

〈研究ノート〉

## ConvLSTM によるスペースデブリと人工衛星の識別

長尾光悦\* 千葉慎一郎† 大井渚‡

### Differentiate between Space Debris and Satellites using ConvLSTM

Mitsuyoshi NAGAO\* Shinichiro CHIBA† Nagisa OI‡

#### 要旨

スペースデブリ（以下、デブリ）とは、使用済みロケット部品や故障した人工衛星などである。地球を周回する一定以上の大きさのデブリは、地上からの観測によって情報がカタログに登録されているが、現在、登録されていないデブリも数多く存在しており、衝突事故などの危険性が増加している。本論文では、ConvLSTM を用いたスペースデブリと人工衛星の識別を提案する。本研究では、トモエゴゼンと呼ばれる観測装置によって撮影された動画像から機械学習の一つのモデルである ConvLSTM に基づき識別を行う。また、ConvLSTM との性能比較のために VGG16 を用いた識別も実施する。実験結果から、人工衛星とデブリが ConvLSTM と動画像から高精度で識別可能であることを確認した。

#### Abstract

Space debris (hereafter, debris) consists mainly of used rocket parts and defunct satellites. Debris of a certain minimal size orbiting the Earth is cataloged based on observations from the ground. However, there are still many pieces of debris that have not been detected and catalogued, increasing the risk of collisions. This paper proposes the use of ConvLSTM for identifying debris. In this paper, we outline the identification of debris based on videos captured by a space observation device called Tomo-e Gozen, using a machine learning model known as ConvLSTM. Additionally, identification of debris using VGG16 is conducted for performance comparison with ConvLSTM. From the experimental results, we confirmed that satellites and debris can be identified with high accuracy from ConvLSTM and videos.

#### キーワード

人工衛星 (Satellites)    スペースデブリ (Space debris)    ConvLSTM

\* 北海道情報大学経営情報学部システム情報学科教授, Professor, Department of Systems and Informatics (Dept. of BIS.), HIU

† 北海道情報大学経営情報学部システム情報学科, Department of Systems and Informatics (Dept. of BIS.), HIU

‡ 北海道情報大学経営情報学部システム情報学科講師, Lecturer, Department of Systems and Informatics (Dept. of BIS.), HIU

## 1. はじめに

スペースデブリ（以下、デブリ）とは、使用済みロケット部品、故障した人工衛星、ミッション遂行上宇宙機から放出された物体、爆発や衝突により発生した破片などを指す。デブリは2007年頃から急速に増加しており、この原因としては、2007年に中国が行った衛星破壊実験、2009年のアメリカの通信衛星イリジウムと使用済みロシア衛星コスモス 2251との衝突事故、2021年にロシアが行った衛星破壊実験が挙げられる（H. Cowardin 2022）。

地球を周回する一定以上の大きさのデブリは、地上からの観測によって、その情報がカタログに登録され、継続的に追跡が行われている。しかしながら、現在、カタログ化されていないデブリも多数存在している。デブリが増加傾向にある現在、デブリ同士や運用中の衛星との衝突が発生する懸念が増加しており、そのような事態になれば更に多量のデブリが発生する。そのため、デブリの数や運動等を把握することは重要である。

一方、天体観測において、トモエゴゼンと呼ばれる口径 105cm のシュミット望遠鏡に取り付けられた動画像の撮像装置が存在する（S. Sako et al. 2018）。これは、東京大学木曾観測所が 2019 年 10 月に運用を開始した装置である。トモエゴゼンは、84 台の CMOS イメージセンサを用いることにより、広視野の撮影能力を有しており、撮影されたデータは、天文学で広く使用される画像ファイル形式である fits (flexible image transport system) ファイルとして一般に提供されている。このトモエゴゼンにより撮影される動画像データからデブリと衛星を分類するこ

とが可能となれば、デブリの管理上、有効性は高い。しかしながら、トモエゴゼンにより取得される人工衛星とデブリの動画像は非常に類似しており人間がこれを判別することは困難である。

本論文では、ConvLSTM によるスペースデブリと人工衛星の識別を提案する。具体的には、トモエゴゼンによって撮影された動画像から動画像の予測に適した ConvLSTM に基づき識別を行う。また、VGG16 を用いた静止画からの識別、VGG16 を用いた動画像からの識別も実施し、識別性能の比較を行う。

## 2. 深層学習

### 2-1 LSTM

LSTM (Long Short-Term Memory) は、RNN (Recurrent Neural Network) の一種であり、長期依存関係を学習できるように設計された機械学習モデルの一つである（S. Hochreiter et al. 1997）。RNN では、時系列データの処理において勾配消失問題が発生しやすい。この問題を解決するために考案されたのが LSTM である。LSTM は、過去の状態から現在の状態に不要な情報を取り除くための忘却ゲート、現在の入力データと前の時刻の状態をもとに新たに追加する情報を決定する入力ゲート、現在の状態と前の時刻の状態から次の時刻に伝える情報を決定する出力ゲートを導入することによって、重要な情報を保持し、不要な情報を捨てることが可能となっている。LSTM は、時系列データや連続データの依存関係を効果的に学習できるため、自然言語処理、時系列予測、音声認識、動画解析などに利用されている。

## 2-2 ConvLSTM

ConvLSTM (Convolutional LSTM) とは、CNN と LSTM を組み合わせた機械学習モデルである (Xingjian Shi et al. 2015)。LSTM は、時系列データやシーケンスデータに対して有効であるが、画像などの位置的な情報が保持されない。ConvLSTM は、時系列データと空間的な情報を同時に考慮するために設計されたモデルであり、画像の位置的な情報を保持したまま、その時間的な変化を扱うことが可能である。Xingjian Shi らは、気象レーダー画像の時系列データから、短期間における特定地域の将来の降雨強度の予測を実施している。その結果、ConvLSTM が、時空間の相関を効果的に捉え、優れた予測性能を発揮することが可能であることを報告している。

本研究では、トモエゴゼンにより取得された動画像に基づく識別を行うため、ConvLSTM を採用する。

## 2-3 VGG16

本研究では、ConvLSTM との識別性能比較のために VGG16 を用いる。VGG16 は、2014 年にオックスフォード大学の Visual Geometry Group によって発表された CNN (Convolutional Neural Network) モデルの一つである (A. Krizhevsky et al. 2012) (Karen Simonyan et al. 2015)。CNN は、画像処理に使用される深層学習の一つであり、畳み込み層とプーリング層を 1 セットとし、これを複数回連続させたネットワーク構造を持つ。VGG16 は、この畳み込み層のカーネルサイズを  $3 \times 3$  に統一し、畳み込み層を 13 層持つモデルであり、画像分類やオブジェクト検出で高い性能を示している。

## 3. ConvLSTM によるデブリと人工衛星の識別

### 3-1 トモエゴゼンによって取得したデータ

トモエゴゼンは 84 台の CMOS イメージセンサで構成され、合計 1 億 9000 万画素の広視野の画像を毎秒 2 フレームで 9 秒間撮影し、18 フレームの動画を撮影可能な天体観測装置である (図 1)。広範囲の空の網羅的観測を繰り返し、一晩の観測で得られるデータは 30 テラバイトに及ぶ。この観測データにおいて、移動している天体や明滅している天体などを  $100 \times 100$  ピクセルの画像として抽出している。

本研究において用いたデータは、トモエゴゼンによって 2023 年 4 月 9 日に撮影された fits データである。また、トモエゴゼンの fits データに格納された人工物の位置情報と米国で運用される、地球を周回する人工物を識別・監視しているサービスである SpaceTrack の位置情報を比較することで人工衛星かデブリかのラベル付けを行う。この動画を 18 枚の静止画に分解し、利用する。実験に用いたデータ例を図 2 に示す。図 2 において、上段左がデブリの動画例、上段右が人工衛星の動画例を示す。図では、光の大きさが衛星よりデブリの方が小さくなっているが、移動体の角度により光の大きさは変化するため、移動体の大きさと光の大きさが一致するものではない。また、下段が動画像を 18 枚に分解した画像例である。

### 3-2 モデル構成

本モデルの構成を図 3 に示す。図 3 に示すように、18 個の入力、VGG16、パッ

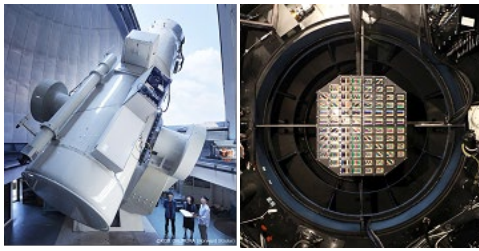


図1 トモエゴゼン

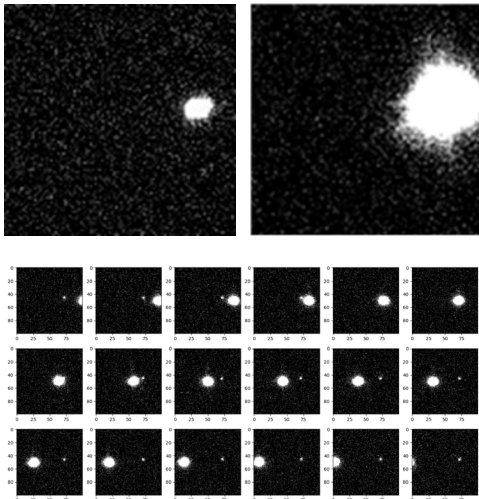


図2 データ例

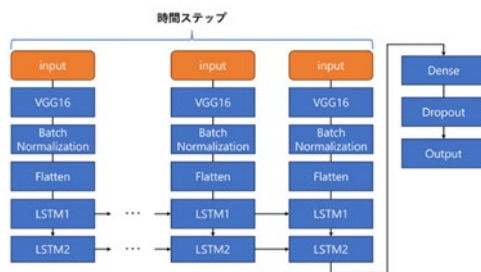


図3 本 ConvLSTM モデルの概要

チ正規化, Flatten 層, 2 層の LSTM 層, Dense 層, Dropout 層, 出力層から構成される。各時間ステップで同じ処理を行えるよう VGG16, バッチ正規化, Flatten 層に対して TimeDistributed レイヤーを適用し, LSTM へ入力することとした。本モデルでは, VGG16 の全層の重みは固定, 中間層 3 層, LSTM 2 層, 出力層を学習させる。中間層のうち 1 層は過学習の抑制のため Dropout 層を導入する。入力として

は, 静止画 18 枚から構成される動画像である。

### 3-3 性能比較のためのモデル

ConvLSTM との性能比較のため, VGG16 を用いた静止画像 1 枚からの識別, 及び, VGG16 を用いた動画像からの識別を行う。静止画像 1 枚で識別するモデルを図 4 に示す。図 4 に示すように, 1 つの入力, VGG16, バッチ正規化, Flatten 層, Dense 層, Dropout 層, 2 つの Dense 層, Dropout 層, 出力層から構成される。ここでは, VGG16 の全層の重みを固定し, 中間層 5 層と出力層を学習させる。中間層のうち 1 層は過学習の抑制のため Dropout 層を導入する。

また, VGG16 を用いた動画像で識別するモデルを図 5 に示す。図 5 に示すように, 18 個の入力, VGG16, バッチ正規化, Flatten 層, Dense 層, Dropout 層, 出力層で構成される。各時間ステップで同じ処理を行えるよう VGG16, バッチ正規化, Flatten 層に対して TimeDistributed レイヤーを適用し, Flatten 層により平滑化した後, 中間層に入力する。ここでは, VGG16 の重みを固定し, 中間層 3 層, 出力層を学習させる。中間層のうち 1 層は過学習の抑制のため Dropout 層を導入する。このモデルは, ConvLSTM における LSTM の有無についての識別精度を比較するために, 構造を ConvLSTM と同様にし, LSTM の代わりに Flatten 層でデータの次元数を削減している。

## 4. 識別精度実証実験

### 4-1 実験方法

実験において使用したデータ数を表 1

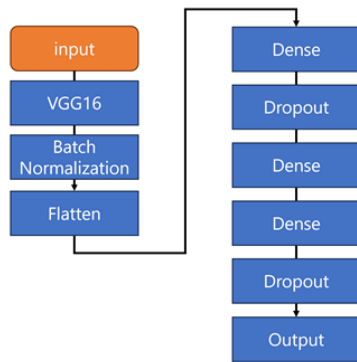


図4 VGG16を用いた静止画像の識別モデル

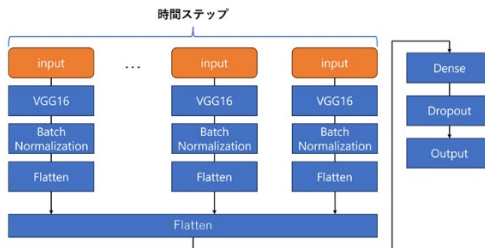


図5 VGG16を用いた動画の識別モデル

表1 使用データ数

	動画 (ConvLSTM)	静止画 (VGG16)
デブリ	294 個	3,828 枚
衛星	308 個	3,948 枚

に示す。生成したデータの8割を学習用、2割を検証用として使用した。実験では、学習率を0.0005、エポック数を200とし、入力画像はグレースケール画像をカラー画像に変換したものを利用する。また、トモエゴゼンにより取得される動画は100×100ピクセルの動画であるが、139×139ピクセルのカラー画像に変換したものを使用する。これは、本研究で利用するVGG16がカラー画像を想定して設計されているため、変換を行っている。また、事前実験により100×100ピクセルの場合には、特徴抽出が適切に行われず、精度が低くなることを確認したため画像サイズの変更を行っている。更に、比較のためのVGG16による識別では、1枚の静止画、及び、動画を利用し

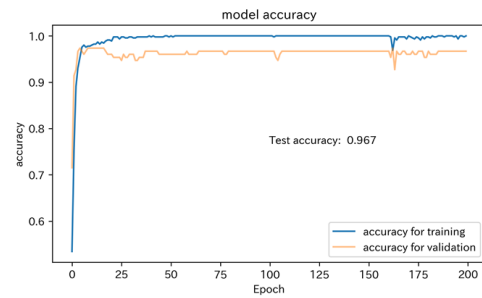


図6 識別精度 (ConvLSTM-動画)

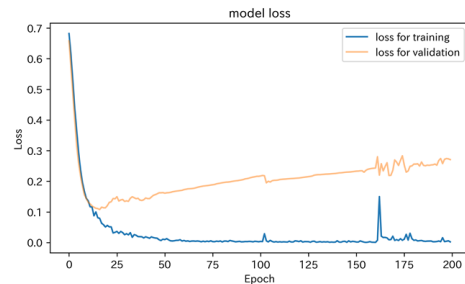


図7 学習時の損失 (ConvLSTM-動画)

た識別を行う。3つのモデル全てにおいてDropout率は0.25とする。

#### 4-2 ConvLSTMを用いた動画からの識別

図6と図7にConvLSTMを用いた動画からの識別精度と学習時の損失を示す。図に示すように、識別精度は、訓練データでは99%に達していたが、検証データでは約96.7%となった。また、エポック数の増加に伴い、訓練データと検証データの識別精度が増加しているが、図7に示すように検証データの損失が増加しており、過学習が発生していることが示された。これは、学習用データ数が少ないことで、過学習が発生し汎化性能が低下した可能性が考えられる。

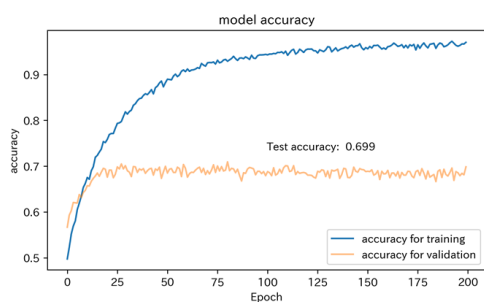


図8 識別精度 (VGG16-静止画)

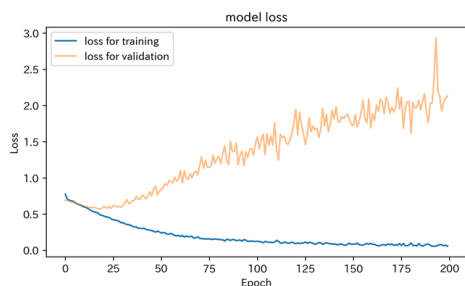


図9 学習時の損失 (VGG16-静止画)

#### 4-3 VGG16 を用いた静止画からの識別

図8と図9にVGG16を用いた静止画からの識別精度と学習時の損失を示す。図に示すように、識別精度は、訓練データでは99%に達したが、検証データでは約69.9%であった。エポック数の増加に伴い、訓練データの識別精度が増加しているが、検証データにおいては20エポック程度で頭打ちとなっている。損失はDropoutを使用したものの訓練データの損失は減少しているが、検証データの損失は増加しており、学習過程において過学習が起きていることが示された。

#### 4-4 VGG16 を用いた動画からの識別

図10と図11にVGG16を用いた動画からの識別精度と学習時の損失を示す。図に示すように、識別精度は、訓練データでは99%、検証データの識別精度は約97.4%と高い値を示した。一方、図11か

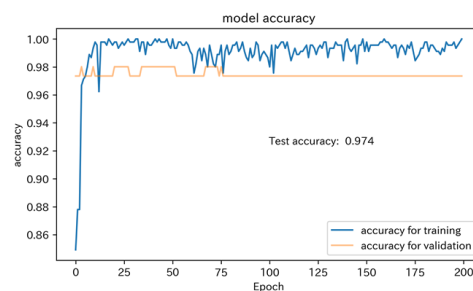


図10 識別精度 (VGG16-動画)

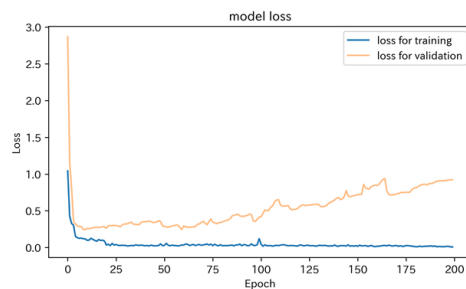


図11 学習時の損失 (VGG16-動画)

ら、検証データにおいて学習後半で損失が増加し、過学習により汎化性能が低下していることが示された。

一方、ConvLSTMによる動画からの識別精度よりVGG16による動画からの識別精度が若干高い精度となった原因としては、移動天体の画像は白と黒の動画であり、その単純さから、VGG16が既に識別性能を有しており、一方で、時間及び空間の両要素を考慮するConvLSTMは、そのモデルの複雑さから識別精度を低下させる方向に働いたと推察される。これは、図10における検証データでの精度が学習初期から高い値を示していることから推察される。このため、入力情報を追加し、データの複雑さを増加させた場合の識別実験を追加で実施した。

#### 4-5 入力情報の追加による識別精度検証

動画からVGG16、及び、ConvLSTMに基づき識別を行った場合、VGG16を用いた場合の方が識別精度が若干高い値と

表2 識別精度の変化

	画像のみ	画像+P
VGG16+S	69.9%	69.9%
VGG16+M	97.4%	96.7%
ConvLSTM+M	96.7%	96.0%
	画像+V	画像+P+V
VGG16+S	69.4%	72.4%
VGG16+M	98.0%	98.0%
ConvLSTM+M	96.7%	98.7%

S : 静止画, M : 動画, P : 位置情報, V : 移動速度

なった。これは、入力データの単純さから生じたものと推察されるため、入力情報を追加し、データの複雑さを増加させた場合の識別精度の変化を調べる実験を行った。実験では、移動天体の移動速度情報と位置情報を付加情報として追加した場合の識別精度を比較する。位置情報と移動速度情報はトモエゴゼンにより取得されている fits データから抽出した。

また、実験では、画像または動画と位置情報の2つを入力、画像または動画と移動天体の速度情報の2つを入力、画像または動画と位置情報、移動天体の速度情報の3つを入力とする実験をそれぞれ行う。

実験結果を表2に示す。表2に示すように、位置情報を追加することにより、VGG16と静止画の場合には、識別精度に変化はなかったが、VGG16と動画の場合とConvLSTMと動画の場合、入力が動画のみの時と比較して、識別精度の低下が確認された。これは、VGG16とConvLSTMの動画像分類において位置情報を追加することでデータが複雑化し、位置情報との関連性を学習しきれなかった可能性が考えられる。

また、速度情報を追加した場合、VGG16と静止画の場合、画像のみの場合

と比較し、識別率が0.5%低下した。これは、画像と速度の関係性を学習できなかったためであると推察される。VGG16と動画の場合には、動画のみの場合と比較し、識別率が0.6%上昇した。これは速度情報から動画像の変動を補完できたため上昇した可能性が考えられる。一方、ConvLSTMと動画の場合には、動画のみの場合と比較し、識別率に変化はなかった。これはConvLSTMの場合には、速度情報が画像から既に学習されていたため変化がなかった可能性が考えられる。

更に、位置情報と速度情報を追加した場合、VGG16と静止画の場合、画像のみの時と比較して2.5%の識別精度の向上が確認された。位置情報と速度情報を追加したことにより、静止画にはない空間的な位置関係や時間的な変化に関する情報を考慮することができたためであると考えられる。VGG16と動画の場合には、動画と速度情報を追加した場合と比較し識別率に変化はなかった。これは動画に位置情報を追加した結果からも示されるように、VGG16においては、位置情報との関係をうまく学習できず、速度情報のみが有効に働いたためと考えられる。一方、ConvLSMTの場合には、動画のみの場合と比較し2%の識別率の向上が確認された。全ての情報を入力することで、デブリと人工衛星の運動パターンなどをより正確にとらえることができた可能性が考えられる。これは、ConvLSTMの方がVGG16を用いた識別よりも時空間的な情報を有効に活用できるため、追加情報の影響がより顕著に表れたためであると考えられる。一方で、VGG16では、動画像内の変動を一定程度考慮するが、ConvLSTMに比べて変動に対する柔軟性が制限されている可能性が考えられる。

## 5. おわりに

本論文では、ConvLSTM とトモエゴゼンの観測データを用いて、デブリと人工衛星の識別を提案し、更に、VGG16 を用いた静止画像からの識別、VGG16 を用いた動画からの識別の識別精度との比較を行った。その結果、人工衛星とスペースデブリの識別が動画から可能であることを確認した。また、追加実験により位置と移動速度の2つの追加することによって ConvLSTM を用いた場合に識別精度が上昇することを確認した。

今後の課題としては、学習過程における過学習の問題があげられる。これは、学習データの水増しやモデルの改良を行うことで改善可能であると考えられる。また、ConvLSTM がどのようにデブリと人工衛星の識別を行っているのかがブラックボックスであるため、識別の可視化を行い、識別の要因を検証する必要がある。

## 参考文献

- A. Krizhevsky, I. Sutskever, et al. (2012)  
ImageNet classification with deep convolutional neural networks,  
NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, pp.1097–1105.
- H. Cowardin (2022) The Orbital Debris Quarterly News, A publication of the NASA Orbital Debris Program Office (ODPO), Vol.26, Issue 1, p10.
- Karen Simonyan, Andrew Zisserman (2015)  
Very Deep Convolutional Networks for Large-Scale Image Recognition,  
International Conference on Learning

Representations, p.14.

- S. Hochreiter, J. Schmidhuber (1997) Long short-term memory., *Neural Computation*, 9 (8): pp.1735-1780.
- S. Sako, R. Ohsawa, et al. (2018), The Tomo-e Gozen wide field CMOS camera for the Kiso Schmidt telescope, *Proceedings of the SPIE, Volume 10702*, id. 107020J, p17.
- Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, Wang-chun Woo (2015) Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, *Advances in Neural Information Processing Systems*, pp.802-810.