Vol.2019-IS-150 No.4
2019/11/23

クラウドソーシングにおける品質改善手法の検討

金地 紗里奈[†1]　小板 隆浩[†2]

概要：現在，人工知能技術の機械学習に用いるデータ収集などで大量のデータを集めるために，ワーカーに対して多様なタスクを容易かつ安価での委託が可能であるクラウドソーシングが利用されている．しかし，クラウドソーシングでは不特定多数のワーカーにタスクを委託することになるので，処理結果の正解率がワーカーの能力ややる気によって大きく異なってしまい品質の面で問題がある．これらの問題を解決するためにはタスク設計や品質管理に対する手法が重要となる．本研究では，クラウドソーシングにおける高品質な処理結果の実現を目的として，タスクやワーカーに対する評価実験を行い，品質改善手法について検討する．

キーワード：クラウドソーシング，Amazon Mechanical Turk，画像認識

1. はじめに

1.1 クラウドソーシング

クラウドソーシングとは，オンライン上で不特定多数のワーカーに人間の能力が必要なタスクを委託することである．クラウドソーシングの中でもマイクロタスク型の実施形態では，人による事前データを必要とする人工知能技術の機械学習に用いるデータ収集などに使われている．機械学習などに用いられるための大量のデータを集めるには，大勢のワーカーの力が必要となる．そこで，ワーカーに対して多様なタスクを容易かつ安価での委託をするためにクラウドソーシングが利用される[1]．クラウドソーシングでは，タスクを委託する人はクライアントと呼ばれ，クラウドソーシングサイトにタスクを委託し，報酬を支払う．タスクを処理する人はワーカーと呼ばれ，クラウドソーシングサイトからタスクを受注し，受けたタスクの処理結果を納品することで報酬を受け取る．こうしてクラウドソーシングは，オンライン上で不特定多数のワーカーに対して容易かつ安価にタスクを委託することができ，コンピュータには難しいタスクを大量に処理することができる．

1.2 研究目的

先行研究では，クラウドソーシングにおける高品質な処理結果の実現を目指すため，ワーカーによる評価要素が品質に及ぼす影響を考察し，クラウドソーシングにおけるワーカーが持つ性別・年齢・タスクの承認数といった評価要素を用いて，品質に及ぼす影響を明らかにした[2]．しかし，国ごとによる認知度の差により，正確な回答を得にくかったことも示された．本研究では，先行研究と異なる条件下で，異なるタスクに対する評価実験を行う．クラウドソーシングにおける品質改善手法の実現を目的として，タスクやワーカーに着目し，品質改善手法について検討する．

2. 評価実験

2.1 実験方法

本研究で用いるタスクは静止画(図 1)と動画(図 2)を用いた対象物の抽出とした．クラウドソーシングサイトとして Amazon Mechanical Turk を用いて，静止画と動画を各 10 人のワーカーに提示し，対象を記述する単語を 10 個記述することを依頼した．ワーカーに対する報酬は 1 タスクあたり 0.4 ドルとした．



図 1　実験で用いた静止画



図 2 実験で用いた動画

†1 同志社大学院理工学研究科
†2 同志社大学理工学部

ⓒ 2019 Information Processing Society of Japan

表１　タスクの処理結果

HIT ID ▲	Worker ID	Lifetime Approval Rate	Tag1	Tag10	Tag2	Tag3	Tag4	Tag5	Tag6	Tag7	Tag8	Tag9
304QEQWKZQROQYJQ4ESTCJQVZ86O0I	A2WGW5Y3ZFBDEC	100% (3/3)	shirt	pants	baton	field	cap	seat	sign	light	wristband	railing
304QEQWKZQROQYJQ4ESTCJQVZ86O0I	A183WYXN12P2TJ	100% (2/2)	stadium	players	baseball bat	audience	lights	field	scoreboard	seats	advertisements	grass
304QEQWKZQROQYJQ4ESTCJQVZ86O0I	A3PGUPNMOU5BPW	0% (0/0)	field	lights	people	hands	shirts	bats	watch	fence	grass	team

2.2 Amazon Mechanical Turk

Amazon Mechanical Turk(Mturk)は，アメリカのアマゾン社の Amazon Web Service の１つであり，コンピュータプログラムを人間の知能と組み合わせて，コンピュータだけでは不可能なタスクを処理するマイクロタスク型のクラウドソーシングサイトである．細かく分割されたタスクを少額の金額で大量のワーカーに委託することができ，人間には可能だが機械には難しいとされる自然言語理解，音声認識などの領域においての機械学習をはじめとするデータ駆使アプローチ[2]に用いられている

3. 実験結果

静止画の実験結果について述べる．実験では，ワーカーは図 1 の静止画像を用いて，対象物の認識を行た．本研究では,ワーカーが正しく返答できれば正解として 10 個の対象の認識を行う．表 1 に処理結果の一例を示す．表中の HIT_ID は，ワーカー名が書かれており，Tag1 から Tag10 までには対象物を記述する単語が書かれている．図 3 に，ワーカーから返ってきた処理結果の対象物ごとの処理結果の人数を示す．5 人以上に認識されたものを表示しており，一番高かったのは light の 9 人だった．ワーカーの正答率は平均 90%とであった．

4. 考察

先行研究では，食品画像を用いてワーカーによる評価要素が品質に及ぼす影響を考察した．以前の食品画像での認識は日本食を用いたことで認知度が低く，正解率はあまり高くなかった．今回のタスク設定では，問題の認知度を高くし，粒度を細かくしたことで正解率は高くなった．しかし，明らかに画像に存在しないものやありそうな対象物を想像して書くワーカーが少数存在したことから，ワーカーの丁寧さが求められる．品質改善のための解決策として，ワーカーの丁寧さの事前質問やタスク自体の属性とワーカーの属性を検討することがあげられる．また，画像認識においては，ワーカーの問題の認知度と粒度の細かさが重要である可能性についても考慮する必要がある．

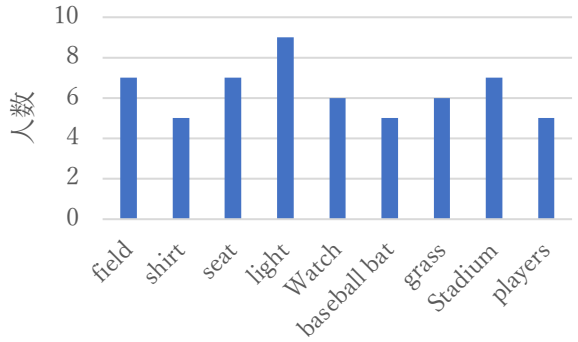


図 3 対象物ごとの処理結果の人数

5. まとめ

本研究では，クラウドソーシングにおける品質改善手法の実現を目的として，タスクやワーカーに対する評価実験を行い，品質改善手法について検討した．先行研究と異なるタスクに対して，正解率の変化が見られた．タスクの粒度を細かくすることで，ワーカーの誤答が減り，品質改善の可能性を示した．今後の課題として，同じタスクでの評価実験の実施や，事前質問によりワーカーの丁寧さの定量的な評価があげられる．

参考文献

[1] 大澤翔吾，柳井啓司，クラウドソーシングによる食事画像認識モデルの自動構築，DEIM2013 The 5th Forum on Data Engineering and Information Management，C4-4，2013.

[2] 金地紗里奈，小板隆浩，クラウドソーシングにおける品質に対するワーカーの影響，研究報告情報システムと社会環境（IS），2019-IS-148，No. 5，pp. 1-6，2019.

[3] 鹿島久嗣，梶野洸，クラウドソーシングと機械学習，人工知能学会誌，Vol. 27, No. 4, pp. 381–388, 2012.

ⓒ 2019 Information Processing Society of Japan