

# ノイズに埋もれた微小信号波形の検出への 変分オートエンコーダを用いた異常検知の応用

Toward Detection of Small Signals from Noisy Data  
Using Variational Autoencoder Based Anomaly Detection

森 雅也<sup>†</sup>  
Masaya Mori

中平 勝子<sup>†</sup>  
Katsuko T. Nakahira

高橋 弘毅<sup>†</sup>  
Hirotaka Takahashi

田中 貴浩<sup>‡</sup>  
Takahiro Tanaka

## 1 はじめに

近年、自然界に存在する微小信号波形の解析による自然災害の予測や、自然現象の解明を試みる研究が数多く行われている [1] [2] [3]. 一般に、微小信号波形の解析を行うためには、観測データから微小信号波形を検出をしなければならない。このため、観測データの中に解析対象である微小信号が含まれているか否かを判定する必要がある。しかし、観測データの中には大小さまざまな定常ノイズや非定常ノイズが含まれており、解析対象である微小信号の正確な波形予測が困難な場合、非定常ノイズの存在は微小信号の識別を著しく困難にする。このことから、大小さまざまな定常ノイズに埋もれている非定常ノイズを同定・識別することは非常に重要であると考えられる。

非定常ノイズは、事前に波形の形状が分かっている既知の非定常ノイズと、正確な波形予測が困難な未知の非定常ノイズの2つに分けられる。深層学習を用いた非定常ノイズの識別を想定した場合、既知の非定常ノイズは、モデル学習のための教師データとして使用することができるが、未知の非定常ノイズは、波形の形状が分からないため、モデル学習のための教師データとして使用することができない。また、非定常ノイズの識別だけでなく、正確な波形予測が困難な解析対象である微小信号を直接識別することを想定した場合、非定常ノイズの識別と同じく、解析対象である微小信号を教師データとして用いることは不可能である。したがって、大小さまざまな定常ノイズを含む既知の非定常信号のみを用いた、定常ノイズに埋もれた未知の非定常信号を識別する方法が必要である。ここで非定常信号とは、正確な波形予測が困難な解析対象である微小信号と非定常ノイズの総称である。

1つの方法として、深層学習モデルに対し、波形予測が可能な既知の非定常信号のみを教師データとした、異常検知問題が挙げられる。これは、定常ノイズに既知の非定常信号が含まれているデータを正常データ、定常ノイズに未知の非定常信号が含まれているデータを異常データとした場合に、正常データのみにも適合する特徴を抽出することで、正常データの特徴とは一致しない異常データを識別するという方法である。

これを実現する手法として、変分オートエンコーダ (Variational Autoencoder ; VAE) を用いた異常検知による、大小

さまざまな定常ノイズに埋もれた未知の非定常信号の有無を判定するモデルの可能性について考える。具体的には、既知の非定常信号が含まれている定常ノイズのみのデータを画像にし変分オートエンコーダに入力することで、既知の非定常信号の特徴を潜在変数に対して標準正規分布に従うよう学習させる。また、最終的な目標として、学習済みモデルの潜在変数を用いて構築した特徴量空間に識別境界を引くことで、未知の非定常信号を含む定常ノイズデータを異常データとして識別を行う。本稿では最終的な目標の前段階として、既知の非定常信号が埋められた定常ノイズと未知の非定常信号が埋められた定常ノイズのデータセットにおける潜在変数の平均と分散、距離分布を確認することで、変分オートエンコーダを用いた異常検知の実現の可能性について考える。

## 2 実験概要

### 2.1 入力データの生成

本稿では、既知の非定常信号を時間変動により周波数が緩やかに変化する波形、未知の非定常信号を減衰信号波形とする。入力データとして、既知の非定常信号を含む定常ノイズのデータと未知の非定常信号を含む定常ノイズのデータに対し、短時間フーリエ変換を行うことで、周波数強度を明度、位相情報を色相としたスペクトログラムを獲得する。

今回は、時間変化により振動数が 50Hz~500Hz 間で緩やかに変化する波形を既知の非定常信号波形とするため、既知の非定常信号波形を、

$$S_{\text{nor}}(t) = \cos\left(2\pi\left\{\alpha f_0 t + \frac{1}{2}\beta(f_0 t)^2 + \frac{1}{3}\gamma(f_0 t)^3\right\}\right), \quad (1)$$

$$f(t) = f_0\{\alpha + \beta(f_0 t) + \gamma(f_0 t)^2\}, \quad (2)$$

に従って生成した。ここで、nor は正常データ、 $f_0$  は基準周波数、 $t$  は時間成分を表しており、今回は  $f_0$  を 500Hz とし、 $t$  の範囲を -1 から 1 までの 2 秒間とした。時間によって変動する周波数  $f(t)$  は  $\alpha, \beta, \gamma$  のパラメータ値によって決定される。今回は各パラメータを、

$$\alpha = \frac{p+q}{2}, \quad (3)$$

$$\beta = \frac{p-q}{1000}, \quad (4)$$

$$\gamma < 8.0 \times 10^{-8}, \quad (5)$$

に従う範囲とし、 $p, q$  が  $0.1 < q < p < 1$  になるように値を決定した。これにより、時間変動により周波数が 50Hz か

<sup>†</sup> 長岡技術科学大学

<sup>‡</sup> 京都大学

ら500Hz間で緩やかに変化する波形を生成することが可能である。

続いて、未知の非定常信号である減衰信号を、

$$S_{\text{ano}}(t) = e^{-(t+1)} \times \sin(2048(t+1)b), \quad (6)$$

により導出した。ここで、anoは異常データ、2048は周波数の上限値を表しており、 $b$ は0から $2\pi$ までのランダムな値とする。これにより、0Hzから2048Hzの範囲で減衰信号を生成することが可能である。

次に定常ノイズの生成を行う。今回は定常ガウス型ノイズを定常ノイズとするので、

$$S_{\text{noise}} = A_{\text{noise}} \times X_i, \quad (7)$$

によって生成した。ここで、 $i$ は時刻ラベル、 $X_i$ は $i$ の時点における標準正規分布に従う独立な乱数を表している。また、 $A_{\text{noise}}$ は定常ガウス型ノイズの振幅を表しており、正常データの場合、

$$\text{SNR} = \sqrt{\frac{8192}{2}} \left( \frac{A_{\text{nor}}}{A_{\text{noise}}} \right), \quad (8)$$

により決定される。8192は非定常信号波形のデータ長、 $A_{\text{nor}}$ は正常の非定常信号波形の振幅を表している。SNRは信号雑音比が1:1の場合は64、1:2の場合は32、1:3の場合は21.3となる。

これらの値を式(8)に代入することで正常データにおける $A_{\text{noise}}$ を導出する。定常ノイズ $S_{\text{noise}}$ に非定常信号波形 $S_{\text{nor}}(t)$ を足し合わせることで、定常ノイズに既知の非定常信号波形 $S_{\text{nor}}(t)$ が埋め込まれた信号 $S_{\text{signal}}$ を生成する。異常データの場合、式(1)と式(6)では正規化の方法が異なるため、SNRが式(8)と異なる。式(6)の場合はSNRをおよそ $\frac{1}{\sqrt{2}}$ 倍とする。

最後に、生成した各非定常信号波形を含む定常ノイズに対し短時間フーリエ変換を行うことで、位相情報を考慮したスペクトログラムを獲得する。短時間フーリエ変換を、窓幅を512、スライド幅を128、サンプリング周波数を4096Hzとして行うことで周波数強度 $S_{\text{power}}$ と位相情報 $S_{\text{phase}}$ を獲得する。獲得した周波数強度と位相情報に対し、

$$R_{\text{cos}} = \frac{\cos(S_{\text{phase}} + 1)}{2}, \quad (9)$$

$$G_{\text{cos}} = \frac{\cos((S_{\text{phase}} - 2\pi/3) + 1)}{2}, \quad (10)$$

$$B_{\text{cos}} = \frac{\cos((S_{\text{phase}} + 2\pi/3) + 1)}{2}, \quad (11)$$

$$P_{\text{color}} = \tanh(\lambda \times S_{\text{power}}) \times \frac{C_{\text{cos}}}{\max(C_{\text{cos}})}, \quad (12)$$

を行うことで、周波数強度を明度、位相情報を色相としたスペクトログラムを生成した。ここで $P_{\text{color}}$ とは、Red/Green/Blueの各色を表しており、 $C_{\text{cos}}$ に $R_{\text{cos}}$ 、 $G_{\text{cos}}$ 、 $B_{\text{cos}}$ を入力することで得られる。また $\lambda$ は、明度を調整する係数であり、信号雑音比ごとに値を決定する必要がある。本稿では、信号雑音比1:1の場合は $3.0 \times 10^{-2}$ 、信号雑音比1:2の場合は $1.5 \times 10^{-2}$ 、信号雑音比1:3の場合は $8.0 \times 10^{-3}$ とした。

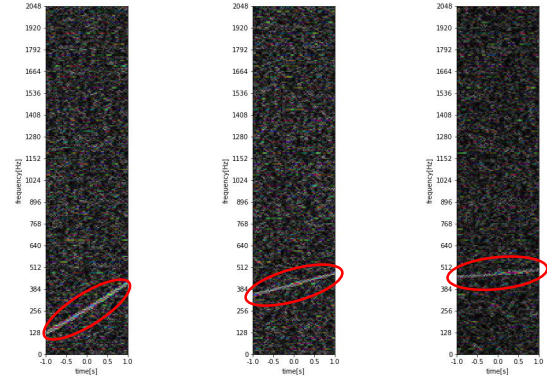


図1 信号雑音比 1:1 図2 信号雑音比 1:2 図3 信号雑音比 1:3

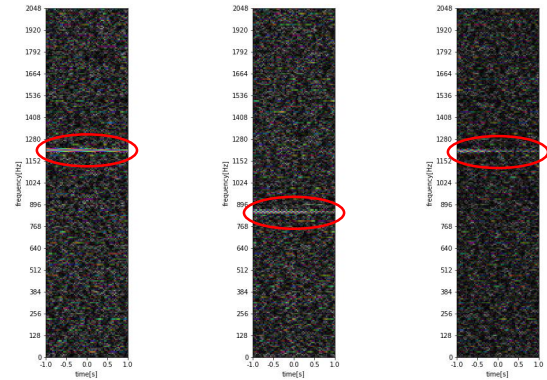


図4 信号雑音比 1:1 図5 信号雑音比 1:2 図6 信号雑音比 1:3

既知の非定常信号波形のスペクトログラムを図1-3、未知の非定常信号波形のスペクトログラムを図4-6に示す。ここで、図1-6の赤枠内に見えるラインが非定常信号である。

## 2.2 変分オートエンコーダの構成

本節では、変分オートエンコーダの構成について説明を行う。変分オートエンコーダとは、エンコーダ部、デコーダ部の2つで構成され、エンコーダ部で入力データの次元数を削減し、デコーダ部で削減された次元数から元の画像を復元する。その際に入力データ $x$ を表現する特徴を潜在変数 $z$ に確率分布の構造で落とし込む手法である。モデルの学習は、

$$L(x) = -E_{p_{\theta}(z|x)}[\log q_{\phi}(x|z)] + \text{KL}(p_{\theta}(z|x)||q_{\phi}(z)), \quad (13)$$

によって行われ、ここで $p_{\theta}(z|x)$ はエンコーダに $x$ を入力したときにおける $z$ の確率分布を表しており、 $q_{\phi}(x|z)$ はデコーダに $z$ を入力したときにおける $x$ の確率分布を表している。 $\theta$ はエンコーダ、 $\phi$ はデコーダのパラメータを示している。つまり、 $E_{p_{\theta}(z|x)}[\log q_{\phi}(x|z)]$ は、入力データ $x$ を潜在変数 $z$ に次元を落とし、潜在変数 $z$ から入力画像 $x$ を復元したときにおける対数尤度の期待値を最大化するという意味が込められている。また、 $\text{KL}(p_{\theta}(z|x)||q_{\phi}(z))$ はカルバック・ライブラー情報量と言い、事前分布 $q_{\phi}(z)$ と事後分布 $p_{\theta}(z|x)$ の類似度を評価している。本稿では、潜在変数 $z$ を標準正規分布に従うように学習させるため、事前分布 $q_{\phi}(z)$ を標準正規分布とする。

今回構築した変分オートエンコーダのエンコーダ部を図7に、デコーダ部を図8に示す。エンコーダの入力は(256,61,3)のカラー画像(周波数強度を明度、位相情報を色相としたスペ

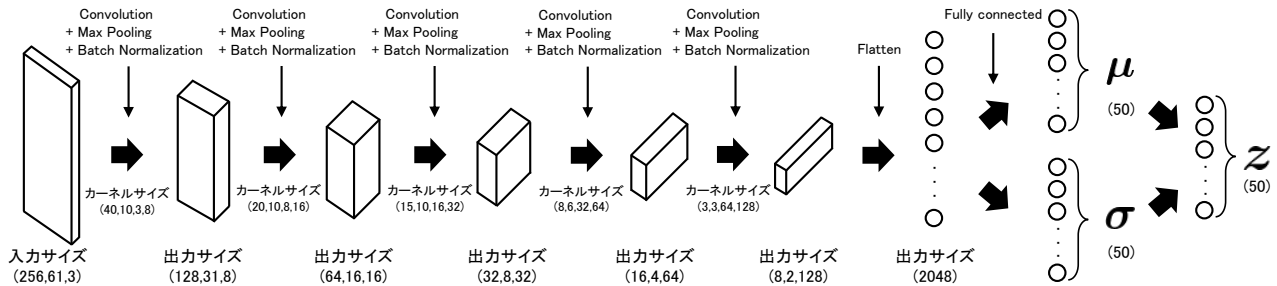


図 7 VAE のエンコーダ部

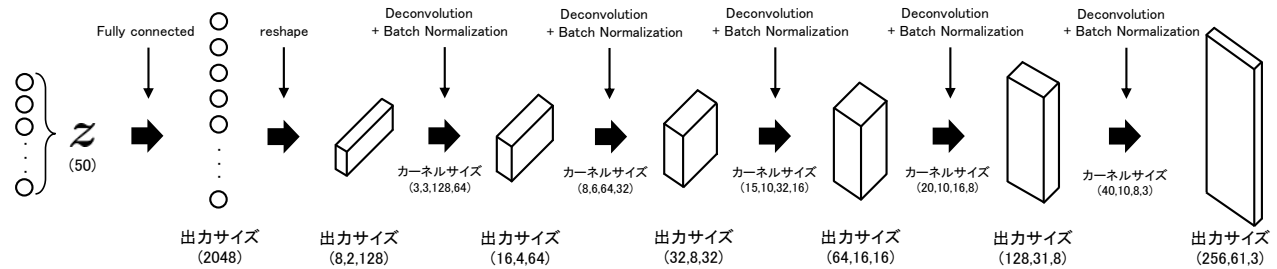


図 8 VAE のデコーダ部

クトログラム) とし、各カーネルフィルタによる畳み込み処理を行った。畳み込み層からの出力に対し、(2,2) のフィルタによるプーリング処理を行うことで次元数を減らした。またバッチ正規化層を加えることで、学習速度の向上や過学習の防止を試みた。最後に全結合層によって、2048 次元から平均ベクトル  $\mu$  として 50 次元、標準偏差ベクトル  $\sigma$  として 50 次元まで次元数を削減し、潜在変数  $z$  を、

$$z = \mu + N \cdot \sigma, \quad (14)$$

により算出した。ここで  $N$  は標準正規分布に従うベクトルを表している。

デコードの入力は 50 次元の潜在変数  $z$  であり、全結合層により 2048 次元まで復元する。その後、サイズを (8,2,128) に再形成し、逆畳み込み処理を行い、バッチ正規化層に入力することで元のカラー画像を復元する。

変分オートエンコーダの構成は、層数やカーネルフィルタのサイズ等を増減させたときにおける評価関数  $L(x)$  の値の減少具合によって決定した。また、各層の出力には活性化関数として ReLU 関数を適用し、 $\mu$  の出力部のみ tanh 関数を適用した。

本稿では、学習データとして、正常データを各信号雑音比ごとに 10,000 データずつ用意し、このモデルに入力することで学習させた。またテストデータとして、正常データを各信号雑音比ごとに 2,000 データずつ、異常データを各信号雑音比ごとに 400 データずつ用意した。学習はバッチ数を 16、エポック数を 50 とした確率的勾配降下法により行った。

### 3 結果と考察

今回は最終的な目標である、変分オートエンコーダの潜在変数  $z$  を用いた異常検知による、非定常信号を含む定常ノイズデータの識別の可能性を確認すべく、正常・異常データにお

表 1 各信号雑音比における正常・異常データの  $z$  の平均・分散

信号雑音比	正常データ		異常データ	
	平均	分散	平均	分散
1:1	$-1.5 \times 10^{-2}$	0.707	$1.5 \times 10^{-2}$	0.493
1:2	$-2.0 \times 10^{-3}$	0.819	$2.8 \times 10^{-2}$	0.853
1:3	$5.6 \times 10^{-6}$	0.842	$2.0 \times 10^{-3}$	0.861

る潜在変数  $z$  の平均と分散、また正常データにおける潜在変数  $z$  の平均から、正常・異常データにおける潜在変数のプロットまでのユークリッド距離の分布を導出した。結果を表 1、図 9-11 に示す。図 9-11 の  $x$  軸は、正常データにおける潜在変数  $z$  の平均から正常・異常データにおける潜在変数のプロットまでのユークリッド距離を、 $y$  軸は確率密度を表している。

本稿では、異常検知問題を行うため、正常データと異常データにおける潜在変数  $z$  の平均が離れており、また、今回は正常データの潜在変数を標準正規分布に従うよう学習させているので、正常データの平均が 0、分散が 1 となっている状態が好ましいと考える。さらに、正常・異常データの潜在変数  $z$  のプロットの距離は互いに離れている状態が好ましいので、正常データと異常データの距離分布が離れている状態を良いとする。

表 1 の正常・異常データの平均に着目してみると、全ての信号雑音比に対して平均値が離れており、信号雑音比 1:3 に関しては正常データと異常データとの距離が約 1,000 倍離れていることが確認された。また正常データの平均、分散に着目すると、平均は限りなく 0 に近い値となっており、分散はおおよそ 1 に近い値であることが確認された。これらのことから、変分オートエンコーダで正常データの大きな特徴を捉えることができ、正常データの特徴とは適合しない異常データを識別できるのではないかと考えられる。

しかし、図 9-11 に着目すると、図 9 の場合は、正常データの距離分布と異常データの距離分布は、お互いに僅かに離れてい

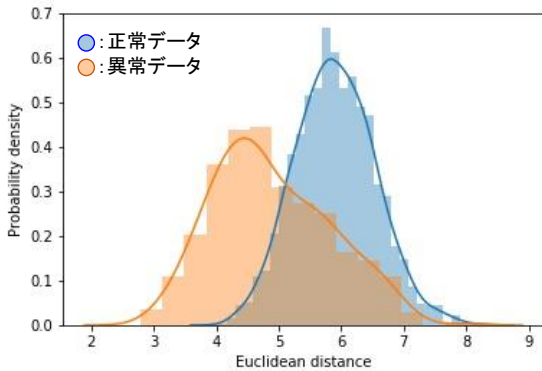


図9 信号雑音比 1:1 の場合における距離分布

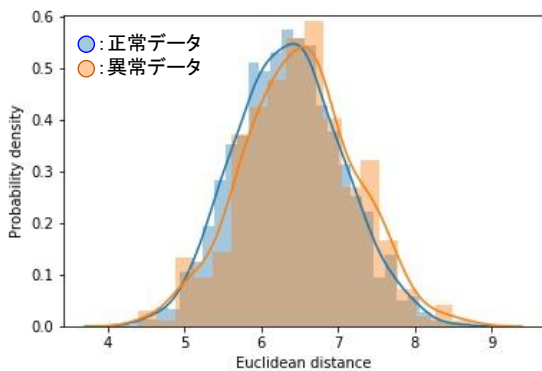


図10 信号雑音比 1:2 の場合における距離分布

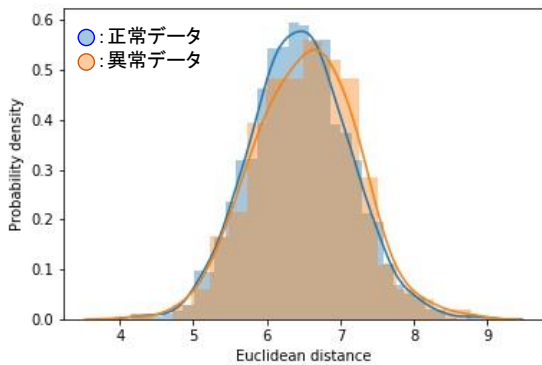


図11 信号雑音比 1:3 の場合における距離分布

るが、図10-11の場合は、正常と異常データの距離分布がほとんど重なっていることが確認された。このことから、空間に識別境界を引いた際に、異常データを識別することは困難だと考えられる。

距離分布が重なる原因として、2つの点が考えられる。1つ目は潜在変数の次元数が高いことが挙げられる。これは、既知の非正常信号を表現する特徴は非常に単純であるため、現在の潜在変数の次元数では、既知の非正常信号を表現する特徴以外の情報も保有してしまっていると考えられる。このことから、正常データと異常データの潜在変数の値が近くなってしまい、距離分布が重なってしまったと考えられる。したがって、潜在変

数の次元数を削減することで、既知の非正常信号を表現する特徴のみを抽出することができ、結果として正常と異常データの距離分布を遠ざけることが可能になるのではないかと考えられる。2つ目は正常と異常データの平均値が近いことが挙げられる。表1から、正常と異常データの平均の値が離れていることは確認されたが、分散の値を考慮すると、平均の値を1以上離す必要がある。しかし、今回は平均を導出するニューラルネットワークの活性化関数として  $\tanh$  関数を用いたため、平均の値を1以上離すことは不可能だと考えられる。したがって、1つ目の原因を解消した後に、平均を導出するニューラルネットワークの活性化関数を変えることで、より正常、異常データの平均の値を離すことができ、結果として点の発生個所を遠ざけることが出来るのではないかと考えた。

#### 4 おわりに

本稿では、既知の非正常信号と未知の非正常信号をカラー画像（周波数強度を明度、位相情報を色相としたスペクトログラム）とし、変分オートエンコーダに入力した際における潜在変数の平均と分散、また正常データの平均からプロットまでのユークリッド距離の分布を確認することで、正確な波形予測が困難な未知の非正常信号を含む定常ノイズデータを異常データとして識別を行うことが可能か検討した。結果は、現在の状態では定常ノイズの振幅が非正常信号に対して大きい場合に、未知の非正常信号を含む定常ノイズデータを識別することは困難であることが確認された。今後の展望は、潜在変数の次元削減や、潜在変数の平均を導出する全結合層の活性化関数を変更することで、潜在変数の平均・分散がどのように変化するか、また正常と異常の距離分布がどれだけ離れるかを確認することで、変分オートエンコーダを用いた異常検知の実現の可能性についてより深い検討を行っていきたい。また、本稿では最終的に潜在変数  $z$  を一変数に落として解析を行ったが、本来ベクトル量である  $z$  をベクトル量で解析を行うことが必須である。この点についてさらに検討を進めて報告する。

#### 5 謝辞

本研究の一部は、文科省科学研究費 17H06358 からの支援を受けています。

#### 参考文献

- [1] 岡田佳之, 福井健一, 沼尾正行, “多次元数値観測量の事象系列に対するクラスタ系列パターンの抽出”, 人工知能学会全国大会論文集, 2015.
- [2] 三代木伸二, “宇宙を観測する新しい窓・重力波”, *RADIOISOTOPES*, vol.66, pp.61-75, 2017.
- [3] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos, S. L. Larson, T. K. Lee, C. Lintott, T. B. Littenberg, A. Lundgren, C. Østerlund, J. R. Smith, L. Trouille and V. Kalogera, “Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science”, *Classical and Quantum Gravity*, vol.34, no.9, 2017.