

深層学習を用いた動画像の輝度変化による咀嚼検出

中田 大樹, 小河 誠巳

東京電機大学理工学部理工学科情報システムデザイン学系

1 はじめに

肥満は糖尿病などや心疾患などの生活習慣病を引き起こす要因とされている。肥満対策を行っていても肥満患者の数は減少してないのが現状である。肥満対策の一つとして、咀嚼回数を増やすことが挙げられる。肥満は早食いの人に多い傾向にあるため [1]、咀嚼の回数を増やすことで消化を助けることができ、さらに満腹感を得られるため肥満対策には効果的である。そのため咀嚼回数を検出する方法を提案することで咀嚼回数向上を促すことを目指す。

2 目的

本研究では、機器装着による利用者の負担を減らし簡易な利用を目指すため、咀嚼の検出を動画を用いて行うことを目的とする。咀嚼の検知、回数計測については動画像から得られる輝度の変化を深層学習によって行う。

3 関連研究

咀嚼回数を計測する研究はこれまでも行われており、咀嚼音から回数を検出する研究 [2] やエッジ検出による上唇と下顎の距離により咀嚼を検出する動画像処理の研究などが報告されている。

深層学習を用いた咀嚼回数の計測の研究として、同じく動画像から作成したデータを深層学習する研究が行われた [3]。この研究では、顔の特異点 (口角) の y 軸方向の移動量をデータとして作成しているが、本研究では動画像の輝度が咀嚼により変化する様子に着目し、データの作成を行う。

4 手法

4.1 動画データ

本研究で使用した動画データは顔が正面を向き、顔から 40~50cm 離れた位置でスマートフォンを机に

置いた状態で撮影した。カメラは Google PIXEL5 の内カメラで撮影した。撮影は 10 秒間で、1~10 回咀嚼した場合の計 10 パターンを行い、合計 1019 個の撮影データを用意した。

以降述べるデータとは、10 秒間のフレーム毎のグレースケールの値をまとめたものであり、10 秒間を 1 データとする。

4.2 データの前処理

この章でのデータとは、動画から得られた 10 秒間のグレースケールの値のことである。

咀嚼検出のためのサンプリング周波数は高フレームレートである必要がないため、FPS を 10 とした。撮影された動画は OpenCV を用いて顔の検出を行い、検出される顔の座標から鼻下から下唇を含む口周りの範囲のみの値を得る。

また、SDTV 規格である BT.601 を用いることで、RGB 画像をグレースケールへの変換を行った。

グレースケール値は 0~255 の値を取るが、事前学習のデータサイズに合わせるため観測された最小値と最大値が-1,1 の範囲になるようデータの正規化を行う (図 1)。データ数 1019 を 8:2 で学習用データと

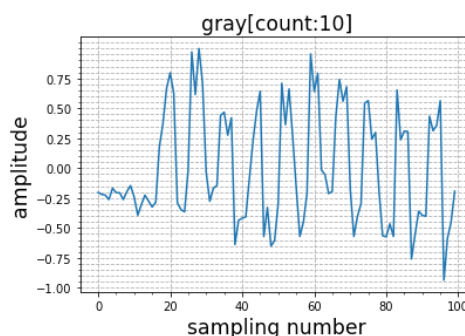


図 1 10 回咀嚼時の波形

テスト用データに分割し学習を進めた。

4.3 事前学習

動画のフレーム毎の輝度の変化量の繰り返し回数を求めることで咀嚼回数の検出を行う。繰り返し回

数はニューラルネットワークで学習して求める。事前学習として波形 20000 個から波の繰り返し回数を求める学習を行う。

ネットワーク構造は 8 層の全結合層で、中間層の活性化関数に ReLU を使用し、出力層の活性化関数にソフトマックス関数を用いた。損失は交差エントロピー誤差で求め、最適化に Adam を用いた。

比較として、3 種類の波形を学習させた。平均 0、振幅が-1~1 の正弦波、0 以下を 0 とした正弦波上半波、正弦波上半波の間隔をランダムにした波形の 3 種類にノイズを加えた波形を用意した。各波形の周波数は 1~10Hz とし、求める繰り返し回数も 1~10 回とする。

4.4 転移学習

4.3 節の事前に学習させたモデルを利用して 4.2 節で作成した動画の波形データを使用して学習を行った。事前に学習する波形の違いで転移学習の結果が影響されるか確認した。

5 結果

事前学習の結果を表 1 に、学習済みモデルで実際の波形を学習させた結果を表 2 に示す。事前学習の正答率では使用データが正弦波のモデルが一番高く、正弦波上半波ランダム間隔のモデルが一番低い (表 1)。逆に転移学習の正答率は、事前に学習した波形が正弦波のモデルが一番低く、正弦波上半波ランダム間隔のモデルが一番高い結果となった (表 2)。事前学習には実際のデータに近いデータで学習することが、実際のデータに適合することが確認できる。

転移学習で一番結果の良いモデルでも咀嚼数を完全に一致させる精度は高くない。しかし、誤差ごとのデータ数 (表 3) や混同行列 (図 2) を見ると、誤差 ±2 以内に 95.6% のデータが存在するため大きく外れた予測は少ないことがわかる。

表 1 事前学習の結果

	正弦波	正弦波上半波	正弦波上半波間隔ランダム
正答率	0.9892	0.9160	0.3912
損失	0.0563	0.3145	0.6165

表 2 転移学習の結果

	正弦波	正弦波上半波	正弦波上半波間隔ランダム
正答率	0.3725	0.4412	0.5343
損失	1.7636	1.6379	0.2590

予測ラベル

	1	2	3	4	5	6	7	8	9	10
1	16	1	0	0	0	0	0	0	0	0
2	8	13	7	0	0	0	0	0	0	0
3	1	4	11	7	0	0	0	0	0	0
4	0	0	8	4	1	0	0	0	0	0
5	0	0	4	5	4	4	1	1	1	0
6	0	0	2	5	3	12	3	1	0	1
7	2	0	1	1	1	7	4	3	0	1
8	0	0	0	0	0	3	4	9	3	2
9	0	0	0	0	0	0	1	3	8	5
10	0	0	0	0	0	0	2	1	5	10

図 2 予測値と正解値の混同行列

表 3 誤差のデータ数 (全体に対する割合 [%])

テストデータ数	正解数	誤差 ±1	誤差 ±2
204	91(44.6)	172(84.3)	195(95.6)

6 まとめ

動画データからフレーム毎のグレースケール値を求め、CSV データに納めた。

そのデータから深層学習により咀嚼回数を学習できている状態を確認できた。

現在、咀嚼回数の計測を行っているのは純粋な咀嚼時のみであるが、食事中には会話や口に食べ物を運ぶシーン等も含まれる。そのため、食事中における咀嚼区間の検出を行うことを検討する。

謝辞

本研究の一部は、東京電機大学総合研究所研究 Q22J-06(研究課題番号) として行ったものである。

参考文献

- [1] 厚生労働省:平成 21 年度国民健康・栄養調査
- [2] 三井秀人, 中野和哉, 磯山直也, ロベズ・ギヨーム:咀嚼回数向上支援システム, 「マルチメディア, 分散, 協調とモバイル (DICOM2017) シンポジウム」, Jun., 2017
- [3] 初澤颯人, 小河誠巳:深層学習による繰り返し回数計測の手法, 情報処理学会第 84 回全国大会, 愛媛大学, Mar., 2021