

敵対的生成ネットワークを用いた動画異常検知における背景除去の有用性の検討

Usefulness of Background Removal for Video Anomaly Detection Using Generative Adversarial Networks

学 ○塚越 健太郎 (中央大学) 橋本 慧志 (中央大学)
正 梅田 和昇 (中央大学)

Kentaro TSUKAGOSHI, Chuo University, tsukagoshi@sensor.mech.chuo-u.ac.jp
Satoshi HASHIMOTO, Chuo University, hashimoto@sensor.mech.chuo-u.ac.jp
Kazunori UMEDA, Chuo University, umeda@mech.chuo-u.ac.jp

Research on anomaly detection has been conducted to capture anomalies in various situations in daily life. In this study, we aim to build a more accurate anomaly detection system that can detect people and cars behaving abnormally in the video. The background information is considered to be unnecessary for anomaly detection. Therefore, we propose a method to multiply the difference between the generated image and the true value of the optical flow by the image with background removal. We confirm that detection results with high accuracy are produced compared with the conventional method.

Key Words: Anomaly detection, Deep learning, Image processing

1. 序論

日常生活における様々な場面での異常をとらえる異常検知に関する研究[1,2,3]が行われている。一般的に異常検知とは、大多数のデータとは振る舞いが異なるデータを検出する技術のことである。近年、深層学習技術の発展に伴い、異常検知分野でも生成モデルを活用した深層学習による手法が登場している。特に、教師なし機械学習の手法を用いて正常な分布からの異常度を定義することで異常を発見するデータマイニング手法が多く用いられている。異常検知は、侵入検知や詐欺検知といった犯罪行動の検知や、日常生活での転倒事故の検知を可能にするなど多方面での活用が期待されている。

本研究では、動画画内での異常な行動をしている人や車などの、より高精度な異常検知システムの構築を目的とする。具体的なタスクとして、監視カメラから得られるデータを基に、群衆内の自動車や自転車、スケートボードに乗っている人などを検出すべき異常対象とする。

既存手法[2]では、監視カメラ映像を入力とし、生成画像と真値を比較することによって異常検知を行っている。このとき、背景情報は異常検知に影響しないと考えられる。そこで本論文では、前景領域である動物体のみを抽出(以下、背景除去という)した2値画像データを、optical flowの推定結果と真値との差に掛けることを提案する。これによってより高精度な異常検知が可能であることを実験により検証する。

以下、最初に敵対的生成ネットワークを用いた教師なし異常検知のフレームワークについて説明する。次に、検証実験を行い、最後に本手法の有効性を検証する。

2. 提案手法

2.1 pix2pixを用いた異常検知

既存手法であるRavanbakhshら[2]の異常検知手法は、図1のようにGAN(Generative Adversarial Networks)[4]を用いて画像の生成を行う。この手法では二つのGeneratorとDiscriminatorが存在する。一つはframe画像からoptical flowを生成するもので、もう一つはその逆である。これら二つのGeneratorはそれぞれpix2pix[5]の枠組みで学習される。識別フェーズでは、optical flowを予測するモデルからの推定結果と真値との差を

求める。次にframeを予測するモデルからの推定結果と真値をそれぞれAlex-net[6]に入力して得られる中間表現の差を求める。そしてこの2つの差を融合して異常度の算出を行う。

この手法では監視カメラ映像を入力として異常検知を行っているが、我々は、背景情報が異常検知に影響しないと考える。

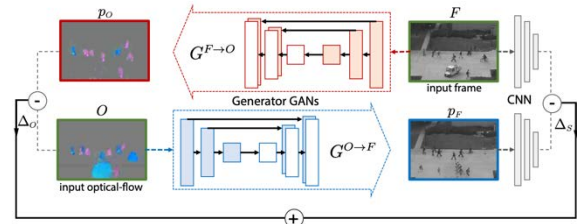


Fig. 1 Conventional framework [2]

2.2 提案手法の概要

提案手法の流れを図2に示す。まず、データに背景除去処理を行い、背景情報を除去した背景除去データを作成する。学習フェーズでは、正常データのみを学習する。識別フェーズでは、異常も含まれたデータを入力とする。optical flowを予測するモデルで出力されたデータと真値の差の絶対値を求め、背景除去データを掛け合わせることで異常度を算出する。背景除去データを掛け合わせることで、推定したoptical flowの動物体のみを注目し、異常検知に不要である背景情報のノイズを減らすことが期待できると考える。

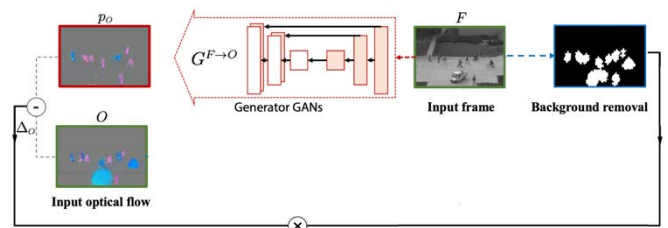


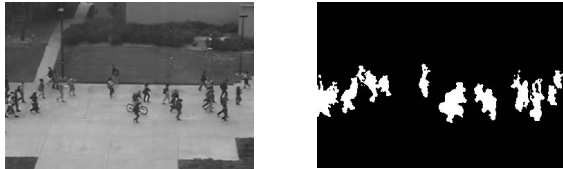
Fig. 2 Proposal framework

2.3 LSBPを用いた背景除去

背景除去にはLSBP(Local SVD Binary Pattern)[7]を用いる。

図 3(a)に対し LSBP により求めた前景領域を図 3(b)に示す。

一般的な背景除去は、色やテクスチャの特徴を利用しているが、シーンごとの照明の変化やノイズ画素の影響により、複雑な環境下ではうまく背景除去できないことが多い。この問題を解決するために、LSBP が提案されている。この手法の特徴は、特異値分解を用いている点である。与えられた画像の局所領域の潜在的な構造を記述することができ、照明変動、ノイズ、影に対するロバスト性を向上させている。



(a) Original image (b) Foreground regions obtained by LSBP
Fig. 3 Background removal

3. 検証実験

3.1 実験概要

提案手法の有効性の確認および pix2pix を用いた従来手法との比較を目的として、検証実験を行った。

監視カメラ映像は公開データセットである UCSD データ[3]を用いる。評価には、モデルの精度を評価する指標である ROC(Receiver Operating Characteristic)曲線が囲う面積の大きさである AUROC(Area Under ROC)を用いる。AUROC が 1 に近いほど高精度な異常検知ができていていることを表す。

まず、背景除去データを取得するために、UCSD データに OpenCV に実装されている LSBP を用いて処理を行った。背景除去したのはテストデータのみである。また、取得した背景除去データには、ノイズや動物体の抽出不足な点があることから、膨張・収縮処理を行う。具体的には、縦横方向に 15pixel 膨張し、円状に 3 pixel 収縮した。図 3(b)を膨張・収縮した結果を図 4 に示す。

続いて、入力に取得した背景除去データを用いて異常検知を行った。異常検知には、2.2 節で説明した pix2pix の手法を用いる。学習フェーズでは、既存手法と同様に学習を行う。識別フェーズでは、まず optical flow を予測するモデルからの推定結果と真値との差を求める。その後、背景除去データを掛け合わせることで、推定した optical flow の背景情報を除去する。そして得られたデータの異常度を計算して、異常かどうかを適切に検知できるか検証を行った。



Fig. 4 Image with erosion and dilation

3.2 実験結果

AUROC の値を表 1 に示す。提案手法では、異常な行動を既存手法より高精度に判別することができた。異常データの例を図 5 に示す。自転車や車が異常である。また、pix2pix を用いた optical flow の生成結果を図 6 に示す。上から順番に、入力データ、出力された生成データ、入力データと出力データの差となっている。

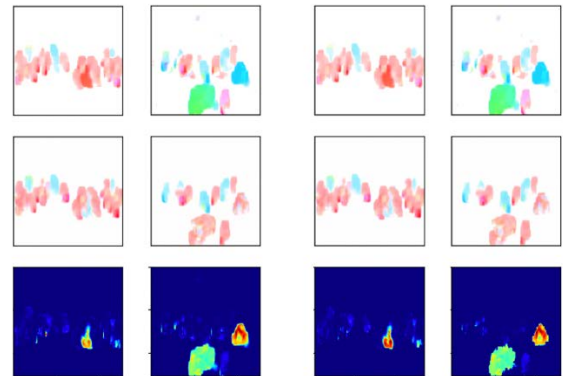
背景除去データを掛けたことにより、背景情報のノイズを除去し、異常箇所がより鮮明になっていることがわかる。

Table 1 Result of AUROC for UCSDped2

	Conventional method [2]	Proposed method
AUROC	0.935	0.953



Fig. 5 Example of abnormal data



(a) Conventional method (b) Proposed method

Fig. 6 Input, output, and difference results in optical flow

3.3 実験考察

3.2 節の結果より、提案手法は、高精度な動画異常検知手法として有効であると言える。異常を検知するフェーズにおいて背景除去データを活用することにより、従来の Optical Flow の差分のみを用いた手法と比較しても高い精度を出すことができた。

生成画像と真値画像との差に背景除去データを掛け合わせることで、差分画像での背景情報のノイズを削除することができ、結果として高精度な異常度の算出に繋がったと考える。

4. 結論

本研究では、背景情報を考慮しない異常検知システムを構築した。具体的なタスクとして、監視カメラ映像での異常検知精度の向上を目指し、従来の手法と比べ高い精度での異常検知を可能にした。

今後の展望として、背景除去手法の改善により、さらに高精度な動物体抽出を行うことや、他のデータセットでも高精度に異常検知を行う異常検知手法の検討がある。

参考文献

- [1] Venkatesh Saligrama, et al., "Video anomaly detection based on local statistical aggregates," CVPR, 2012.
- [2] Mahdyar Ravanbakhsh, et al., "Abnormal Event Detection in Videos using Generative Adversarial Nets," ICIAP, 2017.
- [3] Vijay Mahadevan, et al., "Anomaly detection in crowded scenes," CVPR, 2010.
- [4] Ian Goodfellow, et al., "Generative Adversarial Networks," NeurIPS, 2014.
- [5] Phillip Isola, et al., "Image-to-image translation with conditional adversarial networks," CVPR, 2017.
- [6] Alex Krizhevsky, et al., "ImageNet classification with deep convolutional neural networks," NeurIPS, 2012.
- [7] Lili Guo, et al., "Background Subtraction using Local SVD Binary Pattern," CVPR, 2016.