としたとき、 $v_{ij} = x$ として表現した。映像データ全体に含まれる器具集合 \mathbf{D} は、フレーム j での器具集合を \mathbf{V}_j とすると、次式で表現される。

$$\begin{aligned} \mathbf{D} &= \{v_{ij} | i = 0, \dots, 23, j = 0, \dots, N - 1\} \\ &= \{\mathbf{V}_j | j = 0, \dots, N - 1\} \end{aligned} \quad (1)$$

この時系列データ \mathbf{D} に対して、術者の手が検出された時を処置の区切りとして、時系列データを分割した (図 2)。図 2 の黒く塗りつぶされた部分は術中に術者が器具を取る際に撮影された手が検出された期間である。この期間を境にバット上の器具の種類が変化していると考えられるため、その時を境として \mathbf{D} を分割し、 $\bar{\mathbf{D}} (= \{d_k | k = 0, 1, \dots, K\})$ を得た。ただし、 K は術者の手が映像に現れた回数とする。また、 $\mathbf{n} (= \{n_k | k = 0, 1, \dots, K\})$ は、 $\bar{\mathbf{D}}$ それぞれの区間におけるフレーム数を表す。

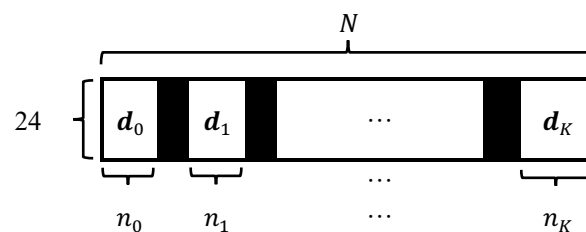


図 2: 時系列データ \mathbf{D} の分割

手が検出されてから次に手が検出されるまでの間に検出されている器具数の、フレームあたりの平均値器具集合 $\bar{\mathbf{D}}$ は区間 k での器具集合を $\bar{\mathbf{V}}_k$ とすると次式で表現される。

$$\begin{aligned} \bar{\mathbf{D}} &= \frac{\mathbf{d}}{\mathbf{n}} \\ &= \{\bar{\mathbf{V}}_k | k = 0, 1, \dots, K\} \end{aligned} \quad (2)$$

こうして得られた 2 種類の時系列データ \mathbf{D} と $\bar{\mathbf{D}}$ を各診療を表現する時系列データとした。

2.4 時系列データの学習

実際に保存科外来診療室で行われた 49 回の診療の時系列データ \mathbf{D} と $\bar{\mathbf{D}}$ に 4 種類の診療内容 {P: 10 回・CR: 7 回・RCT: 21 回・RCF: 11 回} の正解ラベルを結びつけ、LSTM(Long short-term memory)を用いた学習と Leave-one-out 交差検証を行った。Nvidia-Docker 上で Tensorflow-gpu(1.14.0)、CUDA(10.0)、cuDNN(7.6)、ライブラリとして Keras(2.3.0)を用いた。

LSTM は CuDNNLSTM (入力次元数 24、出力次元数 3) を用い、分類のための全結合層 (入力次元数 3、出力次元数 4、活性化関数 softmax) とした (図 3)。なお学習時のデータ割合を 80% とし、検証データ割合は残りの 20% とした。

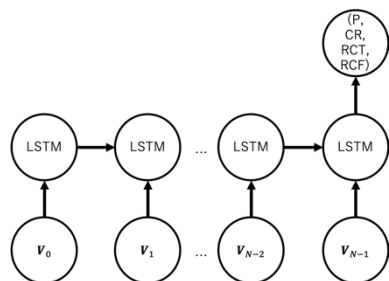


図 3. LSTM ネットワークの模式図

3 結果

表 1 は、観察された 49 診療の時系列データと正解ラベルとの組み合わせを LSTM で学習させた判別器での Leave-one-out 交差検証結果である。真陽性率は、各正解ラベルにおいて正しく陽性と判定した割合を示し、偽陰性率は誤って陰性と判定した割合を示す。

表 1: 時系列データからの診療内容推定手法の Leave-one-out 交差検証結果
数字は分類された診療数を示している。
{P: 歯周処置、CR: う蝕修復処置、RCT: 根管貼薬処置、RCF: 根管充填処置}

正解ラベル	RCT	0	0	11	0	0.00%	100%
	RCT	0	0	21	0	100%	0.00%
	CR	0	0	7	0	0.00%	100%
	P	0	0	10	0	0.00%	100%
		P	CR	RCT	RCF	真陽性率	偽陰性率
予測ラベル							

LSTM ネットワークはどの診療に対しても RCT ラベルを予測結果として出力した。

式 (3) で表される overall accuracy は 42.9% となった。ただし $TP_{(X)}$ は処置ラベル $X (= \{P, CR, RCT, RCF\})$ についての真陽性数を表し、 A は全診療数 (= 49) を表す。

$$\text{overall accuracy} = \frac{TP_{(P)} + TP_{(CR)} + TP_{(RCT)} + TP_{(RCF)}}{A} \quad (3)$$

4 考察

LSTM モデルでは必ずサンプル数が最多の項目を出力する結果となり、正しく学習が行えていないことが判明した。この結果は図 4 のように LSTM を 2 層化しても同じであった。

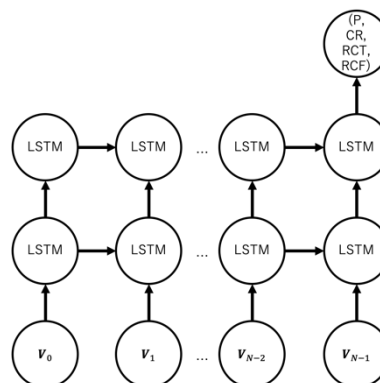


図 4. 2 層 LSTM ネットワークの模式図

今回用いた診療データは各ラベル間でのサンプル数に偏りがあるため、学習時に損失関数が適切に機能せず、最適化が十分行われなかった可能性が考えられた。そこで、学習時に例数の少ないラベルの重みを増すために LSTM ネットワークの class_weight に例数の逆数と等しくなるようにした初期値を与えた結果、サンプル数が最も少ない CR が出力され、現実的な解が得られなかった。

次に RCT ラベル数を減らし、{P: 10 回・CR: 7 回・RCT: 12 回・RCF: 11 回} となるデータで学習および評価を行ったが、これもすべて RCT ラベルと出力される結果となった。

原因を探索するため、単純なサンプルを作成し同様の LSTM ネットワークで基礎実験を行った。xy 平面上のランダムな座標を始点とし、その点から x 座標および y 座標の値が 1 もしくは -1 ずつ単調に変化する系列データ (長さ 20) を作成し、それを 4 クラス分類する LSTM ネットワークの学習および評価を行った (図 5)。

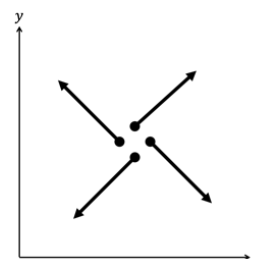


図 5. 単純なサンプルのイメージ
始点からの座標変化は 4 クラスに分類できる。

サンプル数を各クラス等しく 10 個とした場合、overall accuracy は 90%であったが、あるひとつのクラスのサンプル数のみを 20 個とすると overall accuracy は 50%に低下し、そのクラス以外の真陽性率が低下した。しかしこの傾向はサンプル数を 100 個単位に増大させることで改善し、サンプル数をあるクラスは 200 個、その他は 100 個とした場合、overall accuracy は 99.5%となった。

以上のことから、器具の使用順序から歯科診療を分類するためには、器具の使用回数のみから分類する場合と比較して、より多くのサンプル数が必要であることがわかった。

謝辞

本研究の一部は、日本電気株式会社共同研究費「スマートデンタルホスピタルに関する研究」の助成による。

参考文献

- [1] Glaser, B., Dänzer, S. & Neumuth, T.: Intra-operative surgical instrument usage detection on a multi-sensor table, *International Journal of Computer Assisted Radiology and Surgery* Volume 10, Issue 3, pp 351–362, (2015)
- [2] 江口郁子, 佐藤友正, 森武俊: 診療行為記述化のための視覚理解機能, *日本ロボット学会誌*, Vol. 15, No. 5, pp766~772, (1997)
- [3] 矢入(江口)郁子, 佐藤友正, 森武俊: 計算機理解に基づく診療行動の蓄積・提示支援システムに関する研究, *日本ロボット学会誌*, Vol. 20, No. 4, pp437~445, (2002)
- [4] 岡真太郎, 野崎一徳, 林美加子: 使用歯科器具判別器を用いた処置の言語化とその分類, *医用人工知能研究会*, SIG-AIMED-007-01, 2019年3月7日, 神奈川.
- [5] Joseph Redmon, Ali Farhadi. YOLOv3: An Incremental Improvement. University of Washington. [<https://pjreddie.com/media/files/papers/YOLOv3.pdf> (cited 2018-Aug-30)]