

クラウドソーシングを用いた異常検知におけるメタ属性活用の有効性

Effectiveness of Meta-Attribute Utilization in Anomaly Detection by Crowdsourcing

玉野 樹[†] 板野 竜也[†] 谷津 帆乃果[‡] 小板 隆浩[†]
Tatsuki Tamano Ryuya Itano Honoka Tanitsu Takahiro Koita

1. はじめに

近年, IoT の普及に伴い監視カメラが低価格で手に入るようになり, 公共の場だけでなく一般家庭にも設置されるようになった. しかし, 監視カメラの台数は増加しているが, 犯罪の検挙率は依然として低い水準を保っている [1]. この理由として, 監視カメラ動画の事後確認では, 犯罪の立証が難しいということが挙げられる. 犯罪の検挙率を向上させるためには, 監視カメラを常時監視する人員が必要となるが, 莫大なコストがかかってしまう. そのため, 犯罪などの異常を自動で検知するシステムが必要となる.

異常を自動で検知するシステムとして, 板野らは深層学習モデルとクラウドワーカーの協調による異常検知手法を提案している [2]. 板野らの手法では, 深層学習モデル層とクラウドソーシング層の 2 層構造で異常検知を実現しており, まず深層学習モデル層で異常検知を行い, 異常がない可能性が高い部分をクラウドソーシングで異常検知を行う. 板野らの手法では異常検知を, 自動かつ高精度, 低コストで実現している. しかし, 板野らの手法ではクラウドソーシング層で提供する動画が無加工であり, クラウドワーカーは回答に時間がかかってしまうため非効率であった.

先行研究では, クラウドワーカーに提供する動画の再生速度を変更することによる精度や効率の変化を評価した [3]. 先行研究の手法では, 異常がある動画の場合, 再生速度を上げることで, より高精度かつ高効率な結果を示した. 一方, 異常がない動画の場合, 再生速度を上げることで, 板野らの研究よりも低精度かつ低効率な結果を示した. 原因として, 動画内での異常判定が難しかったからであると考えられる. これは, 再生速度を上げていない 1 倍速の場合でも, 回答精度が低い結果になったからである. そのため, クラウドワーカーによる異常検知において, 判定が難しいタスクの回答精度を向上させる必要がある.

本研究では, クラウドワーカーによる異常検知において, 判定が難しいタスクの回答精度を向上させるために, クラウドワーカーのメタ属性を活用する. メタ属性とは,

回答に関する情報のことである. これは, 回答者の実績や確信度, 回答時間などの, 回答に関する補足的な情報を提供する. メタ属性を活用することにより, 判断が難しいタスクの回答精度を上げることを目的とし, 提案手法の評価実験を行い, 精度について評価する. 本研究では, 以下の項目に貢献している.

- クラウドワーカーによる異常検知において, 判断が難しいタスクの回答精度を向上にワーカーのメタ属性を活用することは有効であること示した.
- 回答に対する確信度を活用した手法よりも, 回答時間を活用した手法の方が精度向上に繋がるという結果を示した.

2. 関連研究

2.1 クラウドソーシングによる異常検知手法

先行研究では, クラウドワーカーに提供する動画の再生速度を変更することによる精度や効率の変化を評価した. 先行研究の手法では, 犯罪が行われている動画の場合, 再生速度を上げることで, より高精度かつ高効率な結果を示した. 一方, 犯罪が行われていない動画の場合, 再生速度を上げることで, 板野らの研究よりも低精度かつ低効率な結果を示した. この原因として, 犯罪が行われていない動画内で, 犯罪が行われているかの判定が難しかったからであると考えられる. これは, 再生速度の変化に関わらずクラウドワーカーの回答の精度が低かったからである. そのため, クラウドワーカーによる異常検知において, 判定が難しいタスクの回答精度を向上させる必要がある.

2.2 確信度を活用した手法

小山らは, クラウドソーシングで回答を集める際に, クラウドワーカー自身の回答に対する確信度を用いることで精度の向上を示した [4]. 具体的には, 確信度によってクラウドワーカーの回答の重みを変更する仕組みになっている. 例えば, 回答に対する確信度が大きい場合は, その回答は結果に大きく反映される. 一方で, 回答に対する確信度が小さい場合は, その回答は結果に小さく反映される. このようにクラウドワーカーの自己申告で確信度を回答してもらうことで, 精度の向上に繋がることを示した. このことから, クラウドワーカーの確

[†] 同志社大学大学院理工学研究科 〒610-0394 京田辺市多々羅都谷 1-3

[‡] 同志社大学理工学部 〒610-0394 京田辺市多々羅都谷 1-3

信度のようなメタ属性の活用が回答精度の向上に有効であることが分かる。しかし、小山らの研究では、画像分類タスクや一般知識を問うタスクなど限定的な条件で実験を行っており、異常検知の場合もメタ属性の活用が有効であるか確かめる必要がある。また、確信度の他にもクラウドワーカーの回答時間といったようなメタ属性の活用も回答精度の向上に有効であるか確かめる必要がある。

3. 提案手法

本研究では、メタ属性として、クラウドワーカーの回答時間や回答に対する確信度を活用することで、判別が難しいタスクにおける回答精度の向上への有効性を評価する。具体的なメタ属性の活用手法としては、クラウドワーカーの回答時間や回答に対する確信度を0から100に数値化し、その数値に伴ってクラウドワーカーそれぞれの回答の重みを変更させる。例えば、クラウドワーカーの回答時間によって回答の重みを変更させる場合は、式(1)のような計算を行う。ここで、 w_j は対象のクラウドワーカーの回答の重み、 t_j は対象のクラウドワーカーの回答時間、 n はタスクを依頼するクラウドワーカーの数を示す。

$$w_t = \frac{t_j}{\sum_{i=0}^n t_i} \quad (1)$$

また、回答の確信度によって回答の重みを変更する場合も式(1)と同様に、式(2)のような計算によって表すことができる。ここで、 c_j は対象のクラウドワーカーの回答に対する確信度を示す。

$$w_c = \frac{c_j}{\sum_{i=0}^n c_i} \quad (2)$$

式(1)及び式(2)の計算によってクラウドワーカーの回答時間や確信度によって回答の重みを変更することで、回答の精度向上が期待できる。例えば、図1のようにクラウドソーシングで回答を収集する際によく用いられる多数決を考えてみる。それぞれのクラウドワーカーの回答の重みの合計を1としたとき、それぞれの回答の重みは0.33となり、全てのクラウドワーカーが同じ能力を持っているという仮定に従っている。ただし、クラウドソーシング環境のように、クラウドワーカーの能力や誠実さに大きな差がある場合には、クラウドワーカーごとに回答の重みを変更する必要があると考えられる。そのため、図2, 3のようにクラウドワーカーごとに回答の重みを変更することで回答精度を高めることができると考える。

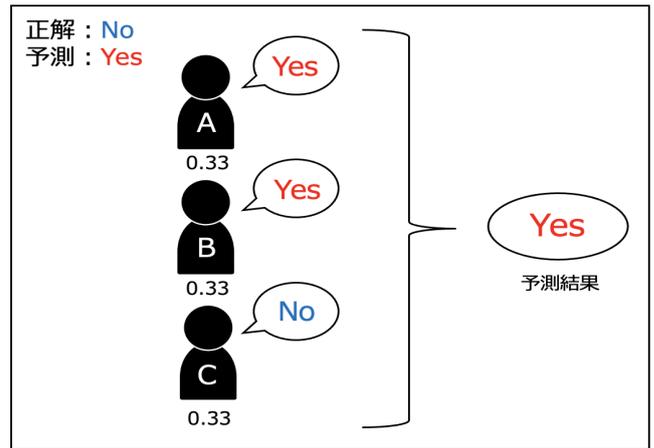


図 1: 多数決による回答収集

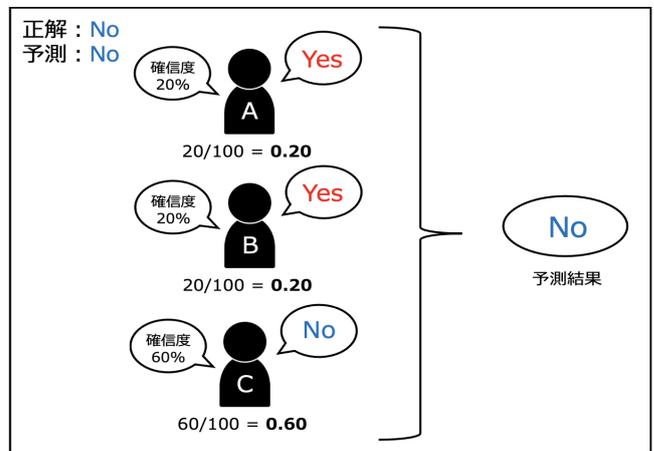


図 2: 確信度を活用した回答収集

4. 評価実験

4.1 実験方法

本研究では、板野らの先行研究と同様に、異常検知用データセットとしてUCF-Crime Dataset*を使用して実験を行う。データセットは、950本の正常な動画画像、及び950本の異常な動画画像で構成されており、異常な動画画像には、公共安全に大きな影響を及ぼす主要な13種類の異常（虐待、逮捕、放火、暴行、事故、強盗、爆発、喧嘩、泥棒、銃撃、窃盗、万引き、破壊行為）のいずれかが含まれている。本研究では、予備実験としてUCF-Crime Datasetから本実験で使用する判断が難しい動画画像10本を決定する。10本の動画画像は、異常がある動画画像を5本、異常がない動画画像を5本にするようにする。本実験では、予備実験で決定した10本の動画画像を使用して、メタ属性活用が異常検知精度向上に有効であるか評価する。板野らの先行研究に従って、予備実験及び、

*<https://www.crcv.ucf.edu/projects/real-world/>

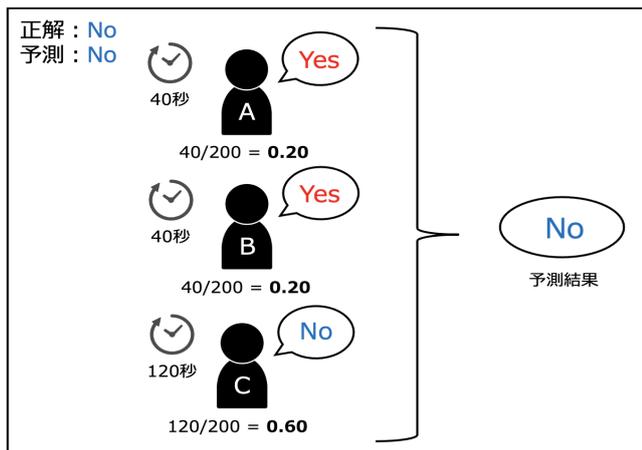


図 3: 回答時間を活用した回答収集

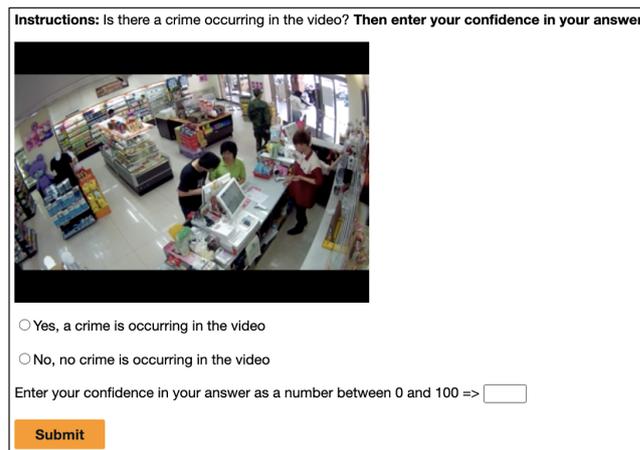


図 4: タスクの概要

本実験で使用する動画像の長さは全て 10 秒程度に設定する。

本研究ではクラウドソーシングプラットフォームとして、Amazon Mechanical Turk (AMT)[†]を使用する。AMT では常時 2,000 人以上のワーカーが待機しており、他のクラウドソーシングプラットフォームと比較して一定レベルのリアルタイム性を確保することができる。ワーカーに与えるタスクは、説明書きや動画などを HTML や CSS で記述して自由に配置できる。

4.2 予備実験

予備実験では、UCF-Crime Dataset から異常がある動画像を 10 本、異常がない動画像 10 本使用し、クラウドワーカー 10 人を対象に異常があるか判断してもらい、タスクそれぞれの正解率を算出した。この結果から、正解率が 30~70% となった動画像を判断が難しい動画像とし、その中から異常がある動画像 5 本、異常がない動画像 5 本の計 10 本の動画像を本実験で使用する。

4.3 本実験

本実験では、予備実験で判断が難しい動画像と決定した 10 本の動画像を使用し、クラウドワーカー 500 人を対象に動画像に異常があるか判断してもらう。1 本の動画像につき図 4 のようなタスクを作成し、クラウドワーカー 50 人に回答してもらう。タスクでは異常があるかどうか Yes/No で答えてもらうと同時に、その回答に対する確信度を 0 から 100 までの数値で入力してもらう。この回答結果から、回答時間や回答に対する確信度を取得し、式 (1) 及び、式 (2) からクラウドワーカーそれぞれの回答の重みを算出する。算出した回答の重みを用いて全てのクラウドワーカーの回答を収集する。

また、タスクあたりのワーカー数が正解率に与える影

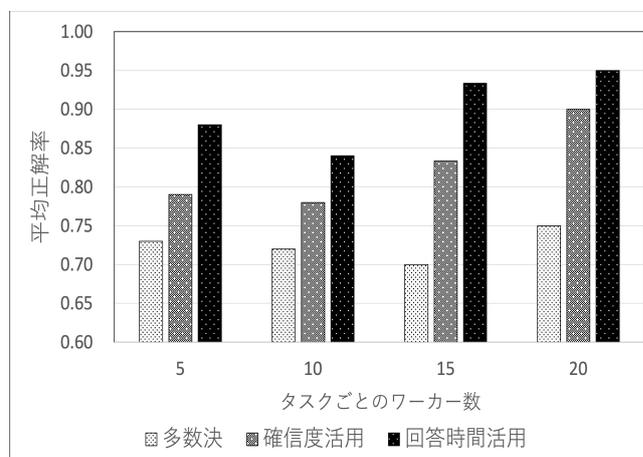


図 5: 実験結果

響を調べるために 50 人のクラウドワーカーを一定数のグループに分割し、各グループで異常検知を行いその正解率を平均した。正解率を平均した結果を平均正解率とする。つまり、10 タスクを依頼する場合、7 タスクで正しい結果が得られ、3 タスクで正しくない結果が得られた場合は、平均正解率は 70% となる。5 人、10 人、15 人、20 人ずつのグループにおいて単純な多数決、確信度を活用した手法、回答時間を活用した手法の 3 つを比較する。

4.4 実験結果

実験結果を図 5 に示す。実験結果から多数決と比べて、メタ属性活用によって平均正解率が向上していることが分かる。また、確信度を活用する手法よりも回答時間を活用する手法の方がどのワーカー数のグループでも高い平均正解率を得ることができた。

[†]<https://www.mturk.com/>

4.5 考察

実験結果から、クラウドソーシングで判断が難しい動画画像から異常検知を行う際に、確信度や回答時間といったメタ属性を活用すること回答精度の向上に有効であることが分かった。今回の実験では、予備実験で正解率が0~30%の動画画像が得られなかったため、判断が難しい動画として、正解率30~70%の動画画像を使用した。しかし、正解率が0~30%程度の動画画像に関しては、本手法が有効であるか評価を行っていないため、今後は、より正解率の低い判断の難しいタスクにも本手法が有効であるのか検証する必要がある。

5. まとめ

本研究では、クラウドワーカーによる異常検知において、判定が難しいタスクの回答精度を向上させるために、ワーカーのメタ属性の活用は有効であるか評価した。結果として、クラウドソーシングで回答を収集する際に一般的に用いられる多数決と比較して、メタ属性の活用は有効であるという結果が得られた。また、回答に対する確信度を活用した手法よりも、回答時間を活用した手法の方が精度の向上に繋がるという結果が示された。今後は、正解率が0~30%程度のより判断が難しい異常検知タスクの精度向上に取り組む必要がある。

参考文献

- [1] 法務省, 令和3年版犯罪白書,
https://www.moj.go.jp/housouken/housouken03_0049.html
- [2] 板野竜也, 野原智哉, 小坂隆浩他, “深層学習モデルとクラウドワーカーの協調による動画画像異常検知手法,” 研究報告オーディオビジュアル複合情報処理 (AVM), vol. 2022, no. 2, pp. 1-6, 2022.
- [3] 玉野樹, 板野竜也, 小坂隆浩, “クラウドワーカーによる異常検知に向けた動画画像提供方式の評価,” 電子情報通信学会学生会研究発表講演会, 2023.
- [4] 小山聡, 馬場雪乃, 櫻井祐子, 鹿島久嗣. “クラウドソーシングにおけるワーカーの確信度を用いた高精度なラベル統合.” 人工知能学会全国大会論文集 第27回 (2013), 一般社団法人人工知能学会, 2013.