

令和 5 年 6 月 21 日現在

機関番号：33908

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K12038

研究課題名（和文）深層学習による視覚的感性数値化のための感性データ拡張手法

研究課題名（英文）Data augmentation method for quantification of visual aesthetics using deep learning

研究代表者

道満 恵介（Doman, Keisuke）

中京大学・工学部・准教授

研究者番号：90645748

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：人が料理写真を見た際の視覚的感性（魅力度）を高精度に数値化する機械学習手法を目指し、画像とその魅力度評価値をペアで生成するためのデータ拡張手法を開発した。また、視線情報を利用した画像特徴抽出手法と併用することで、魅力度推定精度が向上することを明らかにした。加えて、深層学習モデルの枠組みの一つであるマルチタスク学習への適用により、推定精度がさらに向上することを明らかにした。さらに、料理写真以外の対象への応用可能性を探るべく、料理写真を含む料理レシピの魅力度を分析・推定する手法を検討し、レシピタイトルと料理レシピに対する注目度の関係を分析した。

研究成果の学術的意義や社会的意義

本研究では、視覚的感性、特に料理写真を見た際に感じる魅力度という人の曖昧な感覚量を深層学習により定量化する手法を開発した。この技術は、これまでデータセットの質・量が足かせとなっていた分野でも深層学習ベースのアプローチを検討する際の一助となり、関連の各学術分野での技術発展を促進することが期待される。また、人の感性を踏まえた振る舞いが可能な人工知能系の基盤構築の一助となることも期待され、社会と人工知能の新たな関係を形成する一助となることも期待される。

研究成果の概要（英文）：This research has developed a data augmentation method for generating pairs of an image and its attractiveness value, aiming at a machine learning method for accurately quantifying the visual aesthetics (attractiveness) for food photos. This research has also shown that the accuracy of attractiveness estimation can be improved by using the developed method in combination with an image feature extraction method based on eye gaze information. In addition, this research showed that its application to multi-task learning, which is one of the frameworks of deep learning models, further improves the estimation accuracy. In addition, to explore the possibility of applying the method to data other than food photos, this research studied a method for analyzing and estimating the attractiveness of food recipes including food photos, and analyzed the relationship between the recipe title and the degree of attention to the food recipe.

研究分野：画像パターン認識

キーワード：視覚的感性 魅力度 深層学習 データ拡張 料理写真

1. 研究開始当初の背景

(1) 人工知能の要素技術として、コンピュータビジョン分野では画像からの物体検出・認識に関する研究が盛んに行われている。深層学習により各種タスクにおける精度が飛躍的に向上したが、深層学習で高い汎化性能を得るには質と量を併せもった学習用データセットが必要となる。これに対して、学習用画像に回転や拡大縮小等の画像変換を適用する等により大量の画像を生成して学習データ不足を補うデータ拡張が一般的に行われている。

(2) 一方で、より高度な人工知能の実現のためには、人の感性や心理を理解する技術も重要となる。例えば情景写真が入力された際に、そこに写る物体やシーンを認識するだけでなく、それを人が見た際にどの程度魅力的に感じるか（以降、「魅力度」と呼ぶ）という視覚的感性を数値化（図1）できれば、より深い意味での画像理解が可能となる。画像特徴と魅力度の関係のモデル化にも深層学習は有効なアプローチであるが、それには画像と魅力度のペアが大量に必要なことになる。よって、検出・認識タスクで適用される画像のみの単純なデータ拡張では必要な学習用データを用意できない。

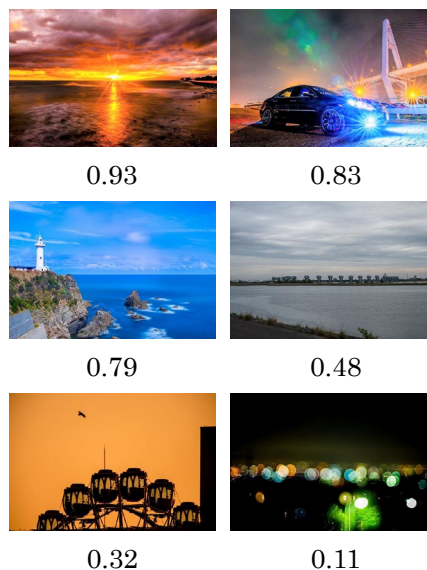


図1 情景写真の魅力度の数値化例
（高：1～低：0）

2. 研究の目的

人が画像を見た際にどの程度魅力的に感じるかという視覚的感性（特に料理写真の魅力度）を数値化する機械学習手法に注目し、そこで必要となる学習用データセット構築のためのデータ拡張手法（達成事項1）およびそれを用いて魅力度を高精度に数値化するための深層学習モデル（達成事項2）を確立する。また、料理写真以外の対象または魅力以外の視覚的感性への応用可能性を模索（達成事項3）する。

3. 研究の方法

(1) 達成事項1について、これまで研究代表者らが研究してきた魅力度付き料理画像データの拡張手法を、同じくこれまで研究代表者らが研究してきた視線情報を利用した画像特徴抽出手法に適用し、定量評価した。これまで研究してきた料理画像データの拡張手法では、見えの変化が知覚できない程の微小な画像変換を元画像に適用することで、元画像と同じ魅力度を生成画像に付与できるという考え方に基づくものであった。また、これまで研究してきた視線情報を利用した画像特徴抽出手法は、人が料理画像を見た際の視線停留分布を画像特徴抽出における重み付けに利用することで、魅力度推定に有効な画像特徴を得るというものであった。前者のデータ拡張手法を、後者の手法に適用することで、大規模な視線情報付き画像データセットを低コストで構築する手法を検討した。

(2) 達成事項2について、深層学習モデルの枠組みの一つであるマルチタスク学習の導入を検討した。加えて、これまで研究代表者らが利用してきた小規模ながら評価値の信頼性が高いデータセットのみならず、評価値の信頼性が多少劣るものの大規模なデータセットを構築した。それぞれのデータセットを利用して深層学習モデルを学習・評価した。さらに、前者のデータセットでモデルの事前学習をしてから後者のデータセットで追加学習するという転移学習も検討した。

(3) 達成事項3について、料理写真以外の対象への応用可能性を探るべく、料理写真を含む料理レシピの魅力度を分析し、推定する手法を検討した。料理の種類や特徴を表現するレシピタイトルが料理レシピの魅力度を大きく左右することに注目し、レシピタイトルと料理レシピに対する注目度の関係を分析した。

4. 研究成果

(1) 3. の(1)の方法を実施し、下記のような知見が得られた。

① 視線停留分布に対するデータ拡張の有無による魅力度推定精度を調査したところ、表 1 のような結果が得られ、当該データ拡張により魅力度推定精度が向上したことが確認された。これにより、本アプローチの有効性が確認された。

② 関連するものの独立してこれまで研究を進めてきた 2 つのアプローチを融合し、より統括的な評価を実施し、その有効性を定量的に示したという点で意義深い成果であると考えられる。

表 1 視線停留分布に対するデータ拡張の有無による魅力度推定精度 (MAE: 平均絶対誤差)

料理	データ拡張	
	なし	あり
鰹のたたき	0.138	0.154
カレーライス	0.113	0.094
鰻丼	0.101	0.066
ビーフシチュー	0.106	0.114
ハンバーグ	0.129	0.12
天丼	0.101	0.092
カツ丼	0.145	0.129
鉄火丼	0.095	0.052
チーズバーガー	0.100	0.100
フィッシュバーガー	0.112	0.087
平均	0.114	0.101

(2) 3. の(2)の方法を実施し、下記のような知見が得られた。

① マルチタスク学習の有無による魅力度推定精度を調査したところ、表 2 および表 3 のような結果が得られ、当該データ拡張により魅力度推定精度が向上したことが確認された。これにより、本アプローチの有効性が確認された。なお、表 2 はマルチタスク学習におけるサブタスクとして、料理の種類分類、魅力度のクラス分類、撮影仰角のクラス分類、撮影回転角の角度差推定を採用したモデルと、従来のシングルタスクモデルの精度を比較した結果である。また、表 3 は、マルチタスク学習と転移学習を併用した際の精度比較であり、表 2 で最も高精度であった「料理の種類分類」をサブタスクとしたときの結果である。転移元によらず料理の種類分類をサブタスクとするマルチタスク学習によって推定精度が向上しており、本アプローチの有効性が確認された。

② 上記アプローチの有効性について議論の余地は残されているものの、データ拡張手法と深層学習モデルの有効な組み合わせ方に関する知見を与えるものであり、意義深い成果であると考えられる。

(3) 3. の(3)の方法を実施し、下記のような知見が得られた。

① 料理写真を含む料理レシピの魅力度を分析したところ、料理レシピタイトルの長さが料理レシピの魅力度に影響を与えていることが判った。また、料理レシピの魅力度はタイトルだけではなく、画像や調理手順などから総合的に決まるものであることが示唆された。この結果を踏まえ、本研究課題で開発した料理写真の魅力度推定技術を応用することで、料理レシピの魅力度推定に利用することができると考えられる。

② 上記の検討は、本研究課題の研究成果である種々の知見や深層学習モデルの応用先の広さを示唆するものであり、意義深いものであるといえる。

表2 マルチタスク学習の有無による魅力度推定精度 (MAE: 平均絶対誤差)

料理	シングルタスクモデル	マルチタスクモデル			
		サブタスク			
	魅力度推定のみ	料理の種類分類	魅力度のクラス分類	仰角のクラス分類	回転角の角度差推定
鯉のたたき	0.169	0.154	0.167	0.158	0.169
カレーライス	0.111	0.114	0.119	0.107	0.111
鰻井	0.087	0.091	0.095	0.084	0.084
ビーフシチュー	0.088	0.098	0.109	0.105	0.087
ハンバーグ	0.118	0.120	0.137	0.141	0.122
天井	0.104	0.109	0.118	0.130	0.106
カツ丼	0.164	0.155	0.164	0.169	0.163
鉄火丼	0.067	0.063	0.053	0.052	0.067
チーズバーガー	0.103	0.102	0.103	0.116	0.105
フィッシュバーガー	0.112	0.080	0.100	0.094	0.116
平均	0.112	0.109	0.117	0.116	0.113

※ 赤色は料理種内で最高精度であることを示す

表3 UEC FOOD 256 データセット (内の一部の和食画像) に魅力度を付与したデータセットを用いて転移学習した場合の魅力度推定精度 (MAE: 平均絶対誤差)

転移元	シングルタスクモデル		マルチタスクモデル (サブタスク: 料理の種類分類)	
	ImageNet	UEC FOOD 256 (和食)	ImageNet	UEC FOOD 256 (和食)
MAE	0.528	0.500	0.448	0.394

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計5件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 高木七海, 久徳遙矢, 道満恵介, 駒水孝裕, 川西康友, 平山高嗣, 井手一郎
2. 発表標題 品詞に着目した料理レシピタイトルの魅力度分析
3. 学会等名 電子情報通信学会 メディアエクスペリエンス・バーチャル環境基礎研究会
4. 発表年 2023年

1. 発表者名 宮崎 光明, 道満 恵介, 井手 一郎, 目加田 慶人
2. 発表標題 マルチタスク学習を用いた料理写真の魅力度推定における有効なサブタスクの検討
3. 学会等名 第19回情報学ワークショップ
4. 発表年 2021年

1. 発表者名 宮崎 光明, 道満 恵介, 井手 一郎, 目加田 慶人
2. 発表標題 料理写真の魅力度推定に適したマルチタスク学習の検討
3. 学会等名 電子情報通信学会 メディアエクスペリエンス・バーチャル環境基礎研究会
4. 発表年 2022年

1. 発表者名 宮崎 光明, 服部 竜実, 道満 恵介, 平山 高嗣, 川西 康友, 井手 一郎, 目加田 慶人
2. 発表標題 料理写真の魅力度推定精度向上のための視線停留分布に対するデータ拡張手法の検討
3. 学会等名 動的画像処理実利用化ワークショップ2021
4. 発表年 2021年

1. 発表者名 宮崎 光明, 道満 恵介, 平山 高嗣, 川西 康友, 井手 一郎, 目加田 慶人
2. 発表標題 料理写真の魅力度推定において有効な画像特徴量の検討
3. 学会等名 第18回情報学ワークショップ
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	井手 一郎 (Ide Ichiro) (10332157)	名古屋大学・大学院情報学研究科・教授 (13901)	
研究分担者	目加田 慶人 (Mekada Yoshito) (00282377)	中京大学・工学部・教授 (33908)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------