

令和 3 年 6 月 22 日現在

機関番号：82670

研究種目：若手研究

研究期間：2019～2020

課題番号：19K20310

研究課題名（和文）敵対的生成ネットワークによる不審動作検出器および生成器の学習

研究課題名（英文）Generative adversarial networks for unnatural human motion detection and generation

研究代表者

三木 大輔（Daisuke, Miki）

地方独立行政法人東京都立産業技術研究センター・開発本部開発第三部情報技術グループ・副主任研究員

研究者番号：70757343

交付決定額（研究期間全体）：（直接経費） 1,900,000円

研究成果の概要（和文）：人物の関節位置に関する時空間的な情報を扱う深層ニューラルネットワーク（Deep Neural Network, DNN）モデルおよびその学習手法を提案することで、人物行動ローカライゼーションを可能とした。さらに人物動作データからの異常検知が可能であることを確認した。これらの結果は、人物の危険な動作や不自然な挙動といった曖昧な動作を含むデータや局所的なアノテーションの付与が困難な膨大なデータを用いて特定の動作の識別および異常の検知が可能であることを示唆するものである。

研究成果の学術的意義や社会的意義

申請者らはこれまでに監視カメラ映像から人物の挙動や軌跡等の動作特徴量を取得するための画像処理技術の開発に従事してきた。監視カメラの設置台数は年々増加し、膨大な数の映像が蓄積されていることから、これらの映像に含まれる人物の動作特徴量から異常を検知する手法を確立できれば、映像監視の自動化が期待される。本研究では、人物動作特徴量のような多変量時系列データから異常を検出するための、DNNモデルおよびその学習手法の確立を行い、人物動作データからの異常検知を可能とした。さらに数に限りのあるデータセットを用いて異常検出器の汎化性能を改善するための、学習データの拡張手法を確立した。

研究成果の概要（英文）：A deep neural network (DNN) model and a training method for analyzing spatial-temporal human motion data are proposed in this study. We confirmed that the proposed method can detect anomalies in human motion data, and the DNN model trained using our multi-instance learning-based training method can detect ambiguously defined motions, such as unnatural human motions. The results indicate that it is possible to detect anomalies and identify human motion when dealing with large datasets that are difficult to annotate locally or when the data contains ambiguous motions, such as dangerous or unnatural motion. To improve the generalization performance of the anomaly detector in the case of a limited number of datasets, we establish a training data generation method using generative adversarial networks.

研究分野：情報工学

キーワード：人物動作解析 異常検知 深層学習 敵対的生成ネットワーク

1. 研究開始当初の背景

監視カメラの設置台数は年々増加し、膨大な数の映像が蓄積されている。これらの映像に含まれる人物の動作を解析し異常を検知することができれば、映像監視の自動化が期待される。申請者らはこれまでに監視カメラ映像から人物の挙動や軌跡等の動作特徴量を取得するための画像処理技術の開発に取り組んできた。得られた時間的な情報をもつ特徴量を解析する手法ととして、統計的手法や機械学習手法を応用した時系列データを解析する方法があるが、特に深層ニューラルネットワーク (Deep Neural Network, DNN) モデルを用いた手法はその表現能力の高さから、データに含まれる複雑な特徴を抽出し、検知する上で有効な手法として期待されている。時系列データからの異常検知に DNN モデルを用いる方法は、多変量データを取り扱う上で有効であるほか、煩雑な特徴量エンジニアリングの作業が少なく済むため実応用の上でも有利である。DNN モデルを用いる課題として、事前に大規模なデータセットを用いた学習により、そのパラメータの最適化が必要なことが挙げられる。特に DNN モデルを時系列データの解析に適用するためには、一連のデータおよびそのデータを説明するアノテーションの対から成るデータセットを用意することが必要となる。現実的に扱われる異常を含む時系列データは、多変量時系列データであることが多いが、このような異常を含む多変量時系列データは一連のデータの中に異常が含まれていることを把握することが可能である場合においても、それらの異常が含まれる箇所やその程度を把握することは一般に困難である。特に人物の不審な動作はその定義が困難であるため、その不審な動作を含む箇所に対してアノテーションを付与することが困難であった。

2. 研究の目的

本研究の目的は監視カメラ映像から人物の不審な動作を検出するための検出器およびその学習方法と、新たなデータの生成手法を確立することである。これまでに、映像中の人物の 2 次元および 3 次元的な姿勢認識を可能としており、特徴量抽出手法については既に開発が済んでいるため、今回の研究では得られた人物動作に関する特徴量から不審な動作を検出するための DNN モデルおよびその学習方法と、その汎化性能を向上させるような学習データの拡張方法を確立する。

3. 研究の方法

人物動作の解析のため、DNN モデルとして人物の関節位置に関する時空間的な情報を入力とし、それぞれの時間における異常の程度の度合を推定値として出力する、多層の一次元畳み込み層および長短期記憶層を持つ構造を採用した。時系列データの解析に用いられる一般的な DNN モデルの学習では、一連の時系列データに対し、その特性を決定付ける上で重要な箇所に対してアノテーションを付与し学習に供することが有効であるが、煩雑な作業を要し、特に人物動作の特徴量のような多変量時系列データにはアノテーションの付与が困難である。そこで本研究ではマルチインスタンス学習に着想を得たランク学習手法 [1] を参考にすることで DNN モデルの学習のための新たな損失関数を提案した。

提案する DNN モデルは長さ T の時系列データ $\mathbf{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^T\}$ を入力した際に $\mathbf{y} = \{y^1, \dots, y^T\}$ を出力する構造とし、入力データにおける異常の有無は既知であるが、それらが含まれる箇所は未知であるものとする。DNN モデルの学習では、まず時系列データに異常を含む (正) または、含まない (負) によってデータを分割し、それぞれのデータを DNN モデルに入力した際に、正のデータに対して高い値、負のデータに対して低い値が出力されることを目標に DNN モデルの重みパラメータを最適化する。各データにおいてその特性を決定付ける上で重要な箇所は未知であるため、ここでは各データにおけるそれぞれ最大の値に着目し、

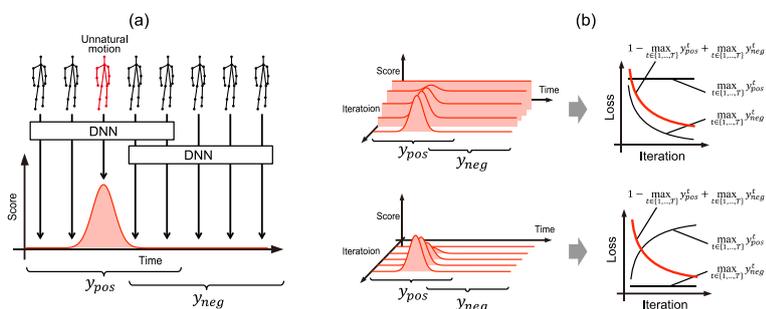


図 1 損失関数の概要 (a) 不自然な動作を含む、または含まないデータに対して推移される値 (b) 推定される値の初期値が (上) 高い場合および (下) 低い場合における推定値と損失の大きさの変化

$$\max_{t \in \{1, \dots, T\}} y_{pos}^t > \max_{t \in \{1, \dots, T\}} y_{neg}^t \quad (1)$$

を満たすように DNN モデルを学習する。ここで、 y_{pos}^t および y_{neg}^t は正および負のそれぞれのデータを DNN モデルに入力した際に出力される $t \in \{1, \dots, T\}$ 時点における推定値である。こ

のような条件を満たす DNN モデルの学習では、損失関数

$$\mathcal{L} = \max\left(0, 1 - \max_{t \in \{1, \dots, T\}} y_{pos}^t + \max_{t \in \{1, \dots, T\}} y_{neg}^t\right) + \lambda \quad (2)$$

を最小化するように DNN モデルのパラメータを最適化する。ここで λ は

$$\lambda = p_1 \sum_{t=1}^{T-1} (y_{pos}^t - y_{pos}^{t+1})^2 + p_2 \sum_{t=1}^T y_{pos}^t \quad (3)$$

で表される正則化項である。提案する DNN モデルの学習は (1) 式の損失関数の最小化問題となるが、これを解析的に解くことは困難であるため、一般的なニューラルネットワークの学習に用いられる数値的解法により反復的にパラメータを更新することで近似的な解を得た。学習データの拡張では敵対的生成ネットワークの仕組みを用いてデータの生成器の学習を行った。生成器は実データの特徴量の分布を学習し、分布の類似したデータを生成する。一方で、識別器では実データと生成データを区別できるように学習されることで、互いの生成精度および識別精度を向上させる。本ステップにおいては、互いの分布に近いデータを生成できるように学習を行い、次のステップで不審動作検知の汎化性能向上のための改良を行った。二者間における敵対的学習では生成された特徴量が実データの存在する局所空間に偏ることが想定されたため、生成器および識別器を、検出器と併せた三者間で敵対的学習を行うことで、検出器の汎化性能が改善されようとして改良した。ここで、

検出器は不審動作の有無を識別できるように最適化されているため、正常動作および不審動作の特徴量が識別困難となるような分布に新たな特徴量が生成される。そこで、検出器と併せた三者間で敵対的に学習させることで、実データ周辺に正常動作と不審動作が区別しづらいようにデータが生成されるように改良した。

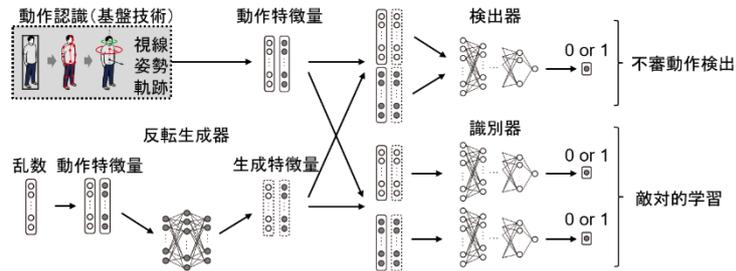


図 2 本研究で提案するネットワーク構造 抽出した特徴量から異常を判定する異常検出器、正常時および異常時の特徴量を生成する生成器、生成データの識別器から成る敵対的生成ネットワークを組み合わせて敵対的学習を行う。

4. 研究成果

人物の歩行動作から不自然な動作が検知可能か確認するために、公開データセット [2] に含まれる「Walk」および「Weird Walks」のタグが付与されたデータを学習用データとして用い提案手法により学習された DNN モデルによる推定値の大きさを比較した。図 3 に不自然な歩行動作および自然な歩行動作を含む一連の人物動作に関する時系列データに対し提案手法により学習された DNN モデルから推定された値を示す。不自然な歩行動作を含むデータに対して推定された値は、多くの時点において自然な歩行動作データから推定された値よりも高いことが確認できる。また、不自然な歩行動作を含むデータに対して推定された値の中でも、幾つかの時点に対しては高い値を示す一方で、他の時点に対しては低い値を示した。本結果から人物の不自然な歩行動作に共通する特徴を含む箇所に対して高い値を推定するように DNN モデルが学習されたと考えられる。以上の結果から、本研究の目的である DNN モデルを用いた多変量時系列データからの異常検知、特に DNN モデルの学習のため異常度の真値を把握することが困難な人物動作データからの異常検知の実現可能性が見出された。特に、このような定義が曖昧な行動の検知技術は、人物が危険な行動、または危険な状態に陥っていないか客観的に把握するシステム等に応用できる可能性がある。

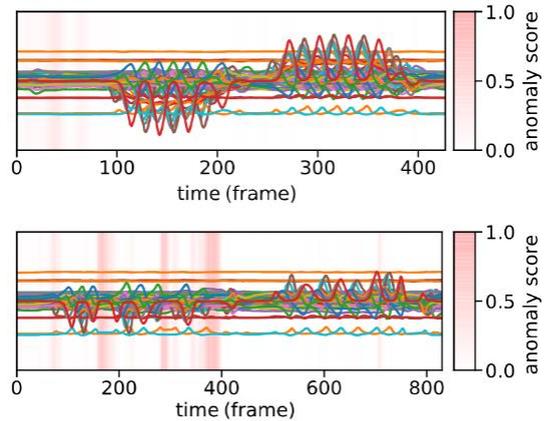


図 3 不自然な動作を (a) 含まないおよび (b) 含む人物動作データおよび推定値

参考文献

- [1] Sultani et al. Real-World Anomaly Detection in Surveillance Videos. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 6479–6488, 2018
- [2] CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu>

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Daisuke Miki, Shinya Abe, Shi Chen, Kazuyuki Demachi	4. 巻 14(4)
2. 論文標題 Robust human pose estimation from distorted wide-angle images through iterative search of transformation parameters	5. 発行年 2019年
3. 雑誌名 Signal, Image and Video Processing	6. 最初と最後の頁 693-700
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s11760-019-01602-5	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 Daisuke Miki, Kazuyuki Demachi	4. 巻 57(9)
2. 論文標題 Bearing fault diagnosis using weakly supervised long short-term memory	5. 発行年 2020年
3. 雑誌名 Journal of Nuclear Science and Technology	6. 最初と最後の頁 1091-1100
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/00223131.2020.1761473	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 Daisuke Miki, Shinya Abe, Shi Chen, Kazuyuki Demachi	4. 巻 7(3), 19-00533
2. 論文標題 Robust human motion recognition from wide-angle images for video surveillance in nuclear power plants	5. 発行年 2020年
3. 雑誌名 Mechanical Engineering Journal	6. 最初と最後の頁 1-8
掲載論文のDOI（デジタルオブジェクト識別子） 10.1299/mej.19-00533	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 -

〔学会発表〕 計3件（うち招待講演 0件／うち国際学会 2件）

1. 発表者名 Daisuke Miki, Shi Chen, Kazuyuki Demachi
2. 発表標題 Weakly Supervised Graph Convolutional Neural Network for Human Action Localization
3. 学会等名 2020 IEEE Winter Conference on Applications of Computer Vision (WACV) (国際学会)
4. 発表年 2020年

1. 発表者名 Daisuke Miki, Shinya Abe, Shi Chen, Kazuyuki Demachi
2. 発表標題 ROBUST HUMAN MOTION RECOGNITION FROM DISTORTED WIDE-ANGLE IMAGES FOR VIDEO SURVEILLANCE
3. 学会等名 27th International Conference on Nuclear Engineering (ICONE) (国際学会)
4. 発表年 2019年

1. 発表者名 Daisuke Miki, Shi Chen, Kazuyuki Demachi
2. 発表標題 Unnatural Human Motion Detection using Weakly Supervised Deep Neural Network
3. 学会等名 2020 Third International Conference on Artificial Intelligence for Industries (AI4I)
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
---------------------------	-----------------------	----

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