

動画画像中の異常検出ネットワークモデルの提案

菅沼 雅徳†

長尾 智晴†

† 横浜国立大学大学院環境情報学府

1 はじめに

監視カメラからの不審者や不審物の検出, 医療分野における病変部位の検出など異常検知システムは様々な分野で必要とされている. しかし, 与えられた環境によって異常の定義が異なることや動的環境では異常の特徴が変化するため, 計算機による異常検知システムの構築は容易ではない. 一方, 我々人間は任意の環境で高精度に異常検知が可能である. これは人間は与えられた環境を繰り返し観測することで環境の通常状態というものを自然に定義し, 異常を非通常状態と定義し注意を向けることができるためであると考えている.

そこで本研究ではこの考えに基づき, 入力画像を繰り返し観測することで環境の通常状態を記憶するネットワークモデルを提案する. 実験では車両や歩行者を検出すべき異常対象とする侵入物体検知問題に提案モデルを適用し, 提案モデルの有効性を検証した.

2 提案ネットワークモデル

2.1 提案ネットワークモデルの概要

環境の通常状態を記憶する提案ネットワークモデルの構造を図1に示す. 提案モデルは通常状態を記憶するネットワークが入力画像中に格子状に整列している. 例えば, 入力画像のサイズを $M \times N$ 画素, 1つの格子領域のサイズを $m \times n$ 画素とした場合, 入力画像中には $\frac{M}{m} \times \frac{N}{n}$ 個のネットワークが配置される. 各ネットワークには入力画像中の各領域から算出した特徴量が入力され, ネットワーク回路を用いて出力値が算出される. 各ネットワークの出力値を入力画像中の各領域の異常度として定義することで, 異常検出を行う.

2.2 各ネットワークの動作

各ネットワークは複数のノードとエッジによって構成されている. 各ノード i は入力特徴量と同次元の重みベクトル \mathbf{v}_i , 各エッジ e_{ij} は結合荷重 w_{ij} をもつ.

提案モデルによる異常検出は以下の手順で行う.

1. 全てのネットワークを以下の通りに初期化する.

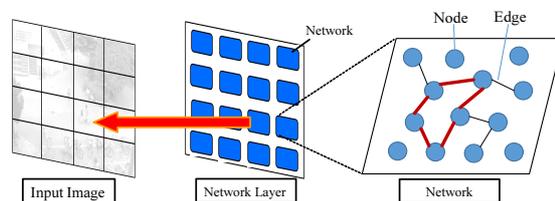


図1: 提案ネットワークモデルの構造

- ネットワークに I 個のノードを生成する.
 - 各ノードの重みベクトルを入力特徴量からランダムに生成する.
- 以下の処理を毎フレーム行う.
 - 各ネットワークが配置された格子領域で特徴量を算出し, ネットワークに入力する.
 - ネットワーク内のノードで入力特徴量と距離 D が小さい上位 S 個のノードを選択する. ノード i の距離計算には次式 (1) を用いる.

$$D_i = \exp(-\|\mathbf{f} - \mathbf{v}_i\|). \quad (1)$$

D_i はノード i の距離, \mathbf{f} は入力特徴量である.

- 距離 D が小さい順に2つのノード間にエッジを生成する. 既にエッジが存在している場合は次式 (2) によって結合荷重を更新する. α は定数, 結合荷重の初期値は 0.01 である.

$$\Delta w_{ij} = \alpha D_i D_j. \quad (2)$$

- 各ネットワークの出力値 O は選択された上位 S 個のノードとエッジの結合荷重を用いて, 次式 (3) によって算出される. β は定数である.

$$O = \sum_{i=0}^{S-1} \beta D_i D_{i+1} w_{i,i+1}. \quad (3)$$

出力値 O がしきい値 O_{th} より小さい場合, その領域を異常であると判定する.

- ネットワーク内で最小の距離 D_{min} がしきい値 D_{th} より大きい場合, 新たなノード k をネットワークに追加する. ノード k の重みベクトルは入力特徴量と最小距離 D_{min} のノードの重みベクトル \mathbf{v}_{min} との平均値を用いる.

A proposal of novelty detection network model for videos

†Masanori Suganuma †Tomoharu Nagao

†Graduate School of Environment and Information Sciences, Yokohama National University

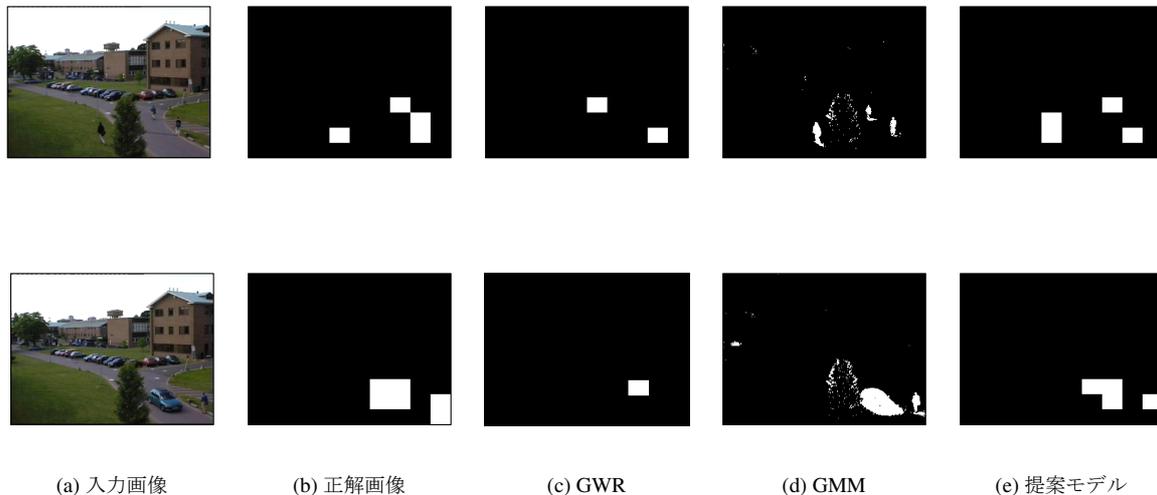


図 2: 検出対象に対する各手法の出力結果例

3 監視カメラからの侵入物体検知実験

3.1 実験設定

本稿では屋外に設置された監視カメラ映像を対象として、映像中に現れる車両と歩行者を検出すべき異常対象とした問題を扱う。対象とした監視カメラ映像として、PETS2001 データセットを用いた。用いた画像は 320×240 画素のカラー画像である。本稿では 32×24 画素領域に 1 つのネットワークを配置したため、ネットワーク数は 100 となった。提案モデルのパラメータは、 $I = 20$, $S = 5$, $\alpha = 0.8$, $\beta = 0.5$, $O_{th} = 3.0$, $D_{th} = 0.9$ を用いた。入力特徴量には平均、レンジなどの基本的な統計量を用いた。

提案モデルの性能を検証するため、正解画像を人手によって作成した。提案モデルが入力画像中の矩形領域ごとに出力をするため、正解画像は矩形領域単位で評価するものを用意した。矩形領域内に占める車両や歩行者といった検出対象画素の割合が 0.1 以上の矩形領域を正解矩形領域とした。正解画像の例を図 2(b) に示す。

実験では映像内の通常状態を提案モデルによって記憶させるために、車両と歩行者が現れないフレーム群を用いてモデルの構築を行った。そして、構築した提案モデルを車両や歩行者が現れるフレーム群に適用することで、性能検証を行った。提案モデルに対する比較手法として、Zivkovic らの混合ガウスモデルを用いた背景差分法 [1] (以下、GMM)、環境の通常状態を表現する先行研究である Grow When Required ネットワーク [2] (以下、GWR) を用いた。

3.2 実験結果と考察

提案モデル、比較手法による出力結果例を図 2 に示す。図 2 の結果から、提案モデルが歩行者や車両に対して適切に反応できていることがわかる。また、手前の樹木は風によって揺れており、比較手法はそのような領域に対して過検出してしまっているのに対して、提案モデルは過検出を抑制することができている。これは提案モデルでは背景の変化に対しても新たなノードの追加や結合荷重の更新によって、適応的に環境の通常状態を記憶することができているためだと考えられる。

4 まとめと今後の課題

本稿では、映像中の通常状態を記憶するネットワークモデルを提案した。提案モデルを侵入物体検知問題に適用した結果、先行研究と比べて高精度な結果を示した。今後の課題は多くの環境への適用実験によってモデルの検証を行い、精度向上のためモデルを改良することが挙げられる。

参考文献

- [1] Zivkovic Zoran. Improved adaptive gaussian mixture model for background subtraction. *International Conference Pattern Recognition*, Vol. 2, pp. 28–31, 2004.
- [2] Marsland Stephen, Shapiro Jonathan, and Nehmzow Ulrich. A self-organising network that grows when required. *Neural Networks*, Vol. 15, No. 8, pp. 1041–1058, 2002.