

# 多様な街路における自動車走行中の異常検知に関する研究

## -リカレント型ネットワークによる未来予測を用いて-

### Research on abnormality detection while driving on various streets

#### -Using future prediction by recurrent network-

○中野 雄大\*<sup>1</sup>, 山田 悟史\*<sup>2</sup>  
Yudai NAKANO\*<sup>1</sup>, Satoshi YAMADA\*<sup>2</sup>

\*1 立命館大学大学院 理工学研究科

Graduate Student, Graduate School of Science and Engineering, Ritsumeikan University.

\*2 立命館大学 理工学部 准教授 博士(工学)

Associate professor, Department of Science and Engineering, Ritsumeikan University, Eng.Dr.

キーワード：建築；情報；深層学習；リカレント型ネットワーク；異常検知

Keywords: Architecture; information; Deep Learning; recurrent network; Anomaly detection.

## 1. はじめに

近年、我が国に限らず、モータリゼーションが進行し、車社会とも呼ばれるまでに我々人々と車は密接にかかわってきている。しかし、車文化の普及とともに車を使用する年齢層も広がり、高齢ドライバーによる接触事故や、自動車運転者による故意的な危険運転や近年特に注目されている煽り運転、その他道路交通法を無視した無茶な運転による事故が多発しており、そういった事故による人々の安全性の侵害が近年問題となっている。

そういった事故を防止するべく、ドライブレコーダーとAIを用いた異常検知による様々なオートブレーキシステム（以下、従来の異常検知と表記）が開発され、実用化に向けて進んでいる。この従来の異常検知は、ドライブレコーダーに映るものをラベリングし、ラベリングされた物の急な飛び出しや、急接近などによって異常発生と認識するシステムになっている。この従来の異常検知は、ラベリングされている物に対しては高い確率で異常検知できる一方で、図1のように落下物やかかんでいる子供の後ろ姿などの想定が困難な障害物にはラベリング処理が弱いことが課題となっている。

この課題を解決する手法として、リカレント型ネットワークである ConvLSTM（Convolution Long Short-Term Memory）<sup>1)</sup>とヒストグラム比較の2つを用いた従来の異常検知に代わり得る、ラベリングを要せず異常を検知できる新たなモデルの構築が考えられる。ConvLSTMを用いて過去の記憶から何事も無ければ来るはずの未来予測画像を生成し、ヒストグラム比較を用いて画像間の類似度を即座に測り、異常を検知するシステムであり、この深層学習によるシステムから直接事故を回避させるように試みる。

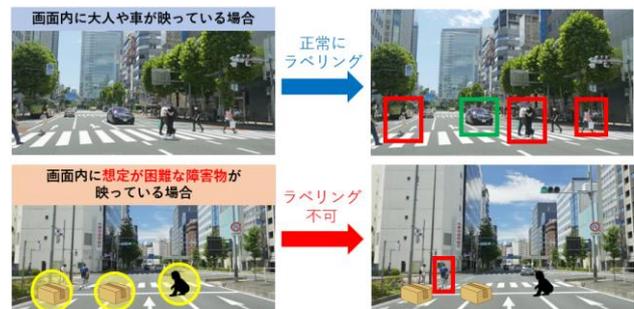


図1. ラベリング処理

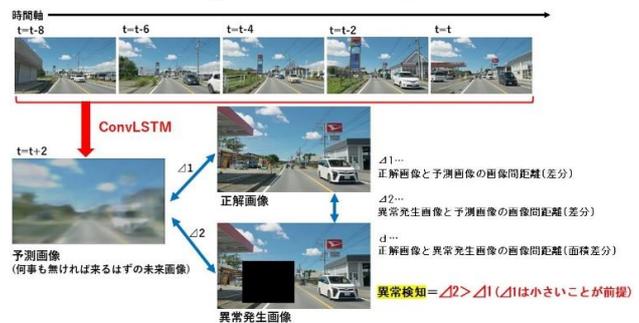


図2. 研究の流れ

## 2. 概要

### 2.1. 提案する異常検知システムの流れ

システムでは車載カメラにより2秒間隔で生成された5枚の連続画像から ConvLSTM を用いて2秒先の6枚目の何事も無ければ来るはずの未来画像（予測画像）を生成する。この予測画像と実際の2秒後の画像（正解画像）との差分を $\Delta 1$ とする。異常が発生した場合の画像（異常発生画像）との差分を $\Delta 2$ とする。 $\Delta 2 > \Delta 1$ となることは異常が無ければ訪れていた未来とは異なる状況になったことを意味している。本研究ではこの状況を異常発生と定義する。但し本研究では異常発生画像は正解画像の一部を任意

面積 (d) の正方形をランダムな単色で重ね描きした画像とする。d を段階的に小さくした際の正しい異常検知率をシステムの検証値とし、これらを様々な街路で検証する。また、検知率の高い  $\angle$  の計測方法の設定も研究課題である。

## 2.2. 街路の種類

本研究で用いる車載動画の街路の種類は、明るい街路と暗い街路、狭い街路と広い街路、景色の移り変わりが激しい街路と景色の移り変わりが緩い街路の2択3種類に分類した。計8通りの街路をそれぞれ5つ車載動画<sup>1</sup>を用意し、素材による偏りが出ないようにする。

## 2.3. dataset の作成

訓練・検証用のデータセットの元となる画像は、既述した5つの車載動画のうち、1つの車載動画をフレームレート毎に静止画を切り出しして生成した。フレームレートは0.5fpsとし、2327枚の連番画像に分割したものを用意した。この画像群を訓練:検証に8:2の比率で先頭から分割した。また、後述の学習モデルに合わせて画像に zero padding と resize を施し、(108×192×3) に形状を変更した。また、残りの4つの車載動画も同様に、同じフレームレート毎に静止画を切り出し、それら画像群をテストデータセットとした。

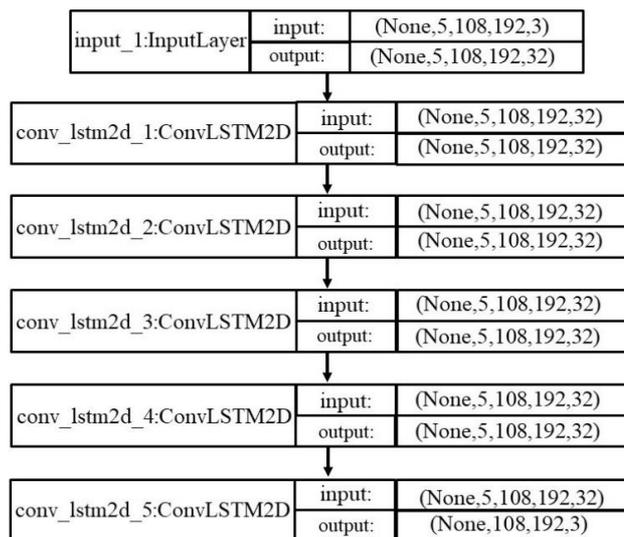
## 2.4. ConvLSTM の構造

ConvLSTM の層構成は5層構成とし、街路の種類ごとにそれぞれ8種類用意する。input の引数は、(サンプル数、フレーム数、画像の高さ、幅、RGB) となっており、最終的な output の引数は、(サンプル数、画像の高さ、幅、RGB) となっている。ランダムに選択された5フレームの連番画像を入力し、2秒後の予測画像を1枚生成するモデルとなっている。今回使用したモデルでは各畳み込み層の間で batch normalization を行っている。これを François Chollet の運営するオープンソースの機械学習フレームワークである keras を利用して、Google 社が提供する Google Colaboratory 上で実行した。

学習時のハイパーパラメータは表1に示す。この各項目については以下で説明する。batch size は1step 内で並列処理するデータ数を指し、本研究では10枚の画像を同時に計算させた。損失関数は正解画像とAIによる予想画像の不一致度合いを示す loss を計算する関数である。評価関数は学習させたモデルの性能の良さを示す関数である。最適化関数は loss が最小になるように学習時の重みを更新する関数である。また、momentum と kernel size に関しては、自身の過去の研究において最も ConvLSTM の精度が高くなった値で実行した。<sup>2)</sup>

## 2.5. 画像類似度の評価方法

画像間の RGB のヒストグラムの類似度とする。画像内の色の分布の様子を、X 軸を RGB ( $0 \leq X \leq 255$ )、Y 軸を分布度合いで示すヒストグラムで表現し、画像間でそのグラフを比較して類似度を測る。ヒストグラム作成には



入力(サンプル数、フレーム数、高さ、幅、RGB)  
最終的な出力(サンプル数、高さ、幅、RGB)

図3. ConvLSTM の構造

表1. 学習時のハイパーパラメータ

項目	値
batch size	10
損失関数	平均絶対誤差(mae)
評価関数	平均二乗誤差(mse)
最適化関数	Adam
学習率	0.001
epoch数	10
momentum	7*7、5*5、3*3、1*1
kernel size	0.4、0.6、0.8、0.99

表2. 各方法の類似度表現

ピアソンの相関係数	-1から1の範囲で類似度を表す 1に近い方が類似度が高い
カイ二乗検定	0以上の値で類似度を表す 0に近い方が類似度が高い
交差法	0以上の値で類似度を表す 0から乖離してる方が類似度が高い
バタチャリア距離	0から1の範囲で類似度を表す 0に近い方が類似度が高い



図4. 異常発生画像の一例

OpenCV<sup>3)</sup>の calcHist 関数を用いる。また、表2に示したようにヒストグラム比較には4種の方法を検討する。

## 2.6. 異常発生素材の作成

前述の方法で作成し、d の割合の水準は40%、35%、30%、25%、20%、17.5%、15%、12.5%、10%、7.5%、5%、2.5% の12段階とする。

### 3. 検証結果

#### 3.1. ConvLSTM による予測画像評価

予測画像と正解画像の比較9セットを図5に示す。上段が正解 (cor) 画像で、下段が予測 (pre) 画像である。形状のエッジが不明瞭ではあるが、街路の種類に関わらず、車や人、建物の配置、道路の中央線や横断歩道の位置はほぼ正確であることが読み取れる。また、直進だけでなく曲がり角についても同等の正確さが読み取れる。また、8種類のConvLSTMにおけるloss値は0.10前後であり、街路の種類による精度のばらつきがないことが確認できた。

#### 3.2. 類似度結果

正解画像と予測画像、異常発生画像（面積dの正方形をランダムな単色で重ね描きした画像）の3枚から構成されるある1セットのRGBのヒストグラムを図6に例示する。グラフはR、G、Bの3種類を重ねてプロットしたものである。また、図6で示した画像を用いて $\angle 1$ と $\angle 2$ の大小関係を画像類似度評価の方法で既述した4種の方法で測った結果を表4に示す。青色で塗られている箇所が $\angle 2 > \angle 1$ が成立しており、正しく異常を検知できたことを示す。反対に橙色で塗られた部分は、異常検知として成功していない部分である。

#### 3.3. 検知率の検証結果

dは異常発生画像の作成で既述したように12水準、テストデータの232セット（1セット正解画像、予測画像、異常発生画像の3枚の画像群）、異常発生パターン数300（単色と位置が一様分布でランダムに生成）、計2,505,600枚（835,200セット）の画像を用いて検知数を算出し、各dの割合毎にテストデータのセット数で割って、8通り4つの街路の計32個の平均検知率を算出する。8通りの街路それぞれにおける検知率の平均値をとってまとめた結果を図7に示す。街路の明るさを明・暗、街路幅を狭・広、景色の移り変わりの度合いを激・緩で分類し、グラフの下にその組み合わせを明記している。表4と同様に、青色のグラフが検知成功、橙色のグラフが検知失敗である。

#### 3.4. 考察

暗い街路は全体的に明るい街路と比較して検知率が高い値を示したが、これはヒストグラム比較という画像内の色の分布の様子を比較する手法の為、視覚的情報が少ない暗い街路の方が検知しやすいことが考えられる。また、全体を通してピアソンの相関係数による検知率が高かったのは、この手法は共分散や標準偏差を用いて類似度を測るため、視覚的類似度が高く読み取れた正解画像のヒストグラムの偏差よりも、ヒストグラムが一部乖離していた異常発生画像の方が予測画像との偏差が大きくなり、正解画像の方が予測画像との相関関係が強くなったためだと考えられる。また、カイ二乗検定は、ヒストグラムの差を二乗するため、グラフの乖離が大きい異常発生画像よりもグラフの一致率が高い正解画像の方が差分が小さく、検知率が



図5. 予測画像 (pre) と正解画像 (cor) の比較

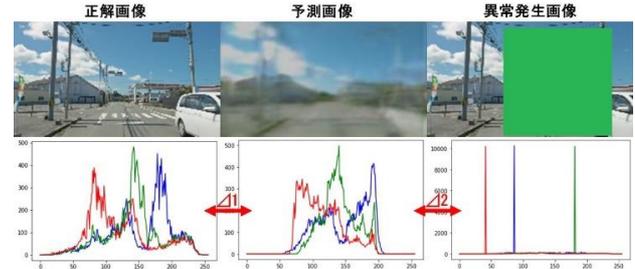


図6. ヒストグラム比較 (X: RGB, Y: 分布度合い)

表4. 類似度結果

比較手法	予測-正解 ( $\angle 1$ に因与)	予測-異常 ( $\angle 2$ に因与)
ピアソンの相関係数	0.755	0.422
カイ二乗検定	15190.357	6719156.668
KLダイバージェンス	73038.897	3261.883
交差法	14774.0	18049.0
パタチャリア距離	0.332	0.255

高くなったと考えられる。一方で街路に関係なく検知率が低かった交差法は、min演算を適用しているため、グラフの乖離が類似度に影響せず、検知率が低くなったと考えられる。

#### 4. まとめ

8通りの街路における検知率の違いについて、明るい街路については、街路幅の広さや景色の移り変わりの激しさはほぼ関係しないことが確認できた。同様に暗い街路における検知率も道幅の広さや景色の移り変わりの激しさはほぼ関係しないことが確認できたが、全体的に検知率が明るい街路に比べて高くなった。

検知率に関しては、本研究で行った条件では、dの割合が40%以下であれば手法に関わらずほぼ確実に異常を検

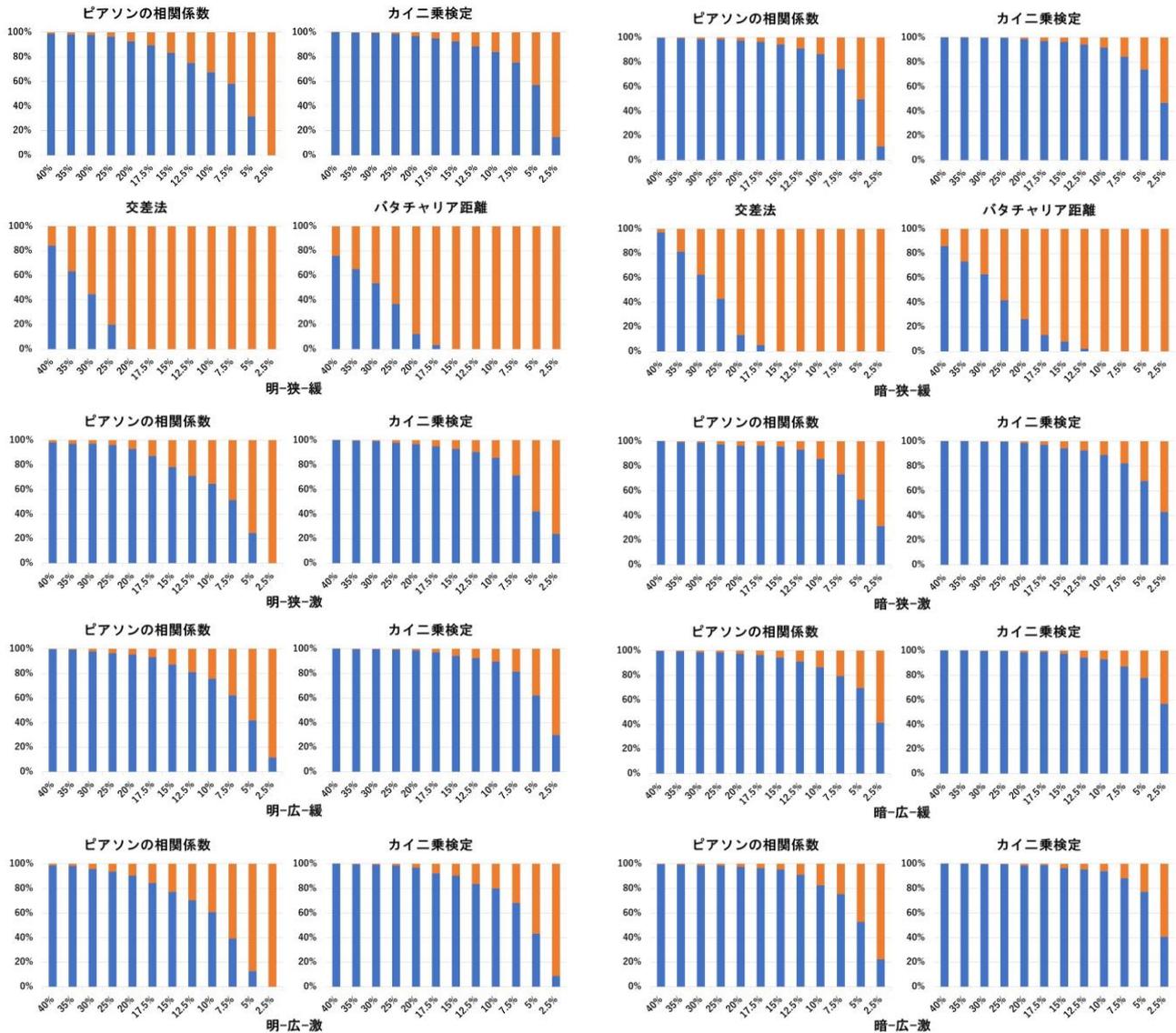


図 7. 検知率結果 (X : d の割合, Y : 検知率)<sup>2</sup>

知ることができることを把握した。また、7.5%付近までであればピアソンの相関係数とカイ二乗検定は過半数以上の検知率を示すが、他の2つの手法については検知率が低くなり、ピアソンの相関係数とカイ二乗検定を使用する異常検知が有効的であることを把握した。

検知時間に関しては、予測画像を出力するのに要する時間は GPU 環境下で 0.97482 秒、検知率を測るのに要する時間は CPU 環境下で 0.00478 秒で計 0.9796 秒であり、リアルタイムシステムに落とし込む場合、2 秒先の異常を検知するには約 1 秒の余裕があることも把握した。

今後の方針としては、今回の様に異常発生箇所を単色の正方形で限定していたため色や形状のパターン数を増加させることや、色の分布に依存するヒストグラム比較以外の類似度評価を用いて検知調査を行うことが挙げられる。また、今回は突発的な異常発生について研究を行ったが、対向車が徐々に中央線を越えてくるといった突発的でない異常に対する検知の検証も挙げられる。

【参考文献】

- 1) Xingjian Shi, Zhouong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, Wang-chun Woo, Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, arXiv:1506.04214v2, 2015
- 2) 日本建築学会 学術講演会梗概集 (情報システム技術), pp. 127-128, 2023.9, 日本建築学会
- 3) OpenCV, [https://opencv.org/], (最終参照日 2023 年 9 月 25 日)

【注釈】

1. 今回使用した車載動画は以下を参考にした  
[https://www.youtube.com/@DashcamRoadshow]  
この中から 8 通りの街路の分類にそれぞれ適当な車載動画を計 32 個選出した
2. 交差法とバタチャリア距離に関しては街路の種類に関わらず検知率が低かった為、明-狭-緩、暗-狭-緩の2つ以外は図から削除した