

令和 4 年 6 月 9 日現在

機関番号：34315

研究種目：若手研究

研究期間：2020～2021

課題番号：20K19831

研究課題名（和文）詳細かつ解釈可能な暴力度レーティング技術の開発

研究課題名（英文）Development of Finegrained and Explainable Violence Rating Approaches

研究代表者

ワン ユ (WANG, YU)

立命館大学・情報理工学部・助教

研究者番号：60724169

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究は、詳細かつ解釈可能な暴力度レーティング技術の開発を目的として、まず、[1]相対的な属性値を用いたラベリング手法のアルゴリズムを具体化した上、実装を完成させた。また、小規模の検証実験において、開発したラベリングツールの有効性を確認した。その上、[2]詳細かつ解釈可能な主観的属性の認識を実現するために、CRAとRMCSAなどの深層モデルのモジュール、及びDPCNNとMSCNNなどの深層モデルのアーキテクチャを開発し、大規模の検証実験において性能の優位性を確認した。

研究成果の学術的意義や社会的意義

本研究で開発した暴力度レーティング技術は実用性が高く、実用できれば、ビデオごとに暴力度の詳細推定及び推定根拠の提示が可能になる。また、この技術により推定した暴力度から、ビデオごとに視聴に適した年齢層を見積もることができるため、青少年をはじめとする多くの利用者にとって、より自由で安全な視聴空間が確保できる。また、本研究で開発したラベリングツール、及び深層モデルのモジュールやアーキテクチャは、画像・映像における主観的属性認識タスクの最も基本的な課題の解決にも貢献できる。研究成果は画像・映像の分類、要約、検索などの研究にも大きな推進力を与えることになる。

研究成果の概要（英文）：In this research project, the objective is to develop fine-grained and explainable violence rating approaches. We mainly worked on the following aspects: (1) We developed and implemented the relative attribute based labeling tool. We evaluated its practical usages on real labeling task with limited scale and confirmed its efficiency. (2) We proposed new modulars (CRA and RMCSA) for deep models, as well as new deep model architectures (DPCNN and MSCNN). Their performance was extensively evaluated.

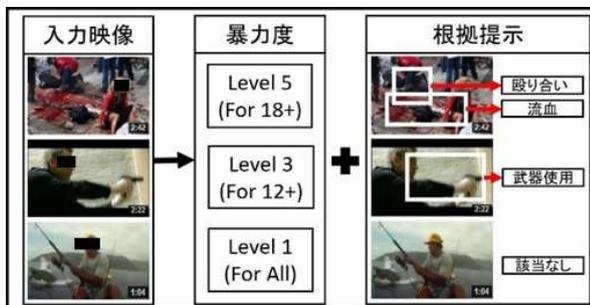
研究分野：画像処理

キーワード：暴力度レーティング 動画画像認識

### 1. 研究開始当初の背景

ビデオコンテンツに含まれる暴力・残虐的表現が、成長過程の青少年にネガティブな影響を与えることは、これまで社会学者・心理学者によって実証されてきた。暴力・残虐表現を含む有害コンテンツから青少年を守るために、放送時間帯の配慮、保護者の番組選択支援など、様々な措置が取られてきたが、インターネットにつながるスマホやタブレットなどの浸透により、ウェブ上に散在するビデオコンテンツにアクセスする手段は多様化・パーソナル化となり、視聴内容をコントロールすることは困難になっている。そのため、有害コンテンツの自動解析・フィルタリング技術へのニーズが高まっている。

動画画像解析分野では、映像特徴を利用した暴力コンテンツ解析の研究が、これまで様々な手法で試みられてきた。しかし、これらの研究は、主にビデオから暴力シーンを検出することに着目しており、ビデオ全体の暴力度を評価する有効手段は開発されていなかった。より実用性の高いフィルタリング技術の実現を目指して、「ビデオ単位の暴力度レーティング」というコンセプトが提案され、初となる暴力度レーティング手法が開発された (Y. Wang ら, ACPR'15)。この手法により、暴力度をビデオごとに評価することができるだけでなく、視聴に適した年齢層も推定することが可能となったため、自動フィルタリングの実用化に向かって確実な前進を成し遂げた。また、その後、さらに精度と速度の両面から暴力度レーティングの改善を繰り返された (Y. Ji ら, ICIP'19)。しかし、推定できる暴力レベルはまだ3段階に止まり、推定根拠も明示することができない。実用化をさらに推進する観点から、暴力度レーティングを、社会一般に認められている分類基準と同等、または、それ以上の細かさでレーティングし、その根拠も併せて提示する段階へ発展させることが必要不可欠である。



### 2. 研究の目的

本研究は、詳細かつ解釈可能な暴力度レーティング技術の開発を目的とする。暴力度は映像の一つの「属性」として捉えることができる。映像コンテンツの属性認識は、コンピュータビジョン分野の最も重要な課題の一つであるが、これまでの研究はほとんど、行動の種類やイベントの種類など、曖昧性・抽象性の低い客観的属性を認識対象として扱ってきた。「暴力度」のような一定の曖昧性をもち、抽象性が高い主観的属性の認識に関する限られた既存研究では、客観的属性の認識手法を流用するレベルに止まっている。本研究の目的を実現するには、幾つか困難な課題が立ちだかる。一つ目は、大規模な学習データセットが必要であるが、映像暴力度の独特の曖昧性から学習データセット作成時のラベリング作業が極めて困難であるという点である。二つ目は、動画画像解析に最も有効とされる深層モデル学習は、そもそもその認識結果を解釈しにくい性質をもっているため、それを採用しながら暴力度の推定根拠を提示しようとするのが困難であるという点である。これらの問題に対して、本研究では、まず、相対的な属性値を用いたラベリング手法を提案し、ラベリング作業の簡易化・安定化を図ることで、高品質かつ大規模なデータセットの作成を可能にする。また、モジュール型深層モデルを開発し、複雑な認識タスクを複数の独立機能をもつモジュールで対応する戦略により、推定結果の解釈を容易にする。

### 3. 研究の方法

本研究では、詳細かつ解釈可能な暴力度レーティング技術を実現するために、ラベリングおよびモデルアーキテクチャーの二つの側面からの2課題を実施した。

#### [1] 相対的な属性値を用いたラベリング手法の開発

ラベリング作業を行う際に、「このビデオなら暴力度はどれぐらいか」という質問には答えにくい、二つの映像から「どちらがより暴力的か」という質問には答えやすい。この知見に基づき、本研究では、まず、ラベリング段階で、絶対的な暴力度の代わりに相対的な暴力度をリストワイズ比較(数個のビデオを比較して暴力度の強弱順に並べ替える操作)により獲得し、さらに、多数のリストワイズ比較の結果から、確率的なグラフィカルモデルと確率伝搬法を用いて、絶対的な暴力度ラベルを算出する戦略を採る。この戦略により、人や時期によって生じるラベリング基準の揺れを吸収することができるため、多数の人員をラベリング作業に投入しても、データセットの品質を維持することができる。と考える。

#### [2] モジュール型深層モデルの開発

暴力度レーティング問題の抽象性・複雑性に対処するために、入力ビデオの局所から全体まで異なるスケールの視覚特徴をキャッチする必要がある。また、暴力度ラベルは離散かつ強弱関係のある変数のため、ラベル内部の構造を学習する仕組みが必要となる。さらに、レーティングパイプラインの処理プロセスに関して、解釈性と透明性も求められる。本研究では、モジュール型深層モデルを開発することで、以上の問題点を克服することを狙う。具体的には、汎用型アテン

ションモジュール，人体パーツを強調する深層モデルおよび複数の受容野を持つ深層モデルの開発に力を入れた。

#### 4. 研究成果

研究課題[1]に関して、「相対的な属性値を用いたラベリング」アルゴリズムを具体化した上，実装を完成させた。計画では20名前後の学生を動員して，実装したラベリングツールを使ってラベリングを行い，手法の有用性を検証する評価実験を行う予定でしたが，コロナ禍が続く中，7名の学生アルバイトしか雇用できず，ラベリング作業は予定通りできなかった。従って，ラベリングアルゴリズムおよびツールの性能評価に関して，大規模の評価実験ができず，小規模の検証実験のみを行った。提案したアルゴリズムおよび開発したツールの有効を確認できたものの，さらなる検証が必要と考えられる。

研究課題[2]に関して，㊸ Contrastively-reinforced Attention (CRA) および ㊹ Recursive Multi-scale Channel-spatial Attention (RMCSA) などのアテンション機構を用いた深層モデルのモジュールや，㊺ Discriminative Part CNN (DPCNN) という人体パーツを強調する畳み込みニューラルネットワーク，㊻ Multi-scale CNN (MSCNN) という複数の受容野を持つ畳み込みニューラルネットワーク，をそれぞれ提案・実装した。各手法の性能を画像・映像認識タスクの標準データセットにおいて大規模の評価実験を行った。また，暴力度の詳細レーティングタスクにおいて，小規模の評価実験ではあったが，既存手法との比較実験を行い，提案手法の性能や解釈性においての優位性を確認できた。

## 5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 0件/うちオープンアクセス 5件）

1. 著者名 ZHAO Longjiao, WANG Yu, KATO Jien	4. 巻 E104.D
2. 論文標題 Rethinking the Rotation Invariance of Local Convolutional Features for Content-Based Image Retrieval	5. 発行年 2021年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 174 ~ 182
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2020EDP7017	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 JI Ying, WANG Yu, KATO Jien, MORI Kensaku	4. 巻 E103.D
2. 論文標題 Predicting Violence Rating Based on Pairwise Comparison	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 2578 ~ 2589
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2020EDP7056	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 王彥, 加藤ジェーン	4. 巻 2
2. 論文標題 行動認識のためのTwo-stream 3D BagNet	5. 発行年 2021年
3. 雑誌名 画像ラボ	6. 最初と最後の頁 34 ~ 39
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 LIU Dichao, WANG Yu, MASE Kenji, KATO Jien	4. 巻 E105.D
2. 論文標題 Recursive Multi-Scale Channel-Spatial Attention for Fine-Grained Image Classification	5. 発行年 2022年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 713 ~ 726
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2021EDP7166	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 WANG Yu, CAO Cong, KATO Jien	4. 巻 E105.D
2. 論文標題 Discriminative Part CNN for Pedestrian Detection	5. 発行年 2022年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 700 ~ 712
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2021EDP7057	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 王 彧、小平 美沙季、加藤 ジェーン	4. 巻 62
2. 論文標題 広域特徴と狭域特徴を用いた画像情報の部分欠損に頑健な歩行者検出	5. 発行年 2021年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 2069 ~ 2078
掲載論文のDOI (デジタルオブジェクト識別子) 10.20729/00214251	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Liu Dichao, Yamasaki Toshihiko, Wang Yu, Mase Kenji, Kato Jien	4. 巻 9
2. 論文標題 TML: A Triple-Wise Multi-Task Learning Framework for Distracted Driver Recognition	5. 発行年 2021年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 125955 ~ 125969
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2021.3109815	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計7件 (うち招待講演 0件 / うち国際学会 3件)

1. 発表者名 Dichao Liu, Yu Wang Kenji Mase and Jien Kato
2. 発表標題 Multi-Task Attention Learning for Fine-grained Recognition
3. 学会等名 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU研究会)
4. 発表年 2020年

1. 発表者名 Longjiao Zhao, Yu Wang, Yoshiharu Ishikawa and Jien Kato
2. 発表標題 Rethinking the local similarity in content-based image retrieval
3. 学会等名 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU研究会)
4. 発表年 2020年

1. 発表者名 内田準也, 王彧, 加藤ジェーン
2. 発表標題 Two-stream 3D BagNetによる人物行動認識
3. 学会等名 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU研究会)
4. 発表年 2020年

1. 発表者名 杉山瑠菜, 王彧, 加藤ジェーン
2. 発表標題 社会関係を活用した映像会話シーンにおける雰囲気への推定
3. 学会等名 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU研究会)
4. 発表年 2020年

1. 発表者名 Dichao Liu, Yu Wang, Jien Kato and Kenji Mase
2. 発表標題 Contrastively-reinforced Attention Convolutional Neural Network for Fine-grained Image Recognition
3. 学会等名 The 31st British Machine Vision Conference (BMVC2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Dichao Liu, Yu Wang, Kenji Mase and Jien Kato
2. 発表標題 Attention-based Multi-task Learning for Fine-grained Image Classification
3. 学会等名 IEEE 2021 International Conference on Image Processing (ICIP2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Longjiao Zhao, Yu Wang and Jien Kato
2. 発表標題 Rotation Invariance Analysis of Local Convolutional Features in Image Retrieval
3. 学会等名 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (国際学会)
4. 発表年 2021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
---------------------------	-----------------------	----

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
---------	---------