

# Deep Learningを用いたJellyfish Sign自動識別システムの開発

川野 晃輔, 飯島 直也, 南木 望, 古居 彬, 曾 智 (広島大学), 森 裕紀 (早稲田大学),  
早志 英朗 (九州大学), 久米 伸治 (広島大学病院), 辻 敏夫 (広島大学)

## Development of Automatic Classification System for Jellyfish Sign Using Deep Learning

Kosuke Kawano, Naoya Iijima, Nozomi Nanki, Akira Furui, Zu Soh (Hiroshima University),  
Hiroki Mori (Waseda University), Hideaki Hayashi (Kyushu University),  
Shinji Kume (Hiroshima University Hospital), Toshio Tsuji (Hiroshima University)

Abstract : This paper proposes an automatic classification system for jellyfish sign that is a kind of carotid plaque. The proposed system uses a deep neural network consisting of a convolutional neural network and a probabilistic recurrent neural network to calculate the posterior probabilities for positive and negative jellyfish signs. The validity of the proposed system is demonstrated via classification experiments.

### 1. はじめに

脳梗塞や動脈硬化の可能性が高い患者の早期発見や経過観察を目的に、非侵襲かつリアルタイムに計測可能な頸動脈超音波検査が広く行なわれている。この検査には、局所的に隆起したプラークとよばれる病変の評価を行なう検査項目が含まれている。プラークは線維性被膜の厚さなどの特徴によって安定性プラークと不安定性プラークに分類され、不安定性プラークは線維性被膜が破綻することで脳梗塞を引き起こす可能性がある。そのため、不安定性プラークの早期発見が重要である。

不安定性プラークの中でも、血流の拍動に応じてプラークの表面が浮き沈みする Jellyfish Sign が近年注目されている<sup>1)</sup>。Jellyfish Sign を有する病変はプラークが破綻している可能性が高く、この特徴が確認された症例のうち54.8%が脳梗塞の発症につながったと報告されている<sup>2)</sup>。しかし、Jellyfish Sign を発見するためには拍動や呼吸に伴う血管全体の動きとプラークの動きを分離して認識する必要があり、診断には高度な専門性が要求される。

この問題に対して、超音波画像内のプラークが Jellyfish Sign を呈するかどうかを客観的に評価できれば有用である。従来、Baroncini ら<sup>3)</sup>や Salem ら<sup>4)</sup>によって画像処理を用いた不安定性プラークの定量評価が行われている。しかしながら、これらはプラークの輝度や均一性、表面形態といった静的特性のみに着目したものであり、Jellyfish Sign のようなプラークの動的特性を評価できるものはなかった。

そこで本稿では、頸動脈超音波検査において計測した超音波動画画像から、ディープニューラルネットワークを用いて Jellyfish Sign の自動識別を行なうシステムを提案する。提案システムでは、計測した超音波動画画像からプラーク

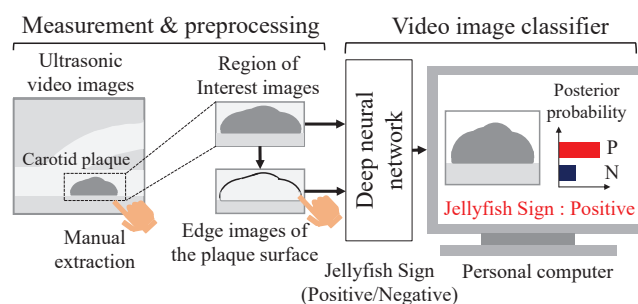


Fig. 1: Overview of the proposed system

動画画像とプラーク表面のエッジ画像から構成される2チャンネル動画画像を作成する。そして、これを畳み込みニューラルネットワークおよびリカレント型の確率ニューラルネットワークで構成されるディープニューラルネットワークへと入力することで、動画画像内のプラークについて Jellyfish Sign の陽性/陰性の自動識別が可能である。

### 2. Jellyfish Sign 自動識別システム

Fig. 1 に、Jellyfish Sign 自動識別システムの概略図を示す。提案システムは信号計測部、前処理部、動画画像識別部から構成される。

#### 2.1 信号計測部

信号計測部では、頸動脈超音波検査によりプラークの動きを計測する。頸動脈へ超音波を当てることでグレースケールの B モード画像 (Fig. 2(a)) を取得し、サンプリング周波数  $f_s$  [Hz] で PC に保存する。このとき検査者には、動画画像の初期フレームにおいて Jellyfish Sign の有無

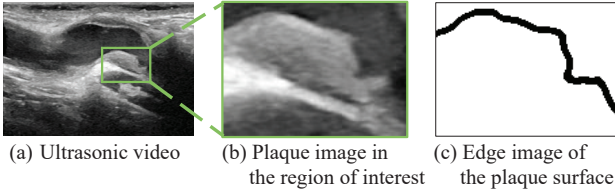


Fig. 2: Examples of captured images

を検査したい領域として、プラーク全体および血流部とプラークの境界が含まれる  $W \times H$  [pixel] の関心領域 (Fig. 2(b)) を手動で選択させる。さらに、関心領域画像においてプラーク表面を曲線でなぞらせ、プラーク表面のエッジ画像 (Fig. 2(c)) を作成する。

## 2.2 前処理部

前処理部では、計測したプラークを含む超音波動画画像およびプラーク表面のエッジ画像に対して画像処理を施し、動画画像識別部に入力する動画画像を作成する。まず、撮影された超音波動画画像の初期フレームから  $T$  [frame] までを解析対象として切り出す。ここで、切り出された超音波動画画像には拍動や呼吸に伴う血管全体の動きが含まれているため、Jellyfish Sign の動きと分離することを目的にテンプレートマッチングを行なう。関心領域画像をテンプレート画像に設定し、動画画像の各フレームにおいてテンプレート画像と最も相関係数が高い領域を切り出す。この処理を全フレームに対して行なうことで血管全体の動きを除いたプラーク動画画像を作成する。次に、プラーク動画画像およびプラーク表面のエッジ画像を  $W' \times W'$  [pixel] にリサイズする。そして、プラーク動画画像とプラーク表面のエッジ画像による 2 チャンネル動画画像  $x_{i,j,k}(t)$  ( $i = 1, \dots, W', j = 1, \dots, W', k = 1, 2$ ) を構成し、これを動画画像識別部への入力とする。

## 2.3 動画画像識別部

### 2.3.1 ネットワーク構造

動画画像識別部は、Convolutional Neural Network<sup>6)</sup> (以下、CNN) および Recurrent Log-Linearized Gaussian Mixture Network<sup>7)</sup> (以下、R-LLGMN) により構成されており (Fig. 3)、入力された 2 チャンネル動画画像から Jellyfish Sign の有無を識別する。

CNN は畳み込み層とプーリング層を含む順伝搬型ニューラルネットワークであり、特徴抽出器の役割を持つ。畳み込み層とプーリング層は対になっており、入力画像  $x_{i,j,k}(t)$  に対し繰り返し適用することで、特徴量の抽出を行なう。畳み込み層では、入力された画像と同チャンネル数のフィルタにより類似した濃淡パターンを検出することで、

特徴的な濃淡構造を画像から抽出する。いま、 $l$  層における画像を  $y_{i,j,k,r}^{(l)}$  とおくと、これを畳み込み層に入力した際に得られる特徴マップ  $u_{i,j,r}$  は、以下のようになる。

$$u_{i,j,r} = f \left( \sum_{k=1}^K \sum_{p=1}^U \sum_{q=1}^U w_{p,q,k,r} y_{i+p,j+q,k,r}^{(l)} + b_{i,j,r} \right) \quad (1)$$

ここで、 $w_{p,q,k,r}$  ( $r = 1, \dots, R$ ;  $R$  はフィルタ数) はサイズ  $U \times U$  の畳み込みフィルタの重みパラメータ、 $K$  は入力画像  $y_{i,j,k,r}^{(l)}$  のチャンネル数、 $b_{i,j,r}$  はバイアス、 $f(\cdot)$  は活性化関数である。また、プーリング層では特徴マップに対し、局所領域内における最も大きな画素値に着目した最大プーリングを適用することで、抽出された特徴量の位置感度を低下させ、出力の不変性を獲得する。これらの処理を 3 回繰り返し、最後に全結合層に通すことで特徴ベクトル  $\mathbf{z} \in \mathbb{R}^d$  ( $d$  は CNN の出力次元数) を算出する。

R-LLGMN は混合正規分布モデルと隠れマルコフモデルを内包したリカレントニューラルネットワークであり、入力の時系列特性を考慮したパターン識別が可能である。R-LLGMN は 5 層構造で、第 1 層目と第 2 層目の間に重みパラメータが存在し、第 3 層目と第 4 層目の間にリカレント構造を有する。この構造の中に、クラス数  $C$ 、各クラスの状態数  $K_c$  ( $c = 1, \dots, C$ ) の連続型隠れマルコフモデルを含んでいる。なお、提案システムにおいては Jellyfish Sign 陽性/陰性の 2 クラス分類を行なうため、クラス数  $C = 2$  となる。各状態における出力確率の密度関数は、 $M_{c,k'}$  ( $k' = 1, \dots, K_c$ ) 個のコンポーネントで近似する。はじめに、CNN で算出された時刻  $t$  における特徴ベクトル  $\mathbf{z}(t)$  ( $t = 1, \dots, T$ ) に対して非線形変換を行ない、得られた  $\mathbf{z}'(t) \in \mathbb{R}^H$  ( $H = 1 + d(d+3)/2$ ) を第 1 層への入力とする。第 2 層のユニットは、第 1 層の出力に重みパラメータを掛け合わせたものを受け取り、指数関数による非線形演算を行ない次の層へ伝える。第 3 層と第 4 層の間ではリカレント構造に基づき時系列情報の処理を行なう。そして、第 5 層では最終時刻  $T$  において各識別対象クラス  $c$  に対する事後確率  $P(c|\mathbf{x}(T))$  を算出する。事後確率  $P(c|\mathbf{x}(T))$  が最大となるクラスを識別結果とすることで、Jellyfish Sign の有無を自動識別可能である。

### 2.3.2 ネットワークの学習則

いま、 $N$  サンプルの入力画像  $\mathbf{X}^{(n)}$  ( $n = 1, \dots, N$ ) に対して、教師ベクトル  $\mathbf{t}^{(n)} = [t_1^{(n)}, t_2^{(n)}]^T$  が与えられた場合を考える。ここで、 $\mathbf{t}^{(n)}$  は one-hot ベクトル形式をとる。ネットワーク全体の評価関数として、次の交差エントロピー誤差を定義した。

$$E = - \sum_{n=1}^N \sum_{c=1}^2 t_c^{(n)} \ln P(c|\mathbf{x}^{(n)}(T)) \quad (2)$$

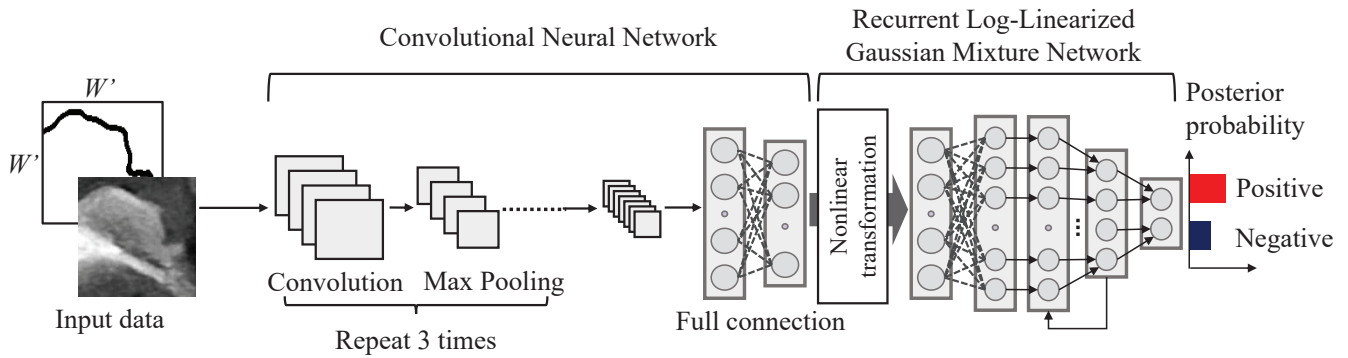


Fig. 3: Structure of the proposed deep neural network

Table 1: Structure of the CNN. The abbreviations CONV, MP, and FC denote convolutional layer, max pooling, and fully connected layer, respectively.

Layer	Filter		Output dimensions	Activation function
	Size	Number		
Input	-	-	$64 \times 64 \times 2$	-
CONV	$5 \times 5$	100	$64 \times 64 \times 100$	ReLU
MP	$2 \times 2$	-	$32 \times 32 \times 100$	-
CONV	$5 \times 5$	150	$32 \times 32 \times 150$	ReLU
MP	$2 \times 2$	-	$16 \times 16 \times 150$	-
CONV	$3 \times 3$	200	$16 \times 16 \times 200$	ReLU
MP	$2 \times 2$	-	$8 \times 8 \times 200$	-
FC	-	-	300	ReLU

この評価関数が最小となるよう、CNN と R-LLGMN の重みパラメータを通時的誤差逆伝播法に基づき一括で更新する。

### 3. 実験

#### 3.1 方法

提案システムの有効性を確認するため、頸動脈超音波検査で計測された超音波動画像を用いて識別精度検証実験を行なった。使用データは、プラークが確認された超音波動画像のうち、診療放射線技師（以下、専門家）1名の判断に基づき分類した Jellyfish Sign 陽性 24 病変および陰性 24 病変の計 48 病変とした。また、計測されたプラーク動画像の画像サイズは  $W = 174 \pm 51$  pixel,  $H = 102 \pm 43$  pixel, サンプリング周波数は  $f_s = 30$  Hz であった。全ての計測実験はヘルシンキ宣言に則り、広島大学疫学研究倫理審査委員会の承認のもと、患者の同意を得た上で実施した。

動画像識別部における各パラメータはそれぞれ  $T = 60$  frame,  $W' = 64$  pixel に設定した。また、提案システムにおける CNN の構造を Table 1 に示す。第 1 層を入力層、第 2, 4, 6 層を畳み込み層、第 3, 5, 7 層をプーリン

グ層、第 8 層を全結合層とし、活性化関数には Rectified Linear Unit (ReLU) を用いた。R-LLGMN においては、コンポーネント数を  $M_{c,k} = 1$ , 状態数を  $K_c = 2$  に設定した。ネットワーク学習時には、Data Augmentation として学習データに対し垂直・水平方向の反転と回転 ( $0-90^\circ$ ) を行ない、学習の繰り返し回数 (エポック数) は 20 回に設定した。

識別では、専門家による評価結果と提案システムによる識別結果を比較し、5-fold 交差検証に基づき精度の検証を行なった。識別精度の指標として、システム全体の識別精度を表す識別率 (Accuracy), Jellyfish Sign 陽性の症例の識別精度を表す感度 (Sensitivity), Jellyfish Sign 陰性の症例の識別精度を表す特異度 (Specificity) を用いた。なお、交差検証は初期値を変え 5 回繰り返し、各精度指標の平均値を算出した。比較として、提案システムにおける R-LLGMN を一般的なりカレントニューラルネットワークである Long Short-Term Memory<sup>8)</sup> (以下、LSTM) に置換したモデルを用いて同様に識別精度を評価した。さらに、プラーク表面のエッジ画像の有無が識別に与える影響を調べるため、提案システムに対する入力としてプラーク動画像のみの 1 チャンネル動画像を用いた場合の精度も同様に算出した。

#### 3.2 結果

Fig. 4 に、エポック数ごとの評価関数の推移の 1 例を各手法について示す。また、Fig. 5 に、5 回の解析における識別率、感度、特異度の平均値と 95 % 信頼区間を示す。識別率、感度、特異度はそれぞれ、提案システムでは 85.8%, 85.8%, 85.8%, 入力データをプラーク動画像のみにした場合では 71.7%, 73.3%, 70.0%, CNN+LSTM では 83.8%, 81.7%, 85.8% であった。また、Fig. 6 に Gradient-weighted Class Activation Mapping (以下、Grad-CAM)<sup>9)</sup> を用いて、Jellyfish Sign 陽性の識別に寄与する領域を可視化した結果を示す。この Grad-CAM は、ニューラルネットワークの重みの勾配を



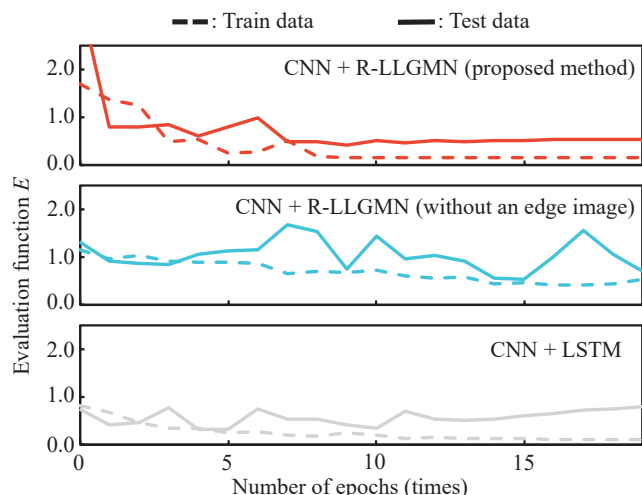


Fig. 4: Transition of evaluation function for train and test data

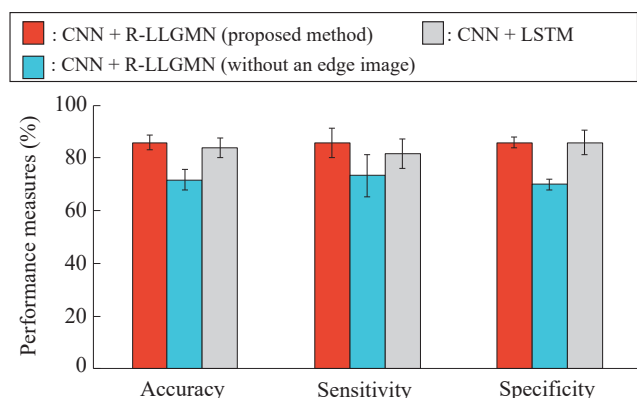


Fig. 5: Results of Jellyfish Sign classification for each method

用いて、画像中の識別への寄与度の高い領域を視覚的に活性化させる手法である。ここでは、提案法においてプラーク表面のエッジ画像を含めた場合と含めない場合のそれぞれについて、陽性の識別に寄与した領域を可視化した結果の一例を示している。

### 3.3 考察

Fig. 4 より、提案システムでは、学習が進むにつれテストデータに対する評価関数の値が減少していることが確認できる。これは、提案システムがテストデータに対する汎化能力を獲得できていることを意味する。一方、プラーク動画画像のみを用いて学習したモデルでは、テストデータに対する評価関数の変動が大きく、識別に有効な学習が行なえていない。また、一般的なりカレントニューラルネットワークである LSTM を用いたモデルでは、13 エポックを終えたあたりからテストデータに対する評価関数値が上昇しており、学習データに対し過学習していることが確認で

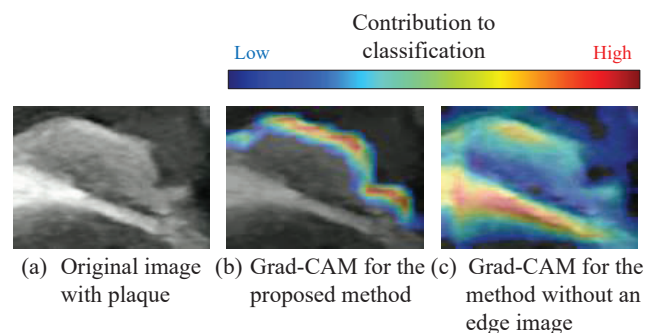


Fig. 6: Examples of Grad-CAM visualizations

きる。

Fig. 5 より、提案システムは全ての識別精度指標において比較手法と同等以上の精度で Jellyfish Sign の有無を識別できていることが分かる。特に、入力データにプラーク表面のエッジ画像を用いなかったモデルと比較し、全ての精度指標が 10% 以上高くなっている。この理由として、プラーク表面のエッジ画像を入力データに加えることで、Jellyfish Sign の特徴であるプラーク表面の浮き沈みをうまく捉えることができたためであると考えられる。これは、Fig. 6 における Grad-CAM の結果からも読み取ることができる。提案システムでは Jellyfish Sign が存在するプラークの表面に沿って活性化領域が広がっているが (Fig. 6(b))、プラークの表面のエッジ画像を用いなかった場合、血管壁のようなプラーク表面以外で高い輝度を示した領域も活性化していることが分かる (Fig. 6(c))。

また、提案法はもう一方の比較法である LSTM を用いたモデルに対しても僅かではあるが高い精度を示した。この比較手法で用いた LSTM は、提案システムにおける R-LLGMN に比べ、単位時間あたりの情報の処理により多くのユニットとパラメータが用いられる複雑なネットワーク構造になっている。そのため、今回の 48 症例のような少数のデータセットに対しては汎化能力の獲得が困難であったと考えられる。一方で、提案システムで使用した R-LLGMN には入力データの分布を近似する統計構造が組み込まれており、少数の学習データに対して高い識別能力を実現可能である<sup>7)</sup>。そのため、今回のデータセットに対しては R-LLGMN を用いた提案システムの方が高い汎化能力を獲得できたと考えられる。

以上より、提案システムを用いることで、比較的高精度に Jellyfish Sign の有無を識別できることが明らかとなった。しかしながら、Jellyfish Sign 陽性の識別精度である感度に注目すると、識別率や特異度に比べ信頼区間が大きい傾向にある。これは、試行ごとの感度の変動が大きく、Jellyfish Sign 陽性の識別が不安定であることを意味している。そのため、血管の拍動など Jellyfish Sign 陽性のプラークの運動に影響を与えられる情報を識別に利用

することで，より安定した識別を行なえる可能性がある．

#### 4. まとめと今後の課題

本研究では，頸動脈プラークにおける Jellyfish Sign の有無を識別可能なシステムを提案した．提案システムでは，頸動脈超音波検査により計測した超音波動画像を多層構造のディープニューラルネットワークへと入力することで，動画像内のプラークについて Jellyfish Sign の陽性／陰性を識別可能とした．実験では，Jellyfish Sign 陽性 24 例を含む 48 症例の頸動脈超音波動画像に対して解析を行い，85.8%の精度で Jellyfish Sign が識別可能であることを確認した．

今後は症例数を増加させるとともに，ネットワーク構造を改良することにより識別精度の向上を行なう予定である．また，Jellyfish Sign の有無の識別だけでなく，陽性領域の自動検出を行なうことで，さらなる診断支援を目指す．

#### 参考文献

- 1) 久米伸治ら： “Jellyfish sign ( 内頸動脈可動性 plaque ) の病理とその臨床像”， *Neurosonology*, 20, pp. 21–24 (2007)
- 2) S. Kume *et al.*: “Vulnerable carotid arterial plaque causing repeated ischemic stroke can be detected with B-mode ultrasonography as a mobile component: Jellyfish sign”, *Neurosurgical Review*, 33, pp. 419–430 (2010)
- 3) LAV.Baroncini *et al.*: “Ultrasonic tissue characterization of vulnerable carotid plaque: correlation between videodensitometric method and histological examination”, *Cardiovascular Ultrasound*, 4, 32 (2006)
- 4) M.K.Salem *et al.*: “Identification of patients with a histologically unstable carotid plaque using ultrasonic plaque image analysis”, *European Journal of Vascular and Endovascular Surgery*, 48, pp. 714–715 (2014)
- 5) M. Hennerici *et al.*: “Spontaneous progression and regression of small carotid atheroma”, *The Lancet*, 325, pp. 1415–1419 (1985)
- 6) Y. LeCun *et al.*: “Backpropagation applied to handwritten zip code recognition”, *Neural Computation*, 1, pp. 541–551 (1989)
- 7) T. Tsuji *et al.*: “A recurrent log-linearized Gaussian mixture network”, *IEEE Transactions on Neural Networks*, 14, 2, pp. 304–316 (2003)
- 8) S. Hochreiter and J. Schmidhuber : “Long short-term memory”, *Neural Computation*, 9, pp. 1735–1780 (1997)
- 9) R. Selvaraju *et al.*: “Grad-CAM: Visual explanations from deep networks via gradient-based localization”, *The IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626 (2017)