

RGBカメラを用いた頸髄症スクリーニング手法の提案

松井 良太*¹ 小山 恭史*² 藤田 浩二*²
齋藤 英雄*¹ 杉浦 裕太*¹

A Screening Method for Cervical Myelopathy Using a Camera

Ryota Matsui*¹, Takafumi Koyama*², Koji Fujita*², Hideo Saito*¹ and Yuta Sugiura*¹

Abstract – Cervical myelopathy (CM) is a pathology caused by cervical spinal cord compression. Spinal surgeons often use the 10-sec grip and release (G&R) test to screen hand disorders, a typical symptom of CM. We propose a screening method for CM based on videos of the G&R test and machine learning. Each patient holds their hand above a smartphone to record the G&R movement as a video with the built-in camera. We use an image-processing framework to obtain feature values of the hand movement. A support vector machine classifier estimates if these feature values suggest any characteristics of CM patients. We conducted a user experiment on 20 CM patients and 15 controls to evaluate our method. As a result, sensitivity, specificity, and area under the curve were 90.0%, 93.3%, and 0.947, respectively. This performance is higher than the conventional methods.

Keywords : screening system, hand tracking, cervical myelopathy

1. はじめに

脊椎(背骨)の中を通る脊髄は、脳と身体の各部をつなぐ中枢神経系を構成している。脊髄が何らかの原因によって圧迫された病態を脊髄症というが、特に脊髄のうち首付近の部位を指す頸髄が圧迫された場合は頸髄症とよばれる。頸髄症は手指のしびれや巧緻運動障害など、さまざまな症状を引き起こす。頸髄症の診断では、MRIなどを用いて頸髄の圧迫を視覚的に確認する画像検査のほかに、症状の有無を簡易に判断するスクリーニング手法が活用される。

頸髄症をスクリーニングする代表的な手法の1つが10秒テストである^[1]。このテストでは、手指の開閉運動(グーパー運動)を10秒間、可能な限り高速に繰り返す。医師は10秒間の反復回数や、運動の敏捷性などを評価して頸髄症症状の有無を判断する。スクリーニングの実施にあたり何らかの装置を必要とせず、非常に手軽に実施できるため、現在でも広く使われている手法である。またスクリーニングの性能向上などを目的に、10秒テストの手法を拡張する後継研究も多くなされてきた。

本研究では、グーパー運動の様子を録画した映像を、機械学習を用いて分析する頸髄症スクリーニング手法を提案する。机の上に置かれたスマートフォンの直上で

グーパー運動を実施し、その様子をスマートフォンの内蔵カメラで録画する。次に画像処理を用い、この映像から手指の動きを特徴量として抽出する。この特徴量は前処理の後に、サポートベクタマシン(SVM)分類モデルを用いて患者と健常者の2クラスに分類される。

2. 関連研究

2.1 頸髄症の検査・スクリーニング手法

人工知能や機械学習を、頸椎外科で用いられる検査やスクリーニングに活用するための検討が近年進められている^[2]。Meraliらは頸髄症の検査に関連し、MRI画像を機械学習で分析することにより、頸髄の圧迫を自動で発見する手法を提案した^[3]。

10秒テストやグーパー運動に着目した同様の検討も進められてきた。加藤らは、Leap Motion Controllerによるグーパー運動の計測と機械学習を組み合わせ、頸髄症症状の有無や重症度を推定する手法を提案した^[4]。Suらは、センサ付きのグローブ型デバイスを用いてグーパー運動を計測し、センサ値の変動をもとに頸髄症をスクリーニングするシステムを開発した^[5]。

本研究ではこれらの研究と同様に、10秒テストやグーパー運動をベースとした頸髄症スクリーニング手法について提案する。他方、提案手法ではこれらの関連研究と異なり、広く普及したスマートフォンを計測デバイスとして利用している。このことから提案手法は、スクリーニングの環境が手軽に構築できる特徴をもつ。

*1: 慶應義塾大学 大学院理工学研究科

*2: 東京医科歯科大学 大学院医歯学総合研究科

*1: Graduate School of Science and Technology, Keio University

*2: Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University

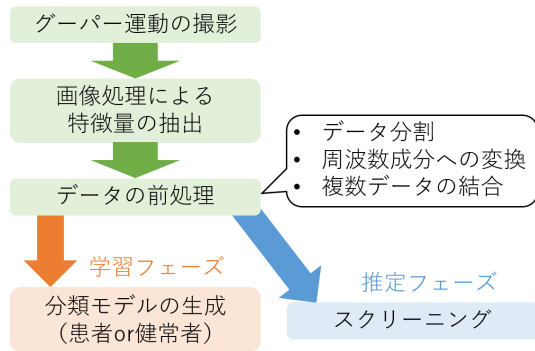


図1 提案手法の流れ
 Fig.1 Flow of proposed method.

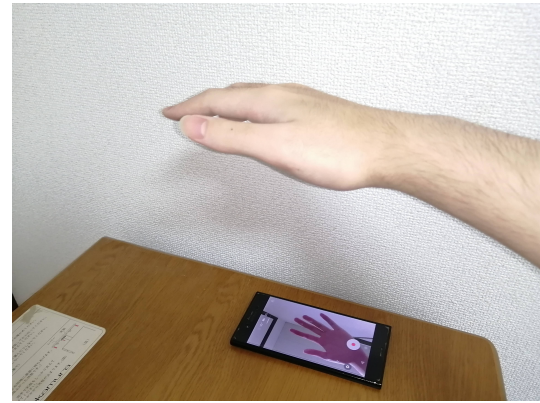


図2 グーパー運動の撮影
 Fig.2 Recording G&R test.

2.2 カメラ映像を用いたスクリーニング

カメラ映像を用いたスクリーニング手法についても、これまでさまざまな病態や疾患に対して研究が進められてきた。Gomez らは、顔の表情を撮影した映像を用い、パーキンソン病の症状である仮面様顔貌（顔面筋の変化による表情の硬直）をスクリーニングする手法について検討を加えた^[6]。Bandini らは、深度カメラを用いて表情の動きを計測し、筋萎縮性側索硬化症の症状を発見する手法を提案した^[7]。本研究はこのようなカメラによるスクリーニング手法を、手指の運動障害を引き起こす頸髄症に応用するものである。

3. 提案手法

本研究における提案手法の流れを図1に示す。提案手法の実装には Python を用いた。

3.1 グーパー運動の撮影

グーパー運動の撮影には、スマートフォンのフロントカメラを用いる。撮影対象者は机の上に置かれたスマートフォンの直上に片手を構え、グーパー運動をできる限り敏捷に繰り返す（図2）。記録者は運動の開始とともに、30 fps のフレームレートで録画を始める。録画時間が 10 sec を超え、かつグーパー運動の反復回数が 20 を超えたところで記録者は録画を停止し、撮影は終了となる。

3.2 画像処理による特徴量の抽出

撮影した映像と画像処理を用い、運動の様子を特徴量として抽出する。ここでは Google が提供する画像処理フレームワークである MediaPipe Hands^[8] を用いる。MediaPipe は手指の特徴点 21 個の位置を時系列で取得する（図3）。この位置は画像の横方向 (x)、縦方向 (y)、奥行き方向 (z) からなる 3 次元座標で表現されるため、この特徴量は $3 \times 21 = 63$ 次元で構成される。このうち x および y については、撮影時のスマートフォンと手の位置関係に依存する。この影響を除くために xy 座標系を、手首を中心とした $x'y'$ 座標系に変換する（図3）。また隣接動画フレーム間の

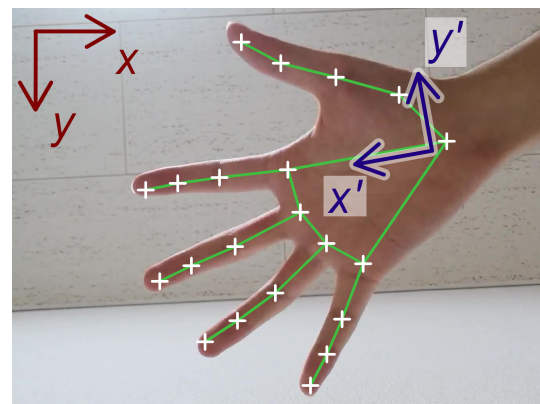


図3 MediaPipe^[8]を用いた画像処理
 Fig.3 Image processing with MediaPipe^[8].

特徴量の差分により、各特徴点位置の速度成分を求める。このようにして得られる位置成分と速度成分の合計 $63 + 63 = 126$ 次元の特徴量を、データとして以降の処理に用いる。

3.3 データの前処理

最初の前処理として、画像処理により得られた特徴量のデータを、20 個のフレームに時間軸で分割し、それぞれ 0 から 19 までのフレーム番号を割り振る（図4）。各フレームの長さは 64 であるのに対し、動画のサンプリングレートは 30 fps であるから、1つのフレームは約 2 sec 間のグーパー運動に対応する。また隣接フレームの始点間の時間差 d は、もとの時系列データと最後のフレームの終点が一致するように調整される。

次に高速フーリエ変換を用い、各フレームにある 126 次元の特徴量をそれぞれ周波数成分に変換する。各次元は 64 個の要素をもつが、対称成分を考慮し低周波 32 成分の振幅成分を抽出する。よって変換後の各フレームは $32 \times 126 = 4,032$ 個の要素で構成される。

複数の試行（グーパー運動の映像）を用いてスクリーニングする場合のみ、最後の前処理としてデータを結合する。各試行の映像から得られたフレームを、

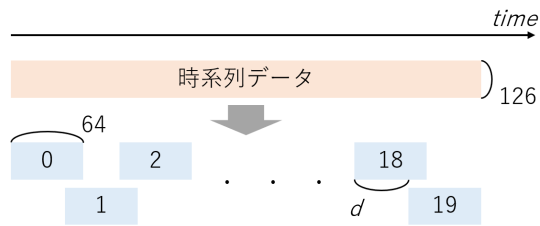


図4 データの分割
 Fig. 4 Segmentation of data.

表1 実験参加者の特徴
 Table 1 Characteristics of participants.

	頸髄症患者		健常者	
	男性	女性	男性	女性
参加者数	10	10	8	7
年齢 (平均)	66.8	71.6	57.9	64.9
年齢 (標準偏差)	7.9	11.1	14.4	12.2

0 から 19 までのフレーム番号ごとに結合する。

3.4 2クラス分類モデル

20 個の前処理後フレームは、それぞれ独立に 2 クラス (患者クラス, 健常者クラス) SVM 分類モデルの生成に用いられる。その際の正解ラベルには、医師による診断の結果を用いる。スクリーニングの際にはこの分類モデルを用いて頸髄症症状の有無を推定する。なお各フレームに対してそれぞれ独立に推定結果を得るため、1 人のスクリーニング対象者から、推定確率に対応する 20 個の決定関数を得ることとなる。最終的な推定結果は、これらの平均により求める。

4. 評価実験・考察

4.1 概要

10 分割交差検証に基づき提案手法の性能を評価するため、35 名 (頸髄症患者 20 名, 健常者 15 名, 全員右利き) の実験参加者からグーパー運動の映像を収集した。参加者の性別, 年齢に関する情報を表 1 に示す。各参加者は右手, 左手 1 回ずつグーパー運動の撮影を受けたため、合計 70 本の動画が収集された。

なお評価には、感度, 特異度, area under the curve (AUC) の 3 指標を利用した。このうち感度および特異度については、2 クラス分類時のしきい値により複数の組み合わせが得られた。本評価では、このうち (感度)² + (特異度)² を最大化する組み合わせを評価指標として選択した。また交差検証の際は、各前処理後フレームを参加者単位で分割した。そのため一つのサブセットにおいて、同じ参加者のフレームが学習データとテストデータの双方に含まれることはなかった。

4.2 特徴量の選定

まず分類性能の向上を目的に、2 クラス分類に用いる特徴量を選定するため、特徴量の種類ごとに分類性

表 2 特徴量の種類と AUC

Table 2 Types of feature values and AUCs.

利用した特徴量	AUC
母指の 4 点	0.907
示指の 4 点	0.900
中指の 4 点	0.877
環指の 4 点	0.853
小指の 4 点	0.870
五指の指先	0.880
五指の DIP 関節	0.867
五指の PIP 関節	0.910
五指の MP 関節	0.930

能を評価した。図 3 に示したとおり、五指それぞれ 4 点 (指先, 遠位指節間 (DIP) 関節, 近位指節間 (PIP) 関節, 中手指節 (MP) 関節) の位置や速度を特徴量として取得している。ここではこのうち、指一本の 4 点のみ、または五指の各 1 点のみの位置や速度を特徴量として利用した場合の性能を評価した。なおここでは、両手各 1 回の試行に基づいた分類モデルを生成、利用した。また母指は他の指と比較して関節が 1 つ少ないため、手首に近い手根中手関節の位置、速度が取得された。本研究では簡単化のため、母指の特徴量も他の指と同様に構成されていると考えた。

各種類に対して評価した際の AUC を表 2 に示す。感度, 特異度はいずれの場合もそれぞれ 90.0%, 86.7% となった。AUC については小指側と比較して母指側, 指先側と比較して手首側の特徴量に対して高くなる傾向が読み取れる。そのため以降の評価では、母指および示指の PIP 関節および MP 関節の特徴量を利用した。この組み合わせは、 $2 \times 3 \times 2 \times 2 = 24$ 次元の特徴量で構成される。

4.3 選定した特徴量を用いた場合の分類性能

4.2 節で選定した 24 次元の特徴量を用い、両手各 1 回ないし右手 1 回のみ, 左手 1 回のみグーパー運動からスクリーニングした場合の性能をそれぞれ評価した。これにより特徴量選定後の分類性能を示し、医療現場などでの応用に向けて十分な性能が得られるかについて検討することを目的とした。

この評価の結果を表 3 に示す。左手のみ 1 回の場合と比較し、右手の試行が考慮された場合に高い性能となる傾向が読み取れる。この要因として、実験参加者の利き手の影響が考えられる。一般的に利き手は非利き手と比べて器用に動かせる場合が多い。この器用さの違いが、頸髄症症状の特徴の見分けやすさに影響を与え、分類性能に影響を与えた可能性があると考えられる。なお本稿では、右手のみ 1 回に対する結果を最終的な性能評価の結果とする。

表3 選定した特徴量に対する分類性能
 Table 3 Classification performance for selected parameters.

試行回数	感度 (%)	特異度 (%)	AUC
両手各1回	90.0	86.7	0.947
右手のみ1回	90.0	93.3	0.947
左手のみ1回	85.0	80.0	0.890

表4 スクリーニング性能の比較
 Table 4 Comparison of screening performance.

手法	感度 (%)	特異度 (%)
10秒テスト ^[10]	70.3-82.1	42.9-74.0
ホフマン反射 ^[11]	44	75
提案手法	90.0	93.3

5. 議論

5.1 既存手法との比較

現在使われている頸髄症のスクリーニング手法には、手指運動に着目した10秒テストや、中指をはじいた際の母指の反射に着目したホフマン反射^[9]などがある。これらの既存手法と提案手法との性能比較を表4に示す。提案手法について、感度、特異度ともに既存手法を上回る性能であることから、実際の医療現場などで活用するに十分な性能を発揮しうると考える。

5.2 制約および今後の展望

提案手法の制約として、現時点では手指の運動障害をもたらしほかの疾患や病態を考慮していないことが挙げられる。このような疾患、病態には、頸髄症のほかにも手根管症候群などさまざまな種類がある。現状の提案手法では、スクリーニングされた症状の原因がどの疾患、病態によるものかを判別することができない。そのため今後は、複数の疾患や病態に適応できるような提案手法の拡張を目指す。これにより症状のスクリーニングだけでなく、鑑別診断での活用も期待できると考える。

本研究の最終的な目標は、実際の医療現場などにおける提案手法の活用である。特にかかりつけ医や介護施設などにおける一次的なスクリーニング手法として提案手法を活用し、頸髄症の早期発見に寄与することを目指している。そこで今後は現場の医療従事者などからの意見を参考にしながら、現場のニーズに合った手法となるようさらなる改良を目指す。

6. おわりに

グーパー運動の映像をベースに、画像処理や機械学習の手法を取り入れた新たな頸髄症スクリーニング手法を提案した。交差検証により提案手法の性能を検証した結果、右手のグーパー運動をもとにスクリーニングした場合の感度、特異度、AUCはそれぞれ90.0%、

93.3%、0.947となった。これは現在使われているスクリーニング手法である10秒テストやホフマン反射の性能を上回っている。今後は複数の疾患や病態に対して提案手法を適応させ、鑑別診断などへの応用を目指す。また実際の医療現場などでの活用を見据え、現場のニーズに適合した手法への改良について検討を進める。

謝辞

本研究は、JSPS 科研費（課題番号 JP21H03485）、および JST AIP-PRISM（課題番号 JPMJCR18Y2）の助成を受けたものである。

参考文献

- [1] Ono, K., Ebara, S., et al.: Myelopathy Hand. New Clinical Signs of Cervical Cord Damage; The Journal of Bone and Joint Surgery. British Volume, **Vol.69**, No.2, pp.215-219 (1987).
- [2] Chang, M., Canseco, J.A., et al.: The Role of Machine Learning in Spine Surgery: The Future Is Now; *Frontiers in Surgery*, **Vol.7** (2020).
- [3] Merali, Z., Wang, J.Z., et al.: A Deep Learning Model for Detection of Cervical Spinal Cord Compression in MRI Scans; *Scientific Reports*, **Vol.11** (2021).
- [4] 加藤, 松本, 他: 手指運動三次元計測による頸髄症スクリーニング手法; ヒューマンインタフェース学会研究報告集, **Vol.22**, No.6, pp.11-12 (2020).
- [5] Su, X.J., Hou, C.L., et al.: Clinical Application of a New Assessment Tool for Myelopathy Hand Using Virtual Reality; *Spine*, **Vol.45**, No.24, pp.E1645-E1652 (2020).
- [6] Gomez, L.F., Morales, A., et al.: Improving Parkinson Detection Using Dynamic Features from Evoked Expressions in Video; 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.1562-1570 (2021).
- [7] Bandini, A., Green, J.R., et al.: Automatic Detection of Amyotrophic Lateral Sclerosis (ALS) from Video-Based Analysis of Facial Movements: Speech and Non-Speech Tasks; 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition, pp.150-157 (2018).
- [8] Google: MediaPipe; <https://mediapipe.dev/>, (accessed 2021-09-06).
- [9] 国分: 頸椎症; *Medicina*, **Vol.32**, No.12, pp.585-586 (1995).
- [10] Machino, M., Ando, K., et al.: Cut Off Value in Each Gender and Decade of 10-s Grip and Release and 10-s Step Test: A Comparative Study Between 454 Patients with Cervical Spondylotic Myelopathy and 818 Healthy Subjects; *Clinical Neurology and Neurosurgery*, **Vol.184** (2019).
- [11] Cook, C., Roman, M., et al.: Reliability and Diagnostic Accuracy of Clinical Special Tests for Myelopathy in Patients Seen for Cervical Dysfunction; *The Journal of Orthopaedic and Sports Physical Therapy*, **Vol.39**, No.3, pp.172-178 (2009).