

画像を用いたラーメン美不味の判断

2020se001 坂友翔 2020se027 窪山和樹

指導教員：野呂昌満

1 はじめに

近年、スマートフォンやタブレット端末などの普及により、情報や画像が多く公開され、どこにいても画像を取得できるようになった。私たちの生活においても、好みの服を探したり、目的地とする旅行先を探したりするのにインターネット上に公開されている画像を用いることが増えている。飲食店においても、テレビやスマートフォンから情報を得て、訪ねる店舗を決めることが多い。

現在、料理の美味い、不味いを判断する方法として、実際に料理を食べに行く、味覚センサ [1] などの味を数値化する機械で判断するなどの方法があるが、どちらも手間がかかる、コストが高いなど、問題がある。いつでも簡単に取得することができる画像から料理の美味しい不味いを判断することができれば、飲食店を選択する際の助けになると考えた。以下、美味しい不味いを美不味と表現する。

本研究の目的は、画像から美不味を判断できる可能性を考察することである。我々は豚骨ラーメンをよく食べるので、本研究では豚骨ラーメンを対象を絞って実験を行う。

研究課題を以下に示す。

RQ1. 見た目が美味なラーメンの考察

RQ2. 特徴量とするデータの定義

RQ3. ニューラルネットワークの設計と実装

RQ4. 特徴量の妥当性の検証

我々が今までに豚骨ラーメンを食べてきた経験から、見た目が美味なラーメンに当てはまる特徴は何か予測し、どのような特徴を持つラーメンが美味であるのか仮説を立てる。考察した美味なラーメンの持つ特徴を詳細化し、豚骨ラーメンの美不味に影響を与える特徴量を定義する。定義した特徴量を抽出し、抽出した特徴量からラーメンの美不味の判断を行うニューラルネットワークの設計・実装を行う。定義した特徴量の抽出には CNN、ラーメン美不味の判断には FCNN(Fully Connected Neural Network) を用いる。CNN では豚骨ラーメンの画像を入力データとし、定義した特徴ごとに特徴量を画像から抽出する。FCNN では、CNN から抽出した複数の特徴量を入力データとし、総合的にラーメンの美不味の判断を行う。設計したニューラルネットワークによって美不味の判断されたラーメンとそのラーメンを提供している店舗の食べログの評価を比較し、特徴量の妥当性の検証を行う。

2 関連研究

土井 [2] は、深層学習の画像の識別分野の事例として、ラーメン画像から店舗を識別するモデルを作成する際の、具体的な作業項目や要所を紹介することを目的としてい

る。ラーメン次郎 40 店舗を対象に約 8 万枚の画像を収集し、モデルの学習、学習したモデルの評価を行っている。学習済み CNN モデルの Inception-v3, ResNet, SE-ResNeXt の 3 つを用いて比較を行った。評価は正解率、適合率、再現率、F-値の 4 つの尺度としている。結果は全てのモデルでそれぞれ正答率、適合率、再現率、F-値の 4 つの尺度で高い精度を得た。最も適合率が低い店舗でも 91.25 % と良い結果が得られており、ラーメン画像から店舗を識別することができていると言える。

佐藤ら [3] の研究は一般の人が料理を魅力的に撮影できるよう支援を目的とし、CNN を用いて料理画像の魅力度を判定する手法を提案、評価することを行っている。学習済み CNN モデルの VGG16, ResNet, Inception-v3 の 3 つを使用し、比較をした。畳み込み層はそれぞれのモデルの出力層側にある全結合層を除いた層からなり、VGG16 は 16 層、ResNet は 50 層、Inception-v3 は 48 層とし、それ以降に 256 次元の全結合層を 3 層追加、魅力度を [0,1] で出力させている。魅力度は一対比較法により、料理画像の魅力度を数値化したものである。実験結果は、魅力度の推定誤差の平均が、VGG16 は 0.092, ResNet は 0.110, Inception-v3 は 0.117 となった。3 つモデルの中で最も単純な構造である VGG16 の推定誤差が一番 0 に近いことから、料理画像の魅力度推定には単純な CNN モデルが適していると考えられる。

3 前提条件と特徴量の考察・定義

3.1 美味なラーメンの持つ特徴量の考察・定義

本研究の目的はラーメン画像から美不味を判断できる可能性を考察することである。RQ1 で美味なラーメンの特徴を「どんぶりの色が黒色」、「チャーシューの色が濃い」、「スープの色が薄い黄金色」、「トッピングの数が多」、「麺線が整っている」と考察した。考察したラーメン美不味の概念図を図 1 に示す。

チャーシューの色が濃い、スープの色が薄い黄金色、トッピングの種類が多いの 3 つの特徴を持つラーメンは市場に多く出回っているため競争率が高くなる。競争率が高くなると、他のラーメンに勝つために味で勝負する必要があり、美味なラーメンが増加する。以上のことから「チャーシューの色」、「スープの色」、「トッピングの種類」を特徴量とした。

どんぶりの色と麺線の綺麗さは直接味とは関係がない。しかし、黒色のどんぶりを使うと、ラーメンのスープやトッピングの色が映え、見た目が美味に見える。店舗側は初期集客率を高めるために、ラーメンが美味に見える黒色のどんぶりをよく使う。我々は、リピート率が高い店舗は

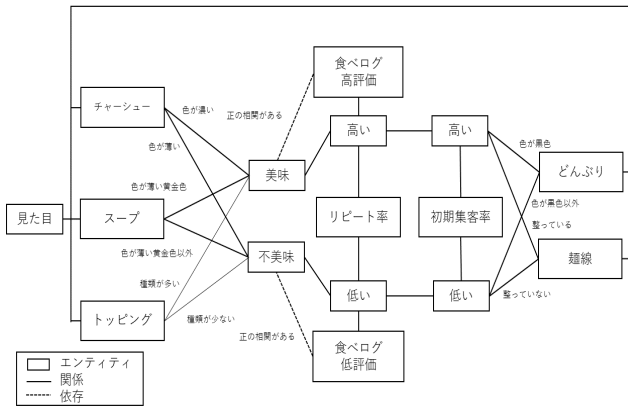


図1 考察したラーメン美不味の概念図

黒のどんぶりを使っているという関係があり、黒色のどんぶりを使っている店舗は美味なラーメンを提供している可能性が高いと考えた。この仮説を検証するために、「どんぶりの色」を特徴量とした。また、麺線が整っている場合も見た目が美味に見える。どんぶりの色と同様に店舗側は初期集客率を高めるために、麺線を整えて提供する。リピート率が高い店舗は麺線が整っているという関係があり、麺線が整っている店舗は美味なラーメンを提供している可能性が高いと考えた。この仮説を検証するために、「麺線」を特徴量とした。しかし「麺線」は、麺線を判断できる画像が少なく、データセットの作成が難しいことから特徴量として選択しなかった。「トッピングの種類」は、ほとんどの店舗がすべての種類のトッピング入れたラーメンを提供していることから店舗ごとの差が生まれにくいことから特徴量として選択しなかった。この考察を踏まえて、本研究では以下の3つを特徴量として定義した。

- チャーシューの色
- スープの色
- どんぶりの色

定義したラーメン美不味の概念図を図2に示す。

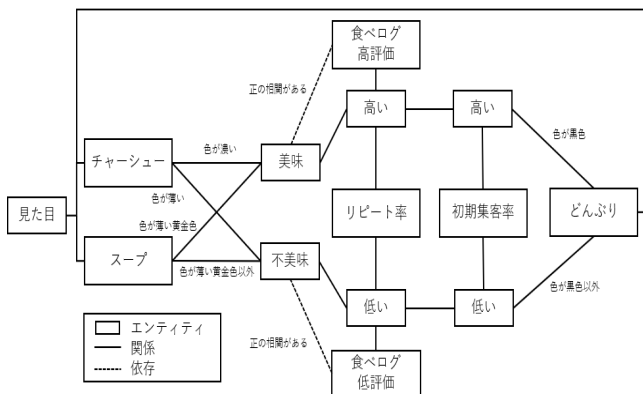


図2 定義したラーメン美不味の概念図

3.2 前提条件

以下の1点を前提条件として定義した。前提条件とした理由を以下に示す。

- 食ベログで高評価な店舗が提供するラーメンは美味である

店舗のリピート率が高ければ、食ベログ評価は高評価になる。また店舗のリピート率が高いければ、ラーメンは美味でもある。以上のことから、食ベログ評価とラーメンの美味とは正の相関関係があると考えた。以上の理由から、食ベログで高評価な店が提供するラーメンは美味であるということ的前提条件の1つとする。

4 データセット

本研究では、美味と定義した「どんぶりの色が黒色」、「チャーシューの色が濃い」、「スープの色が薄い黄金色」の特徴を持つ豚骨ラーメン画像を収集した。また、「どんぶりの色が黒色以外」、「チャーシューの色が薄い」、「スープの色が薄い黄金色以外」の豚骨ラーメン画像も収集した。画像は手作業で収集し、画像枚数が少ないことから画像をランダムに回転、反転させ、データセットの拡張を行い、各CNNごとに訓練データ240枚、テストデータ60枚を用意した。さらに、特徴を際立たせ、正確に訓練させるために、背景除去を行った。

5 ニューラルネットワークの設計、実装

5.1 実装環境

実装した際のハードウェアコンポーネントとソフトウェアコンポーネントをそれぞれ表1、表2に示す。

表1 ハードウェアコンポーネント

システム	名前
os	Windows10
プロセッサ	Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz
メモリ	16.0GB

表2 ソフトウェアコンポーネント

コンポーネント名	バージョン
Python	3.8.9.64
tensorflow	2.13.0
NumPy	1.22.0

5.2 ニューラルネットワークの設計

CNNはニューラルネットワークと異なり、重みを共有している。学習が効率的に行われ、精度の高い特徴を抽出することが出来るので用いる。本研究では、独立した3つのCNNを用いてそれぞれ特徴量を抽出する。独立した3つのCNNを用いる理由は以下の2点ある。

- モデルの分解性の向上

- 異なる特徴抽出の容易性

複数の独立したモデルに分割することで、それぞれのモデルの役割が明確になり、モデルの理解を容易にし、トラブルシューティングがしやすくなる。各 CNN は異なる特徴を抽出する傾向があり、それぞれの CNN が異なる側面やパターンを学習することで、互いに影響を受けずに特徴量を得ることができる。

FCNN は他のニューラルネットワークのから抽出された特徴を組み合わせ、異なるモデルやデータソースからの情報を統合するために使用される。本研究では複数の CNN モデルからの情報を結合して、総合的な判断を行いたいので FCNN を用いる。各 CNN で抽出した特徴ベクトルを FCNN の入力データとして、ラーメンの美不味の判断を総合的に判断する。これらを基に作成した全体図を図 3 に示す。

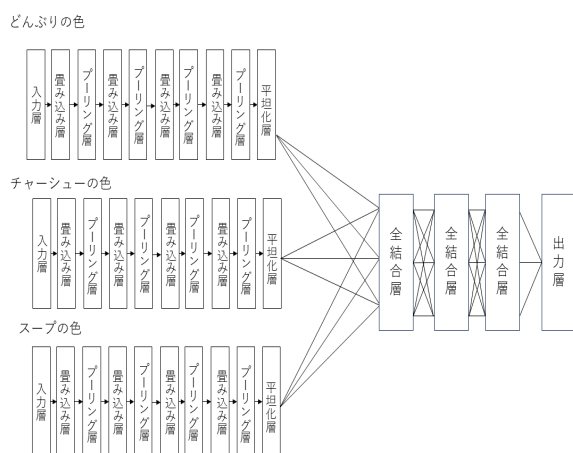


図 3 構造図

5.3 パラメータ

小さなフィルターサイズは画像内の局所的な特徴をより良く捉えることができるので、設計段階では畳み込み層のフィルターサイズ 3×3 、プーリング層のフィルターサイズ 2×2 とし、フィルター数を 32 とした。最終的に使用した CNN のハイパーパラメータは以下の表 3 の通りである。

CNN で実験を行った結果、すべての CNN モデルにおいて、畳み込み層 4 層、プーリング層ともに 4 層の時が最も特徴識別の精度がよかったので、4 層とした。また、フィルターサイズ、フィルター数は最も精度のよかったので、表 3 のパラメータを使用した。入力は、 224×224 の RGB 画像とした。

表 3 CNN のパラメータ

	入力層	畳み込み層	プーリング層	フィルターサイズ	フィルター数	活性化関数
どんぶり	$224 \times 224 \times 3$	4 層	4 層	1, 2, 3, 4 層目 3×3	1, 2, 3, 4 層目 32	ReLU 関数
チャーシュー	$224 \times 224 \times 3$	4 層	4 層	1, 2, 3, 4 層目 3×3	1, 2 層目 32, 3, 4 層目 32	ReLU 関数
スープ	$224 \times 224 \times 3$	4 層	4 層	1, 2, 3, 4 層目 3×3	1, 2, 3 層目 32, 4 層目 32	ReLU 関数

FCNN では全結合層を 3 層用いる。実験を行った結果、

FCNN において、全結合層が 3 層の時が最も特徴識別の精度がよかったので、3 層とした。FCNN でも活性化関数には sigmoid 関数と relu 関数を用いる。sigmoid 関数を用いてラーメンが「美味」、「不美味」の二値分類を行うので、最後の全結合層の出力を 0 から 1 の範囲とした。

6 検証・考察

美味と判断されたラーメンを提供する店舗の食べログ評価を調べる。食べログとはグルメレビューサイトであり、信頼できるお店探しサイトとして定評がある。本研究では、食べログの評価点数が 3.3 点以上の店舗を美味とする。食べログでは、評価点 3.5 点以上は全体の 3% しかないとされている。また 2.5 点未満の評価点も極端に数が少ない。このことから、おおよその評価点が 2.5 点以上 3.5 点未満であると考える。美味の判断基準として 2.5 点以上 3.5 点未満の上位 20% である 3.3 点以上の評価を美味とする。

検証を行ったラーメン画像 10 枚の美不味判断の結果とそれらのラーメンの食べログ評価による判断結果を以下の表 4 に示す。

表 4 検証結果

	FCNN による判断	食べログの評価による判断
店舗 A	○	○
店舗 B	×	×
店舗 C	○	○
店舗 D	×	○
店舗 E	×	×
店舗 F	○	○
店舗 G	×	○
店舗 H	○	×
店舗 I	○	○
店舗 J	×	×

この表は FCNN による美不味の判断、食べログの評価による美不味の判断において美味と判断された場合は「○」、不美味と判断された場合は「×」と記載している。豚骨ラーメンの美不味を画像を用いて判断する精度としては約 60% となった。10 店舗中、店舗 A, B, C, E, F, H, I は黒色のどんぶりが用いられていたが、店舗 B, E では FCNN による判断、食べログの評価による判断ともに不美味という結果、店舗 H では食べログの評価による判断が不美味という結果となった。黒色以外のどんぶりを用いていた店舗 D, G, J, の中で店舗 D, G は食べログによる評価は美味となった。結果、どんぶりの色はラーメンの美不味に影響が少ないと言える。

そこで我々は、どんぶりの色という特徴を除き、再度特徴量とするデータの定義を行った。作成したラーメン美不味の概念図を図 4 に示す。

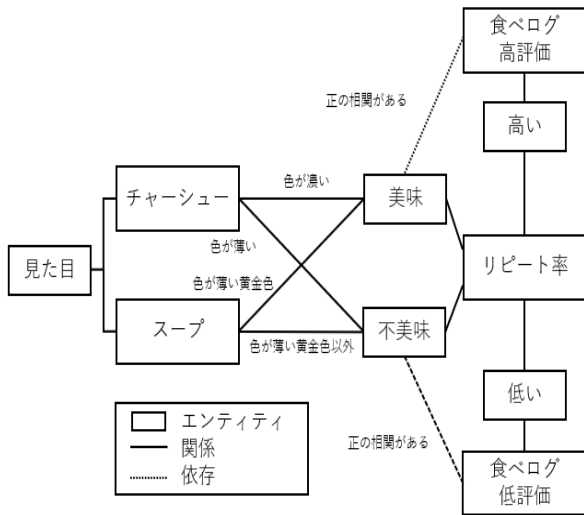


図4 定義したラーメン美不味の概念図

この概念図に基づき、再度ニューラルネットワークの設計を行った。特徴量は「チャーシューの色」、「スープの色」の2つとし、ハイパーパラメータは特徴量3つで行った時と同じパラメータを使用した。この設計に基づき、再度実験を行った結果を以下の表5に示す。黒色のどんぶりが用

表5 検証結果 (特徴量2つの場合)

	FCNN による判断	食べログの評価による判断
店舗 A	○	○
店舗 B	×	×
店舗 C	○	○
店舗 D	○	○
店舗 E	×	×
店舗 F	○	○
店舗 G	○	○
店舗 H	○	×
店舗 I	○	○
店舗 J	○	×

いられていた店舗 A, B, C, E, F, H, I では、FCNN による美不味の判断に変化はなかったが、黒色以外の色のどんぶりが用いられていた店舗 D, G, J の中で、店舗 D, G は FCNN による美不味の判断が美味に変化し、食べログによる美不味の評価でも美味と判断されている。豚骨ラーメンの美不味を画像を用いて判断する精度としては約 80% となった。以上のことから、画像からラーメンの美不味を判断する特徴として「チャーシューの色が濃い」「スープの色が薄い黄金色」は適切であったと言える。

7 おわりに

近年、スマートフォンやタブレット端末などの普及により、情報や画像が多く公開され、どこにいても画像を取得できるようになった。飲食店においても、全国どこからでも店舗が提供している料理の画像を見ることができる。料理の美不味を判断する方法として、実際に料理を食べに行

く、味覚センサ [1] などの味を数値化する機械で判断するなどの方法があるが、食べに行くことは手間がかかる、味覚センサは高価であるのであまり普及していないなどの問題がある。画像から料理の美味しい不味いを判断できれば飲食店を選択する際の支援ができると考えた。本研究では、画像から美不味を判断できる可能性を考察することを目的とした。研究課題として、見た目が美味なラーメンの考察、特徴量とするデータの定義、ニューラルネットワークの設計と実装、特徴量の妥当性の検証を挙げた。

本研究は、各特徴量（「どんぶりの色」、「スープの色」、「チャーシューの色」）を CNN に学習させ、特徴ベクトルを抽出した。それらの特徴ベクトルを、FCNN に入力することによって総合的に美不味の判断を行った。

FCNN によるラーメン美不味の判断の精度は、約 60% となった。実験結果から、どんぶり色は美不味の判断に影響が少ないと言える。再度特徴量の定義を行い、「どんぶりの色」を除いた 2 つの特徴を使って実験を行った結果、約 80% の精度となった。以上のことから、画像からラーメンの美不味を判断する特徴量として「チャーシューの色」「スープの色」は適切であったと言える。

8 今後の課題

本研究では、豚骨ラーメンの「どんぶりの色」、「チャーシューの色」、「スープの色」の 3 つの特徴量とし、美不味の判断を試みたが、これらの特徴以外にも美不味と判断できる特徴があると考えられる。これらの特徴以外で美不味を判断できる可能性がある特徴として、「油の量」があると考えられる。特徴量を増やすことでより正確にラーメンの美不味を判断することができると考える。

参考文献

- [1] Y. Tahara, K. Toko, “Electronic Tongues—A Review,” IEEE Sensors Journal, vol.13, no.8, pp.3001–3011, Aug.2013.
- [2] 土井 賢治, “ラーメン画像からの全店舗識別,” 情報処理論文誌, Vol.59 No.11, pp.971-973, 2018.10.
- [3] 佐藤 陽昇, 道満 恵介, 井手 一郎, 出口 大輔, 村瀬 洋, “畳み込みニューラルネットワークを用いた料理写真の魅力度推定,” 電子情報通信学会技術研究報告, MVE2017-32, Oct (2017).
- [4] 西田 奈生, 金本 玲花, 松本 尚, “深層学習によるインスタグラム画像からの流行抽出,” 研究報告数理モデル化と問題解決 (MPS), pp.1-6, 2020.
- [5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. of the IEEE, pages 2278–2324, 1998.