

Classification Method of Eating Behavior by Dietary Sound Collected in Natural Meal Environment

蒲地 遥 (15816024)

ロペズ研究室

1. はじめに

肥満は糖尿病や高血圧などの生活習慣病を引き起こす恐れがある。厚生労働省は肥満予防のために対策を講じてきたが、肥満の患者数は 10 年前と比べて減少していない[1]。また、早食いの人ほど BMI が高い傾向にあることが示されている[2]。さらに、食事中の会話がある場合、生活習慣が規則正しく、好き嫌いが無いなど健康と関連があることから[3]、食事中の会話を増やすことが望ましい。また、自由な食事環境での詳細な食事行動の識別が可能になれば、食事行動と健康に関連する研究に貢献できると考える。

本研究では、自然な食事環境下での食事行動の定量化を実現することを目的とし、自然な食事環境に対応した食事行動（咀嚼・食べ物の嚥下・飲み物の嚥下・発話・ノイズなどのその他の音）の高精度分類を目標とする。自然な食事環境とは、日常生活における食事環境とする。なお、本研究では、リアルタイムでの分類は行わないため、各行動に該当する音声区間を手動で抽出し、分類に用いる。

2. 関連研究

Amft らは、耳の内側にマイクロフォンを配置することで質の高い咀嚼音を取得可能であると示した[4]。

また、近藤らは、自然な食事環境下で骨伝導マイクロフォンを用いて音声データを収集し食事行動を分類する手法を提案した[5]。4 種類の食事行動の分類を高精度で実現しているが、不均衡なデータセットに対するリサンプリングを行うタイミングにより学習モデルの過学習が起きている可能性がある。

表 1. 各ラベルの合計ラベルデータ数

ラベル名	咀嚼	食べ物の嚥下	飲み物の嚥下	発話	その他
合計	2001	119	83	555	201

表 2. 訓練データセットのデータ数の一例

ラベル名	咀嚼	食べ物の嚥下	飲み物の嚥下	発話	その他
データ数	447	222	295	447	447

3. 自然な食事環境下での食事行動の分類手法

本研究では、自然な食事環境下での食事行動の分類を行うために、自然な食事環境下での食事音声データを収集した。収集した音声データの「咀嚼・食べ物の嚥下・飲み物の嚥下・発話・その他の音」に該当する区間を手動でラベリングした。表 1 にラベリングされたデータの合計結果を示す。

次に、特徴量抽出を行った。Short Term Energy (STE), Power Spectrum Density (PSD), Mel Frequency Cepstrum Coefficient (MFCC) などを含めた合計 75 個の特徴量を抽出した。表 1 が示すように咀嚼以外のデータ数は咀嚼と比べて非常に少ないため、リサンプリングを行いより均衡なデータセットを作成した。リサンプリングは、咀嚼のデータをランダムに 500 個選択した後、機械学習における分類モデル作成の際に訓練データのみを対象に Support Vector Machine (SVM) による Synthetic Minority Oversampling TEchnique (SMOTE) を利用した。表 2 にリサンプリング後の訓練データ数の一例を示す。

次に、食事行動の分類を行うために、最適な分類モデルを選定した。Matlab の「分類学習器」による各モデルの交差検証結果を表 3 に示す。これにより、最も精度が高かった中程度のガウス SVM (rbf カーネル) を用いた。

2019 年度（令和元年度）卒業論文要旨

表 3. 分類学習器によるモデルの交差検証結果

分類モデル		精度[%]
決定木	複雑な木	68.6
	中程度の決定木	62.8
	粗い木	54.6
SVM	線形 SVM	73.1
	細かいガウス SVM	62.8
	中程度のガウス SVM	85.9
	粗いガウス SVM	68.2
KNN	細かい KNN	80.2
	中程度の KNN	67.3
	粗い KNN	54.0
	コサイン KNN	74.2
	3 次 KNN	65.1
	重み付き KNN	75.0
アンサンブル分類器	ブースティング決定木	68.5
	バギング決定木	82.4

表 4. 汎化性能評価結果(平均)

特徴数	F1 値			
	75	48	75	48
咀嚼	0.75	0.78	0.77	0.75
嚥下	食べ物	0.43	0.48	0.28
				0.27
発話	飲み物	0.43	0.48	0.19
	0.29			
発話		0.90	0.89	0.91
その他		0.51	0.50	0.51

4. 分類性能評価

交差検証の結果から選択された分類モデルでの高精度な分類を実現するため、パラメータの調整を行った。SVM の正則化パラメータ C とガウシアンカーネルの幅の逆数を示す γ を調整し、 $C=10$ 、 $\gamma=0.01$ のときが最適であるという結果が得られた。よって、このパラメータを用いて分類モデルの作成を行った。

本研究では、ラベリングした 5 種類の行動の分類と、食べ物の嚥下と飲み物の嚥下を 1 つの嚥下ラベルとした 4 種類の行動の分類を行った。また、抽出した 75 個の特徴量での分類の他に 48 個の特徴量での分類も行った。特徴量が減少した理由は MFCC に関して 39 次ではなく 12 次を用いたからである。今回は訓練データとテストデータの組み合わせを 10 個用意し、分割後に訓練データのみオーバーサンプリングを行ったため、それぞれの条件の 10 回分の平均の結果を表 4 に示す。

5. まとめ

本研究は、自然な食事環境下での食事行動の分類手法を提案した。

発話に関する F1 値はどの条件の場合でも約 90% と良い結果となった。咀嚼も 75% 以上の F1 値となった。

分類する行動を増やすことで嚥下に関する F1 値は下がった。また、特徴量を減らした場合に関して、分類精度の大幅な向上は見られなかった。このことから、MFCC に関して 12 次を使用した場合でも分類精度を維持できることが分かった。

今回の結果から、食べ物の嚥下・飲み物の嚥下に関する分類精度が低いことが分かった。この 2 つの行動は咀嚼と間違っ予測されることが多かった。よって、咀嚼・食べ物の嚥下・飲み物の嚥下の違いが明らかとなるように特徴量を追加・変更する必要がある。また、リサンプリングを訓練データのみに行うことで精度が落ちたことから、先行研究[5]の学習モデルは過学習が起きていたことが考えられる。

今後の展望としては、リアルタイムでの食事行動の分類を目指す。これを実現するために、まず特徴量の変更やラベリングの見直しなど、より高精度に分類できるような手法を検討し、さらにリアルタイムで咀嚼や嚥下などの食事行動に該当する音声区間を自動的に抽出するシステムを設計していきたい。

参考文献

- [1] 厚生労働省: 国民健康・栄養調査結果の概要 (平成 29 年)
- [2] 安藤雄一, ほか: 「ゆっくりと噛んで食べることは肥満防止につながるか?」. ヘルスサイエンス・ヘルスケア, Vol. 8, No. 2, pp.51-63, 2008.
- [3] 岸田典子, ほか: 学童の食事中における会話の有無と健康及び食生活との関連. 栄養学雑誌, Vol. 51, No. 1, pp. 23-30, 1993.
- [4] Oliver Amft, et al: Analysis of chewing sounds for dietary monitoring. In International Conference on Ubiquitous Computing, pp.56-72, 2005.
- [5] Takumi Kondo, et al: Robust classification of eating sound collected in natural meal environment. In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, pp.105-108, 2019.