

修士論文

Variational AutoEncoder を用いた教師なし学習に  
よる肺レントゲン画像の潜在空間モデルの構築

安形 俊輝

奈良先端科学技術大学院大学

先端科学技術研究科

情報理工学プログラム

主指導教員: 金谷 重彦 教授

(情報科学領域)

令和2年3月13日提出

本論文は奈良先端科学技術大学院大学先端科学技術研究科に  
修士(工学)授与の要件として提出した修士論文である。

安形 俊輝

審査委員：

金谷 重彦 教授 (主指導教員)

佐藤 嘉伸 教授 (副指導教員)

小野 直亮 准教授 (副指導教員)

黄 銘 助教 (副指導教員)

# Variational AutoEncoder を用いた教師なし学習による肺レントゲン画像の潜在空間モデルの構築\*

安形 俊輝

## 内容梗概

今日、医師不足の問題を背景に、CT や MRI から得られる画像から深層学習によって診断を自動的に行う研究が盛んに行われている。深層学習モデルは、入力データを表現するのに最適化された特徴量を学習できることが特徴である。これは、深層学習における中間層における表現は、入力データを適切な低次元変数として表すことができることを意味している。ただし、先行研究においては、予めラベルデータが準備されたデータセットに対する教師あり学習での識別が主流であり、画像の細部まで確認が必要な診断には多大なコストが発生していた。そこで、本研究では、Variational AutoEncoder (VAE) を用いた教師なし学習によって特徴量を抽出するモデルを構築した。さらに、この特徴量を PCA によって解析し、分布の平均から差異を表す指標を算出し、医学的な所見との対応関係をフィッシャーの正確確率検定によって確認した。その結果、7つの指標のうち3つの指標において統計的に有意な対応関係が確認できた。

## キーワード

深層学習, Variational AutoEncoder (VAE), レントゲン画像, 潜在空間, 分布解析

---

\*奈良先端科学技術大学院大学 先端科学技術研究科 修士論文, 令和2年3月13日.

# Construction of a Latent Space Model of Chest X-ray Images by Unsupervised Learning Using Variational AutoEncoder\*

Toshiki Agata

## Abstract

In recent years, there has been a lot of research on computer aided diagnosis by deep learning using medical images, such as CT and MRI images in the due to the shortage of doctors. A Advantage of the deep learning model is that it can learn features that are optimized to represent input data. This implies that the representation in the hidden layer in deep learning can represent the input data as appropriate low-dimensional variables. However, in previous researches, discrimination by supervised learning for a data set in which label data was prepared in advance was indispensable, and a large cost for diagnosis was required to evaluate image details. In this study, we constructed a model to extract features by unsupervised learning using Variational AutoEncoder (VAE). The features were analyzed by PCA, an index indicating the difference was calculated from the mean of the distribution, and the correspondence with medical findings was confirmed by Fisher's exact test. As a result, a statistically significant correspondence was confirmed in three of the seven indices.

## Keywords:

deep leaning, Variational AutoEncoder (VAE), X-ray image, latent space, analysis of distribution

---

\*Master's Thesis, Graduate School of Science and Technology, Nara Institute of Science and Technology, March 13, 2020.

# 目次

<b>1. 研究背景</b>	<b>1</b>
1.1 画像データに対する深層学習	2
1.2 医用画像への深層学習の応用	2
<b>2. データセット</b>	<b>6</b>
2.1 データ前処理	6
<b>3. 提案手法</b>	<b>7</b>
3.1 提案手法の流れ	7
3.2 AutoEncoder	7
3.2.1 Encoder	8
3.2.2 Decoder	8
3.2.3 損失関数	8
3.3 Variational AutoEncoder (VAE)	8
3.3.1 損失関数	9
3.3.2 変分下限の解釈	10
3.3.3 Reparameterization Trick	10
3.3.4 作成した VAE	11
3.4 PCA (Principal Component Analysis)	11
<b>4. 実験結果</b>	<b>12</b>
4.1 VAE による再構成画像	12
4.2 潜在変数の主成分分析	13
4.3 潜在空間の重心とのノルム	15
4.4 各主成分の分布	17
4.5 主成分の重心との距離	23
4.6 分布の平均から離れたデータ	30
4.6.1 潜在変数空間と診断の関係	34

<b>5. 考察</b>	<b>35</b>
5.1 第1主成分について . . . . .	35
5.2 1000次元潜在変数の第3主成分 . . . . .	37
5.3 抽出された画像について . . . . .	37
5.3.1 背骨に関する所見 . . . . .	38
<b>6. 総括</b>	<b>40</b>
謝辞	41
参考文献	42

## 目 次

1.1	人口 1000 人あたりの医師数 . . . . .	1
1.2	ニューラルネットワークの中間層における特徴量の可視化（第一層はエッジなどの局所的な特徴量を得ているのに対し、第 5 層では、物体や動物の全体像の特徴量を得ている） [1] . . . . .	4
1.3	潜在変数を連続的に変化させたときの生成画像 . . . . .	5
3.1	提案手法の流れ . . . . .	7
3.2	AutoEncoder の概要図 . . . . .	7
3.3	サンプリングを含む VAE . . . . .	10
3.4	サンプリングを含む VAE . . . . .	10
3.5	本研究で作成した VAE . . . . .	11
4.1	VAE による再構成画像（潜在変数：500 次元，上段：入力画像，下段：再構成画像） . . . . .	12
4.2	VAE による再構成画像（潜在変数：1000 次元，上段：入力画像，下段：再構成画像） . . . . .	13
4.3	PC1, PC2 の分布（潜在変数：500 次元） . . . . .	14
4.4	PC1, PC2 の分布（潜在変数：1000 次元） . . . . .	15
4.5	500 次元の潜在変数から算出された $\delta^{(n)}$ . . . . .	16
4.6	1000 次元の潜在変数から算出された $\delta^{(n)}$ . . . . .	17
4.7	第 1 主成分の分布 (潜在変数：500 次元) . . . . .	18
4.8	第 2 主成分の分布 (潜在変数：500 次元) . . . . .	19
4.9	第 3 主成分の分布 (潜在変数：500 次元) . . . . .	20
4.10	第 1 主成分の分布 (潜在変数：1000 次元) . . . . .	21
4.11	第 2 主成分の分布 (潜在変数：1000 次元) . . . . .	22
4.12	第 3 主成分の分布 (潜在変数：1000 次元) . . . . .	23
4.13	$d^{(n,1)}$ (潜在変数：500 次元) . . . . .	24
4.14	$d^{(n,2)}$ (潜在変数：500 次元) . . . . .	25
4.15	$d^{(n,3)}$ (潜在変数：500 次元) . . . . .	26
4.16	$d^{(n,1)}$ (潜在変数：1000 次元) . . . . .	27

4.17 $d^{(n,2)}$ (潜在変数：1000 次元)	28
4.18 $d^{(n,3)}$ (潜在変数：1000 次元)	29
4.19 $d^{(n)}$ の値が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	30
4.20 第 1 主成分の値が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	31
4.21 第 2 主成分の値が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	31
4.22 第 3 主成分の値が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	32
4.23 $d^{(n,1)}$ が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	32
4.24 $d^{(n,2)}$ が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	33
4.25 $d^{(n,3)}$ が大きい画像（上：潜在変数；500 次元，下：潜在変数；1000 次元）	33
5.1 男女別の第 1 主成分（潜在変数:500 次元）	36
5.2 第 1 主成分	37
5.3 男女別の第 1 主成分（潜在変数:1000 次元）	37
5.4 第 3 主成分（潜在変数：1000 次元）において側弯が確認された画像	39
5.5 第 3 主成分（潜在変数：1000 次元）において側弯が確認された画像	39

## 表 目 次

4.1 異常性が確認された画像の枚数と p 値（潜在変数：500 次元）	34
4.2 異常性が確認された画像の枚数と p 値（潜在変数：1000 次元）	34
5.1 抽出された画像の症例とその件数	38



# 1. 研究背景

今日、医師不足が問題となっている。OECDの調査 [2] によると、人口1000人あたりの臨床医師数はオーストリアやノルウェーと比較して半数以下である。我が国は、世界的に見てCTやMRIの台数が多く、検査件数も世界最多水準であるが、放射線科医の不足が大きな問題となっている。その結果、CTやMRIといった高度画像検査において画像診断報告書がされないことがある。また、1検査の画像診断に費やすことができる時間も短くなり、見落としや誤診の危険性が増すことが指摘されている。このような問題を背景に、診断の自動化が大きな注目を集めている。

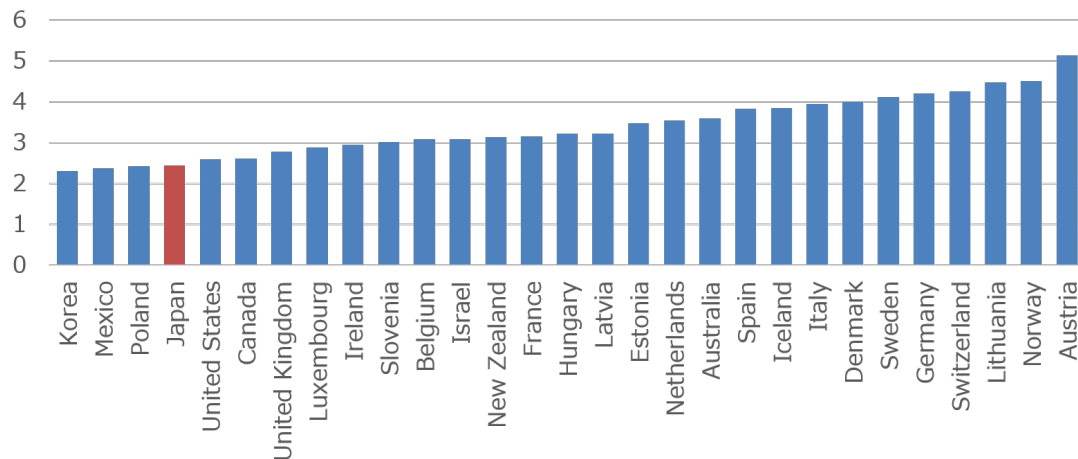


図 1.1 人口1000人あたりの医師数

近年、深層学習を応用した画像診断に関する研究が盛んに行われている。深層学習とは、人間の脳内の神経回路を模倣した機械学習手法である。このモデルはニューラルネットワークと呼ばれ、多段に重ねることによってモデルをより現実問題に即した表現力が得られることが知られている。画像認識においては、全結合型ではないニューラルネットワークである畳み込みニューラルネットワークがよく用いられる。

## 1.1 画像データに対する深層学習

2012年にILSVRC(ImageNet Large Scale Visual Recognition Challenge)での物体カテゴリ認識において、優勝したのは深層学習のモデルの一つであるAlexNet[3]である。これ以前は、画像の局所的特徴量であるScale-Invariant Feature Transform (SIFT)[4]やそれを高速化したSpeeded UP Robust Feature (SURF)[5]やHistograms of Oriented Gradients (HoG)[6]などと機械学習を用いた画像認識が主流であったが、深層学習が大きな注目を集めることになった。深層学習の応用タスクは、前述の画像分類のみならず、物体検出[7]、セグメンテーション[8]、画像生成[9]など様々なものが挙げられる。

## 1.2 医用画像への深層学習の応用

深層学習は、画像データに対する分類やセグメンテーションを高い精度で行うことができ、風景画像等のみならず、医用画像に対しても高精度で識別が可能である。肺のレントゲン画像に対しては、疾患の有無だけでなく、複数の症状が見られるデータセットに対しても高精度での識別[10]ができるようになっている。また、病理画像から細胞等の注目領域を学習し、その領域の切り分けを行うセグメンテーションにも高精度で成功している[11]。

ただし、現段階においては、予め症例が明らかになっているデータから特徴量を抽出し、分類等を行う教師あり学習が主流である。また、中間層の出力にも関心が寄せられており、層を重ねるごとに局所的な特徴量から大局的な特徴量が抽出されることが示されている[1] (図1.2参照)。さらに、最近では、中間層の出力を

応用することで、ノイズの入った画像からノイズのない画像を生成する研究 [12] が行われるなど中間層の出力の有用性が示されている。

ニューラルネットワークの中間層を用いた特徴量抽出としては、AutoEncoder[13] やVAE[14]がある。VAEは、中間層に正規分布を仮定することによって、潜在変数の連続的な分布を学習することができる。潜在変数を連続的に変化させると生成画像も連続的に変化する（図 1.3 参照）。

以上より、肺レントゲン画像を入力データとした際も、潜在変数の連続的な分布を得ることができ、それは、肺レントゲン画像の特徴（正常，異常）と対応していると考えられる。そこで、本研究では、VAEによる教師なし学習によって得られた特徴量の分布を確認する。それと同時に実際の診断との関係性を確認する。



図 1.2 ニューラルネットワークの中間層における特徴量の可視化（第一層はエッジなどの局所的な特徴量を得ているのに対し，第5層では，物体や動物の全体像の特徴量を得ている） [1]

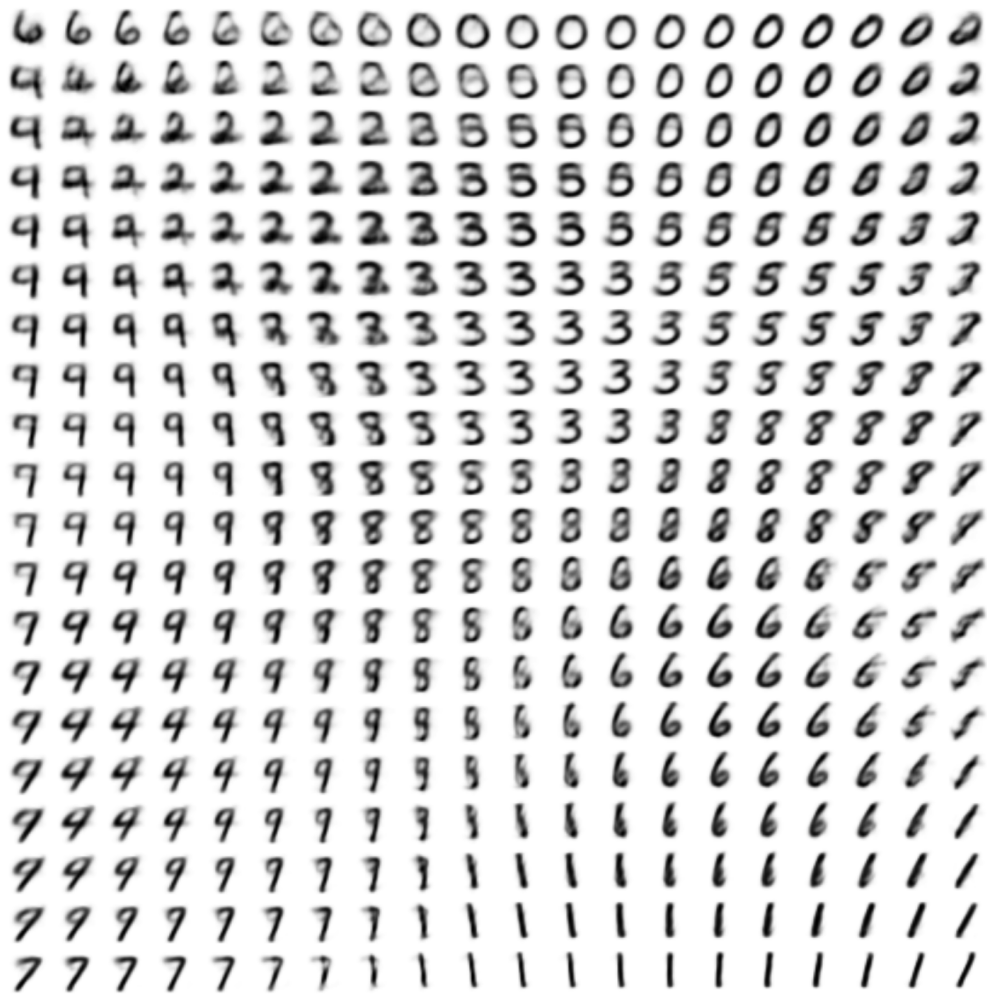


図 1.3 潜在変数を連続的に変化させたときの生成画像

## 2. データセット

本研究で用いるデータセットは東京大学 保健センターより提供された健康診断時に撮影されたレントゲン画像である。この画像の形式はDICOM形式であり、オリジナルのデータは画像と撮影時の情報が紐づいたデータである。学生およびスタッフの健康診断の際に、オプトアウト方式で被験者の同意を得たものであり、提供に際して匿名化されたものを用いている。なお、データの総数は80000枚である。

### 2.1 データ前処理

前述のデータセットは様々なサイズの画像が含まれているため、最も画像枚数の多いサイズが $3052 \times 2047$ である画像(44,725枚)を分布解析の対象とした。このうち、10,000枚をトレーニングデータとし、1,000枚をテストデータとして $240 \times 192$ にリサイズしたものをVAEの入力データ入力画像とした。

### 3. 提案手法

本章では，本研究で用いた学習モデルを紹介し，さらに，本研究で作成したモデルの構造を示す．

#### 3.1 提案手法の流れ

提案手法の流れを図 3.1 に示す．提案手法ではまず，入力画像から 3.3 で述べる VAE を用いて元の画像の大局的な特徴量を反映した分布を得るための学習を行う．その後 PCA を用いて次元圧縮を行い，その分布を解析する．



図 3.1 提案手法の流れ

#### 3.2 AutoEncoder

ここでは，VAE の基礎となる AutoEncoder（自己符号化器）について述べる．AutoEncoder とは，教師なし学習による特徴量抽出のための深層学習モデルである．特に Convolutional AutoEncoder (CAE) は画像に対して有効な畳み込み層から構成されるため，高い精度で意味のある次元圧縮を行うことができる．

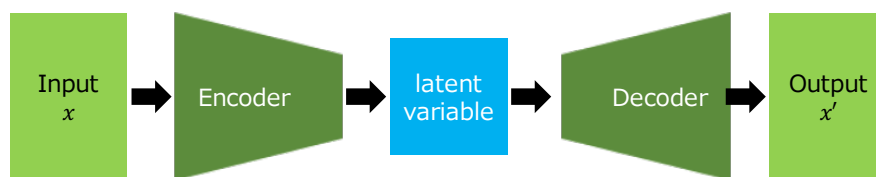


図 3.2 AutoEncoder の概要図

### 3.2.1 Encoder

Encoder は入力データ  $\mathbf{x}$  を低次元の潜在変数  $\mathbf{z}$  へ変換することを目的としたニューラルネットワークである。Encoder 第  $i$  層の出力  $\mathbf{y}_i$  は、入力  $\mathbf{x}_i$ 、重み  $W_i$ 、バイアス  $\mathbf{b}_i$ 、活性化関数  $f$  を用いて式 (3.1) によって表される。

$$\mathbf{y}_i = f(W_i \mathbf{x}_i + \mathbf{b}_i) \quad (3.1)$$

### 3.2.2 Decoder

Decoder は潜在変数  $\mathbf{z}$  を入力データ  $\mathbf{x}$  と等しい次元の空間へ変換することを目的としたニューラルネットワークである。Decoder 第  $i$  層の出力  $\mathbf{y}'_i$  は、入力  $\mathbf{x}'_i$ 、重み  $W'_i$ 、バイアス  $\mathbf{b}'_i$ 、活性化関数  $f$  を用いて式 (3.1) によって表される。

$$\mathbf{y}'_i = f(W'_i \mathbf{x}'_i + \mathbf{b}'_i) \quad (3.2)$$

### 3.2.3 損失関数

AutoEncoder 全体の最適化には、入力データ  $X$  と出力データ  $X'$  の平均二乗誤差  $L(X, X')$  を用いる。ここで、 $N$  はサンプル数を表す。

$$L(X, X') = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}'_i - \mathbf{x}_i\|^2 \quad (3.3)$$

## 3.3 Variational AutoEncoder (VAE)

Variational AutoEncoder (VAE) も AutoEncoder と同様に、Encoder と Decoder を持つ深層学習のモデルであるが、潜在変数に対して下記の仮定を置いている点異なる。

- 潜在変数  $\mathbf{z}$  は確率分布  $p(\mathbf{z})$  に従う。
- 入力データ  $\mathbf{x}$  は条件付き確率分布  $p(\mathbf{x}|\mathbf{z})$  に従う。



潜在変数の分布を求められれば，入力データを次元圧縮された表現である  $z$  の推定ができるが，周辺尤度

$$p(\mathbf{x}) = \int p(\mathbf{x}|\mathbf{z})p(\mathbf{z})d\mathbf{z}$$

や事後分布

$$p(\mathbf{z}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{x})}$$

を直接的に求めることは一般に困難である．そこで，以下に述べる計算によってこの問題を回避する．

### 3.3.1 損失関数

ここでは，潜在変数の真の分布  $p(\mathbf{z}|\mathbf{x})$  を近似する分布を  $q(\mathbf{z}|\mathbf{x})$  とする．VAE においては，このとき， $p(\mathbf{z}|\mathbf{x})$  と  $q(\mathbf{z}|\mathbf{x})$  のカルバック・ライブラー・ダイバージェンス  $D_{\text{KL}} [q(\mathbf{z}|\mathbf{x})||p(\mathbf{z}|\mathbf{x})]$  を考える．VAE の目的は，周辺尤度  $p(\mathbf{x})$  の最大化である． $p(\mathbf{x})$  が最大化されるとき， $\log p(\mathbf{x})$  も最大化されるので，この最大化を考える．

$$\begin{aligned} \log p(\mathbf{x}) &= \log \int p(\mathbf{x}, \mathbf{z})d\mathbf{z} \\ &= \log \int q(\mathbf{z}|\mathbf{x}) \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} \\ &\geq \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} (\because \text{イエンセンの不等式}) \\ &= \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} \\ &= \int q(\mathbf{z}|\mathbf{x}) \left( -\log \frac{q(\mathbf{z}|\mathbf{x})}{p(\mathbf{z})} + \log p(\mathbf{x}|\mathbf{z}) \right) d\mathbf{z} \\ &= -D_{\text{KL}} [q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})] + E_{q(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})] \end{aligned} \tag{3.4}$$

式(3.4)によって示される周辺尤度の下限  $L = -D_{\text{KL}} [q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})] + E_{q(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})]$  は変分下限 (evidence lower bound; ELBO) と呼ばれる．VAE の学習は，この変分下限を最大化するようにパラメータの更新が行われる．

### 3.3.2 変分下限の解釈

変分下限  $L(\mathbf{x}, \mathbf{z})$  は前節で示した AutoEncoder と同様の解釈を与えることができる。すなわち、 $q(\mathbf{z}|\mathbf{x})$  を Encoder,  $p(\mathbf{x}|\mathbf{z})$  を Decoder として解釈すると、VAE のモデルは図 3.3 のように表される。ここで、事前分布  $p(\mathbf{z})$  と事後分布  $q(\mathbf{z}|\mathbf{x})$  について、 $p(\mathbf{z}) = N(\mathbf{z}; 0, I)$ ,  $q(\mathbf{z}|\mathbf{x}) = N(\mathbf{z}; \boldsymbol{\mu}(\mathbf{x}), \boldsymbol{\sigma}(\mathbf{x})^2)$  であると仮定すると、変分下限の第 1 項  $D_{\text{KL}}[q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})]$  は、さらに単純化できる。

$$D_{\text{KL}}[q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})] = -\frac{1}{2} \sum_{j=1}^M (1 + \log \sigma_j^2 - \mu_j^2 - \sigma_j^2)$$

ここで、 $M$  は潜在変数  $\mathbf{z}$  の次元数を表す。また、 $E_{\mathbf{z}}[\log p(\mathbf{x}|\mathbf{z})]$  は AutoEncoder と同様の入力データと出力データの誤差を表す。

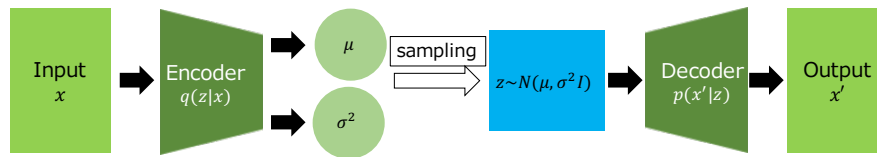


図 3.3 サンプリングを含む VAE

### 3.3.3 Reparameterization Trick

図 3.3 はサンプリングを含むため、誤差逆伝搬による最適化を行うことができない。そこで、 $\mathbf{z}$  のランダムサンプリングの代わりに、ランダムノイズ  $\epsilon \sim N(0, 1)$  を用いて、 $\mathbf{z} = \boldsymbol{\mu}(\mathbf{x}) + \epsilon \odot \boldsymbol{\sigma}(\mathbf{x})^2$  とすることで、誤差逆伝搬による最適化を可能にする。

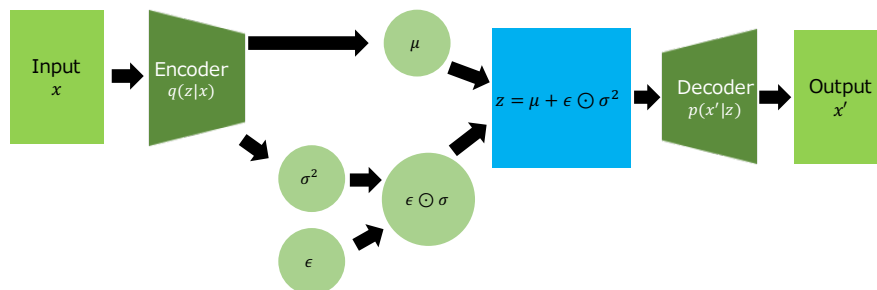


図 3.4 サンプリングを含む VAE

### 3.3.4 作成した VAE

本研究で作成した VAE の構造を図 3.5 に示す。ただし、Conv は畳み込み層、Strides はフィルターのストライドの大きさを表す。

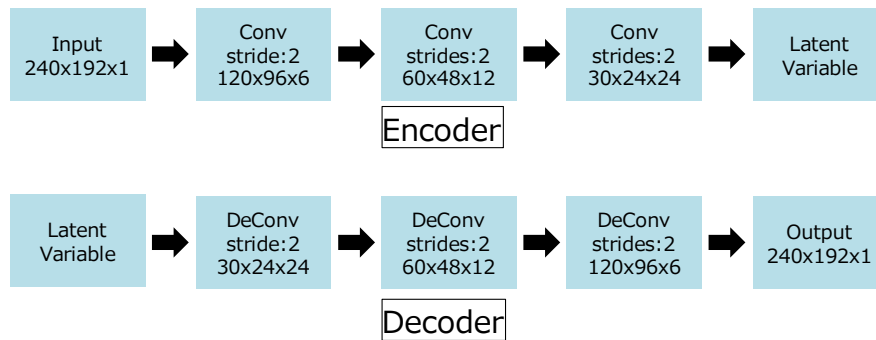


図 3.5 本研究で作成した VAE

本研究では、潜在変数  $z$  の 500 次元と 1000 次元に設定し、特徴量を抽出している。

## 3.4 PCA (Principal Component Analysis)

VAE によって抽出された 500 次元、1000 次元潜在変数は多次元正規分布に従っている。PCA によって、各正規分布の特徴量を組み合わせた特徴量が抽出できると考えられる。そこで、PCA による次元圧縮を行い、その分布の解析を行った。計算量の都合上、5000 枚を主成分の計算に用いた。その後、2 章で述べた全データについて PCA を適用した。以後、この PCA 適用後のデータを PCA データとする。

## 4. 実験結果

### 4.1 VAEによる再構成画像

図4.1, 図4.2にVAEによって再構成された画像を示す. VAEによって概形は再構成できたものの, 詳細部分の再現には至らなかった.

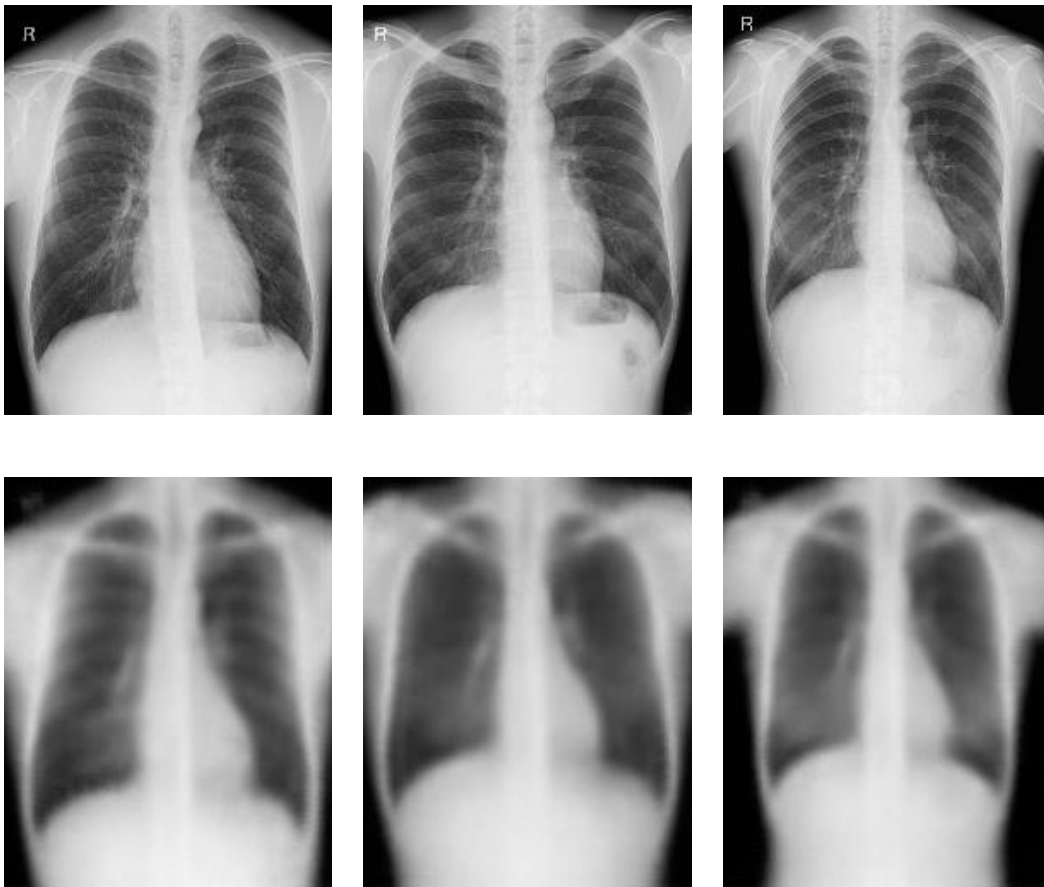


図 4.1 VAE による再構成画像 (潜在変数: 500 次元, 上段: 入力画像, 下段: 再構成画像)

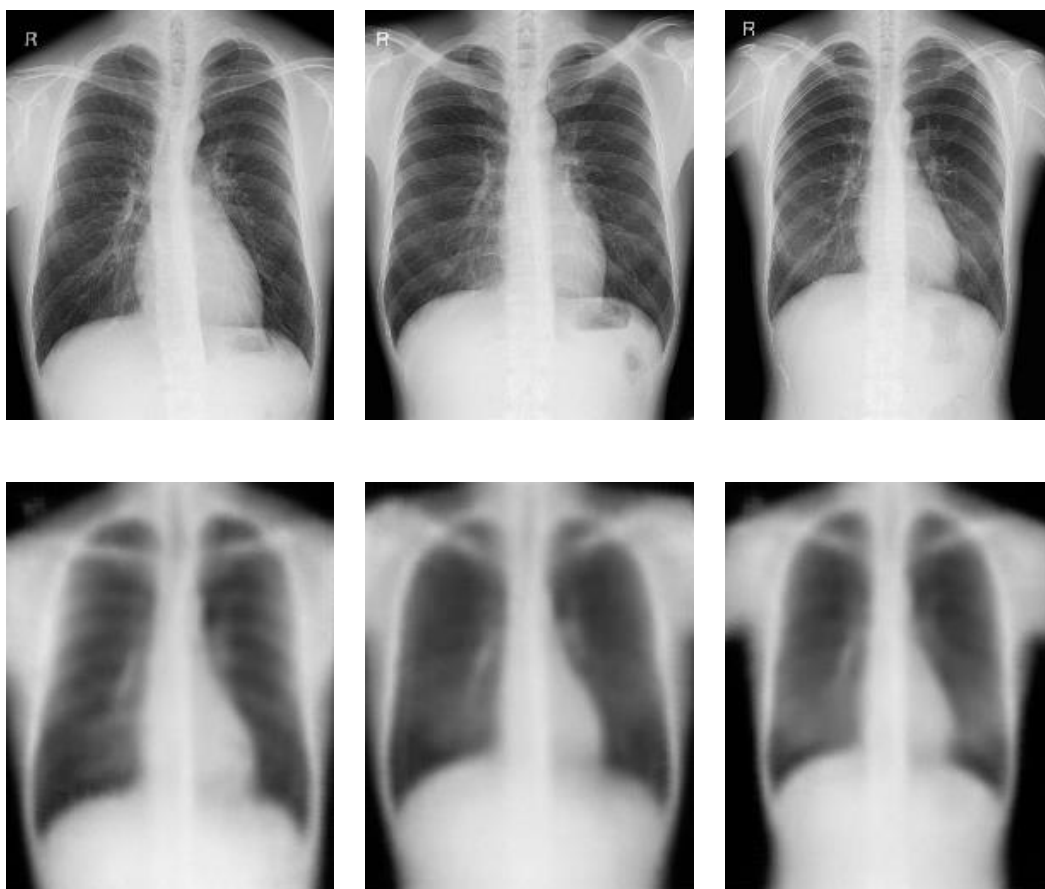


図 4.2 VAE による再構成画像（潜在変数：1000 次元，上段：入力画像，下段：再構成画像）

## 4.2 潜在変数の主成分分析

以降，VAE の潜在変数を  $z^{(n--)}$  とし，PCA データの第  $m$  成分を  $\hat{z}^{(n,m)}$  とする．  
 図 4.2，図 4.4 に横軸に  $\hat{z}^{(n,1)}$ ，縦軸に  $\hat{z}^{(n,2)}$  をとったときの分布を示す．

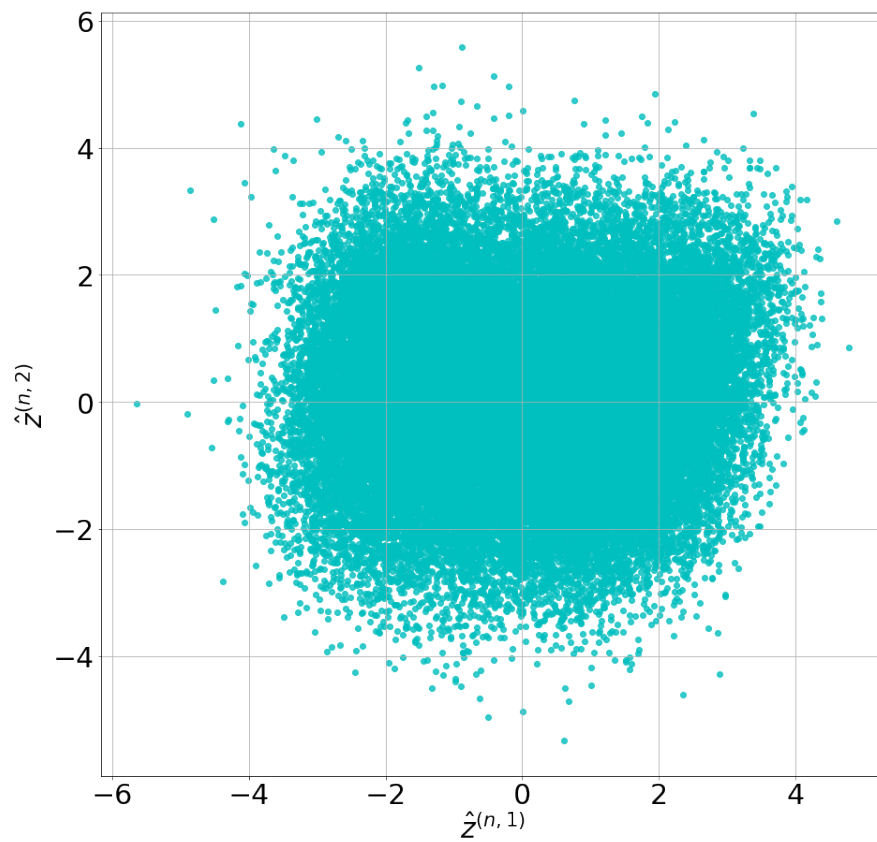


図 4.3 PC1, PC2 の分布 (潜在変数 : 500 次元)

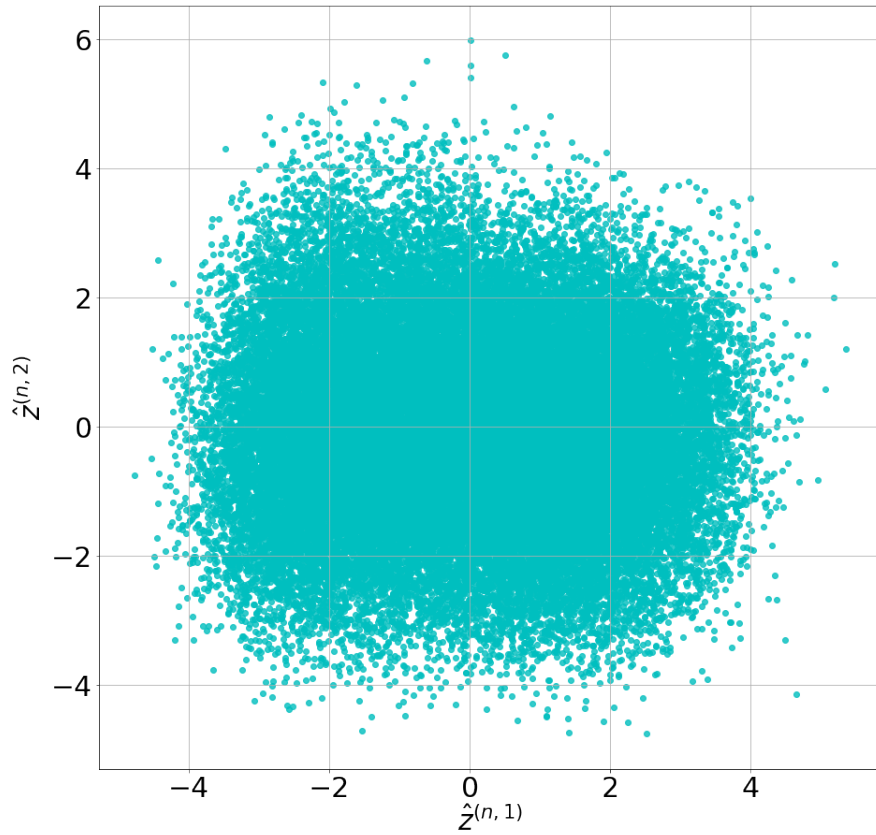


図 4.4 PC1, PC2 の分布（潜在変数：1000 次元）

### 4.3 潜在空間の重心とのノルム

適切に VAE によって入力画像の分布が学習されていれば，分布の重心に近いものは正常な画像が多く分布していると考えられ，重心から離れたものは正常でない画像が多く分布していると考えられる．そこで，式 4.1 によって示される潜在変数の重心  $z_c$  からのノルム  $\delta$  を各変数に対して求め，その分布を確認した．

$$\delta^{(n)} = \|z^{(n)} - z_c\| \quad (4.1)$$

図 4.5 に 500 次元の潜在変数から算出された  $\delta^{(n)}$ ，図 4.6 に 1000 次元の潜在変数から算出された  $\delta^{(n)}$  を示す．

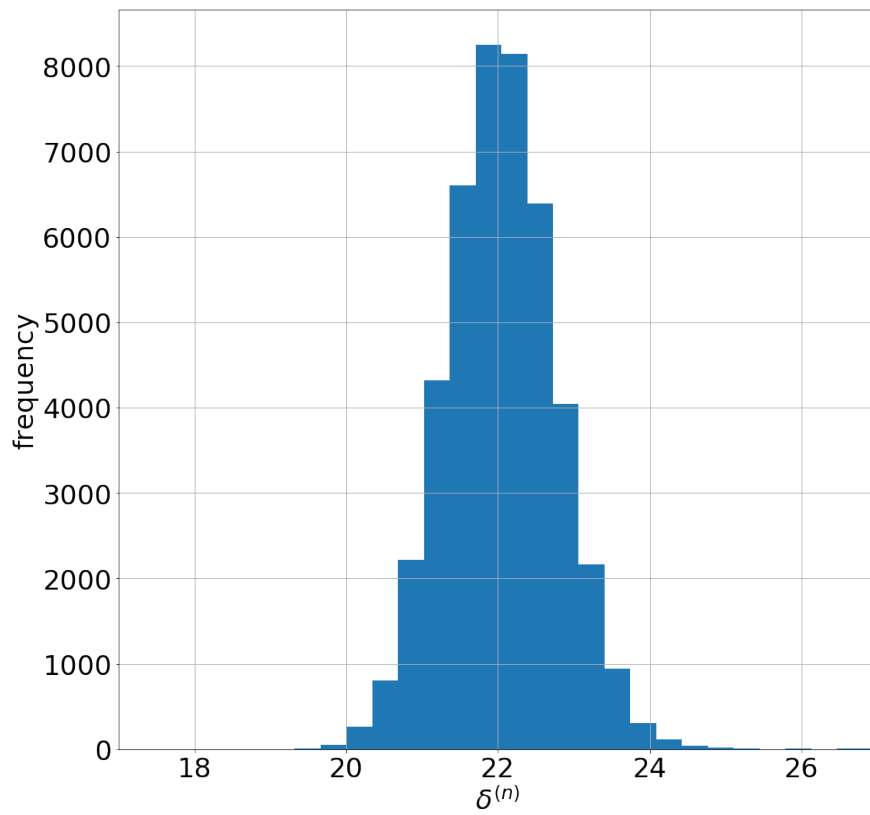


図 4.5 500 次元の潜在変数から算出された  $\delta^{(n)}$



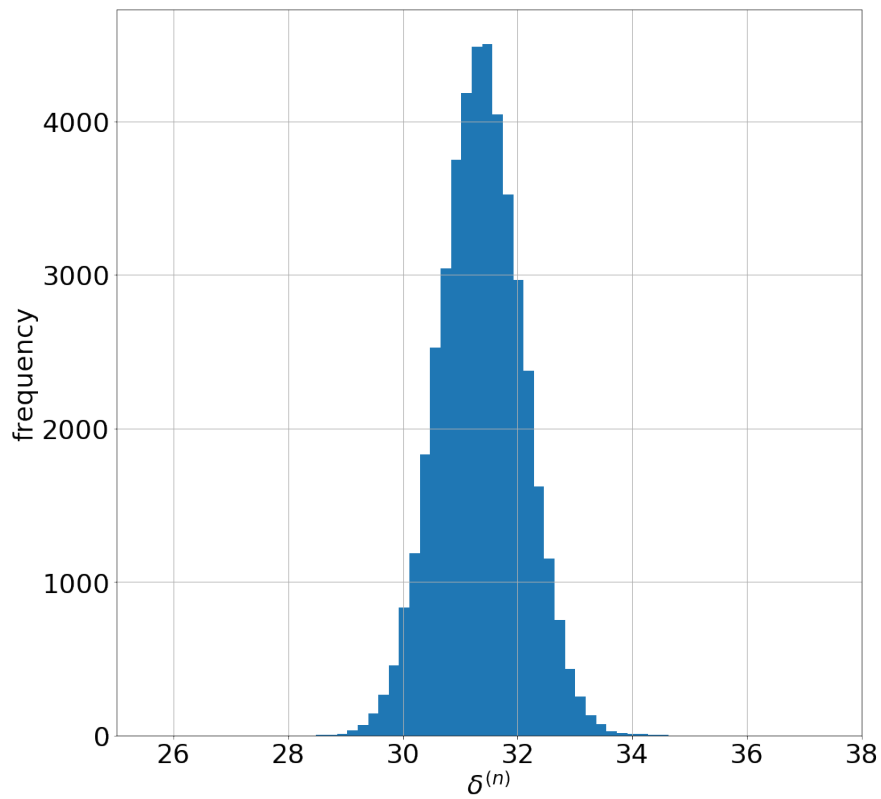


図 4.6 1000 次元の潜在変数から算出された  $\delta^{(n)}$

#### 4.4 各主成分の分布

500 次元の潜在変数の第 1 主成分, 第 2 主成分, 第 3 主成分を図 4.7 – 図 4.9, 1000 次元の潜在変数の第 1 主成分, 第 2 主成分, 第 3 主成分を図 4.10 – 図 4.12 に示す.

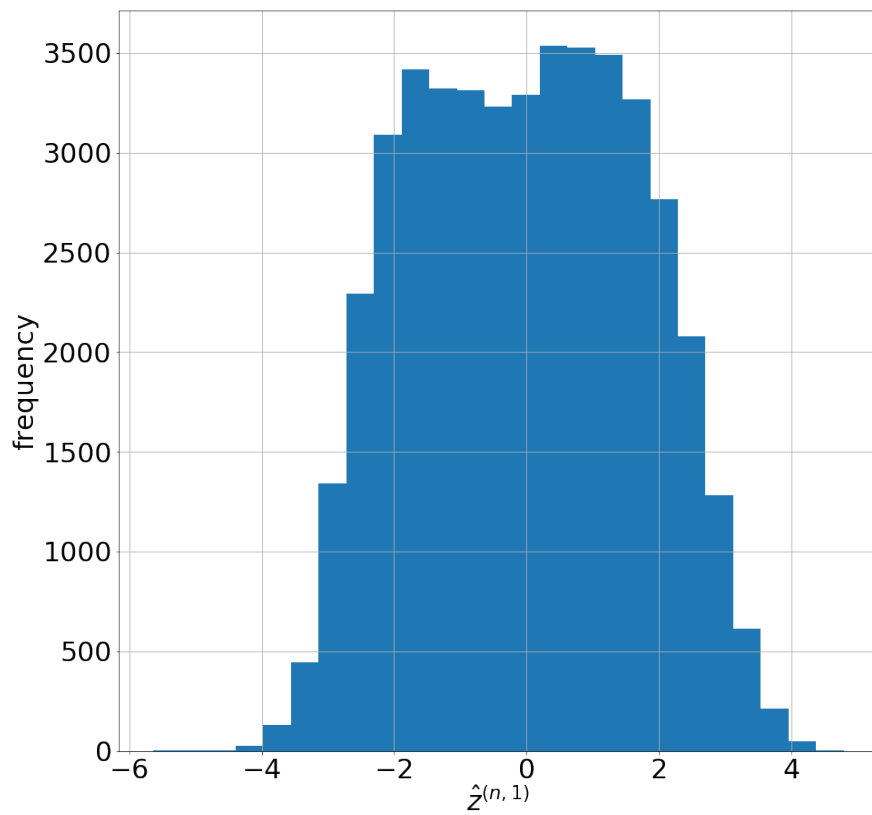


図 4.7 第 1 主成分の分布 (潜在変数 : 500 次元)

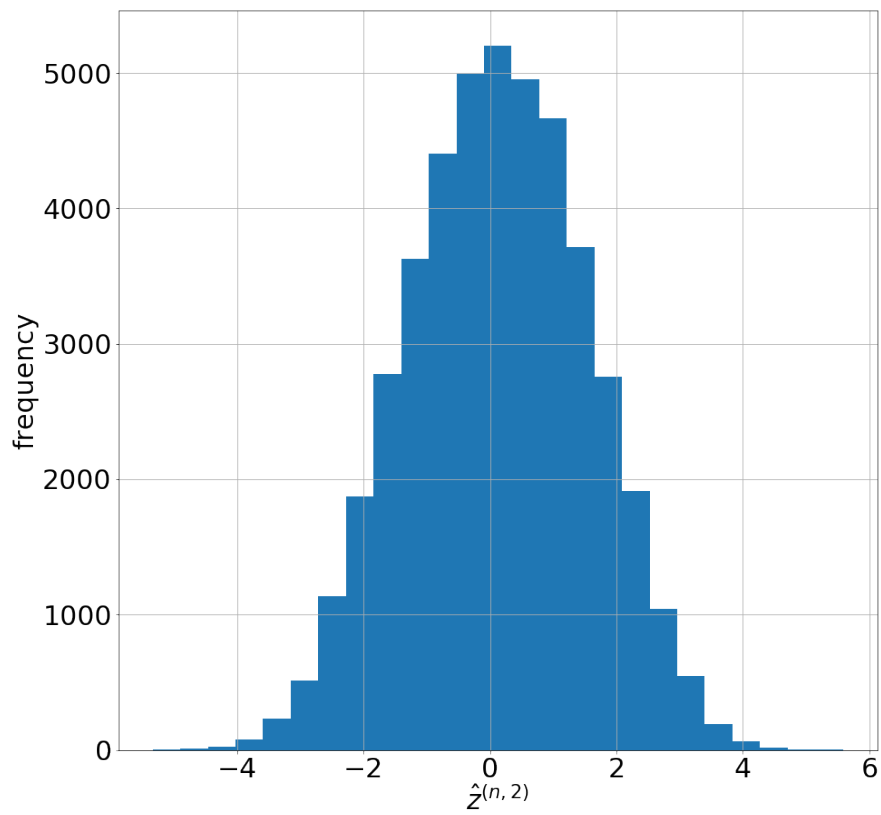


図 4.8 第 2 主成分の分布 (潜在変数 : 500 次元)

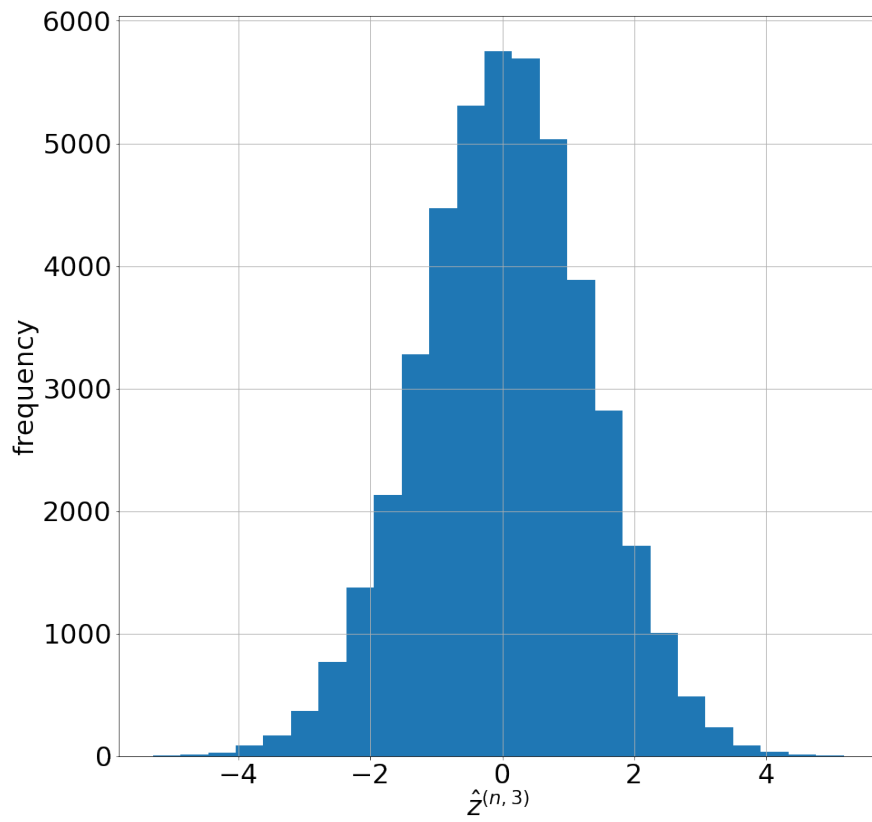


図 4.9 第 3 主成分の分布 (潜在変数 : 500 次元)

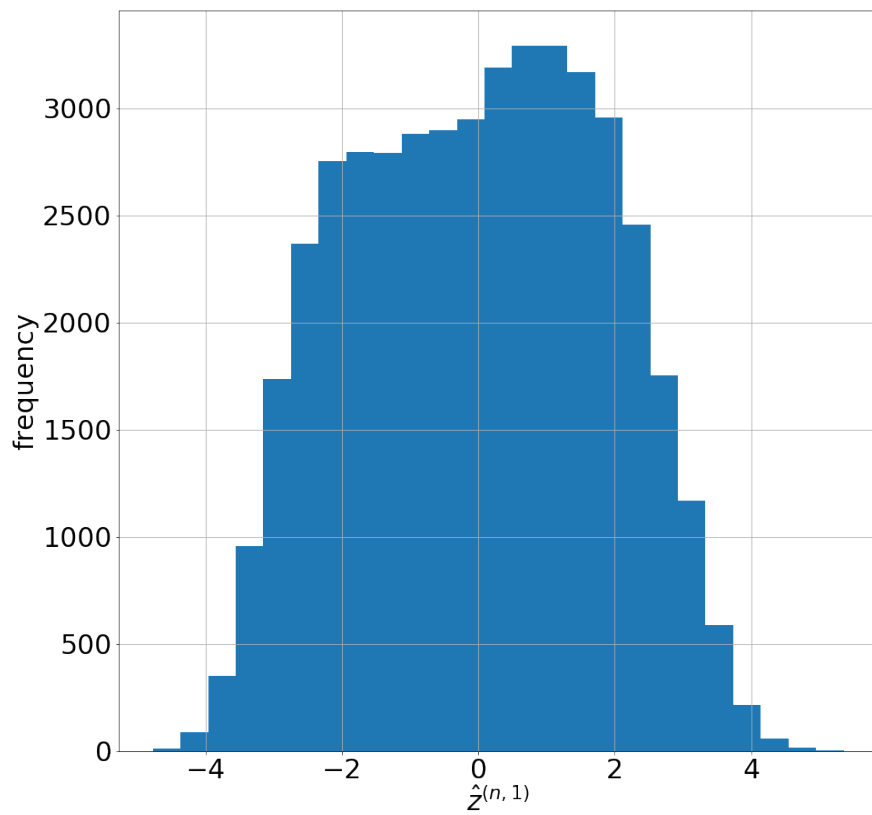


図 4.10 第 1 主成分の分布 (潜在変数 : 1000 次元)

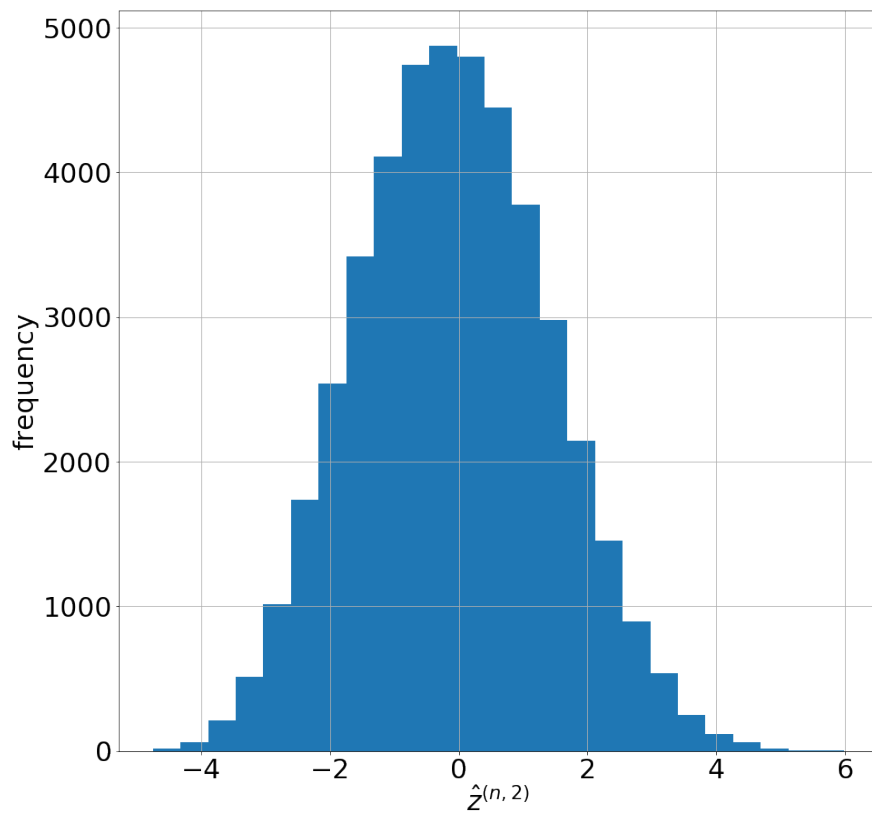


図 4.11 第 2 主成分の分布 (潜在変数 : 1000 次元)

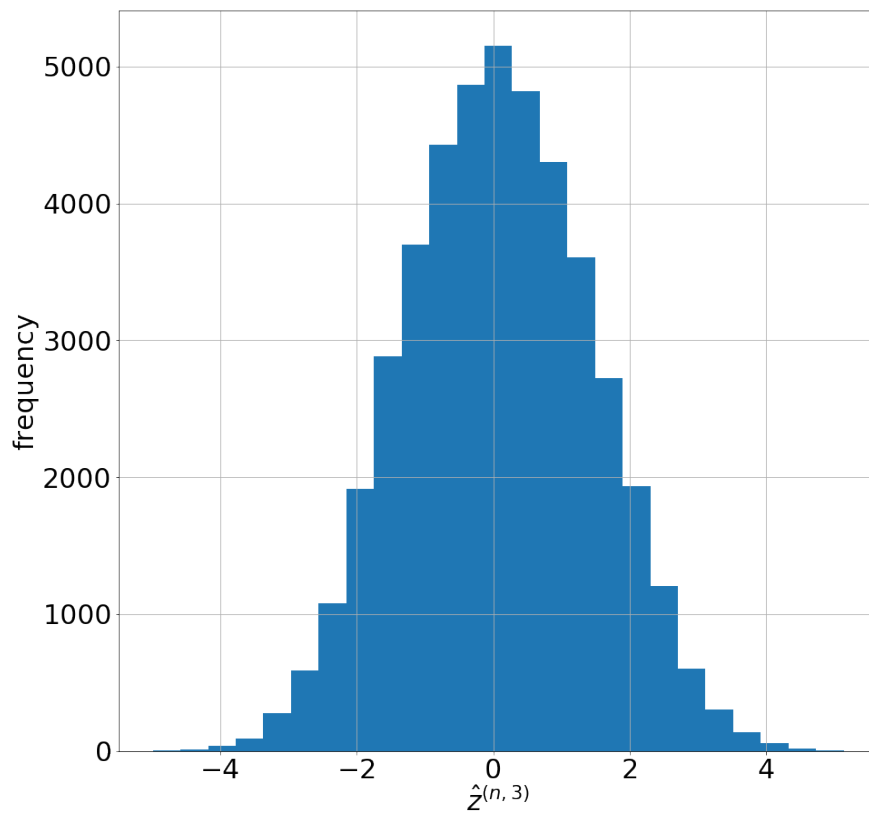


図 4.12 第 3 主成分の分布 (潜在変数 : 1000 次元)

#### 4.5 主成分の重心との距離

第  $m$  ( $m = 1, 2, 3$ ) 成分の重心を  $\hat{z}_c^m$  として, 式 4.2 によって算出される重心からの距離の分布を確認した.

$$d^{(n,m)} = |\hat{z}^{(n,m)} - \hat{z}_c^m| \quad (4.2)$$

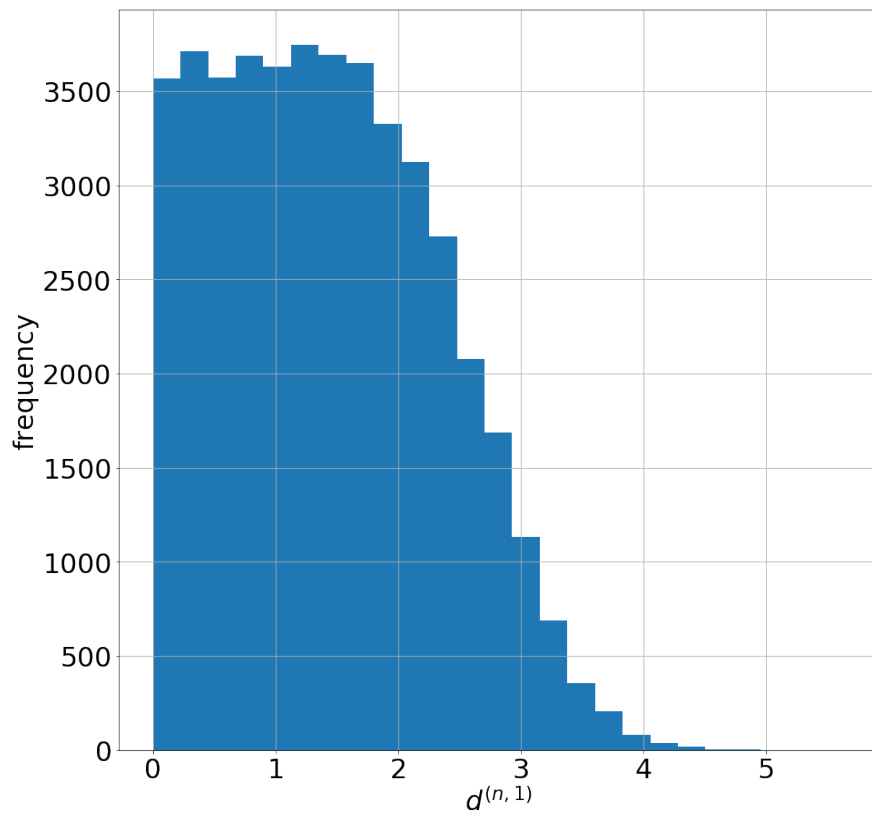


图 4.13  $d^{(n,1)}$ (潜在変数 : 500 次元)



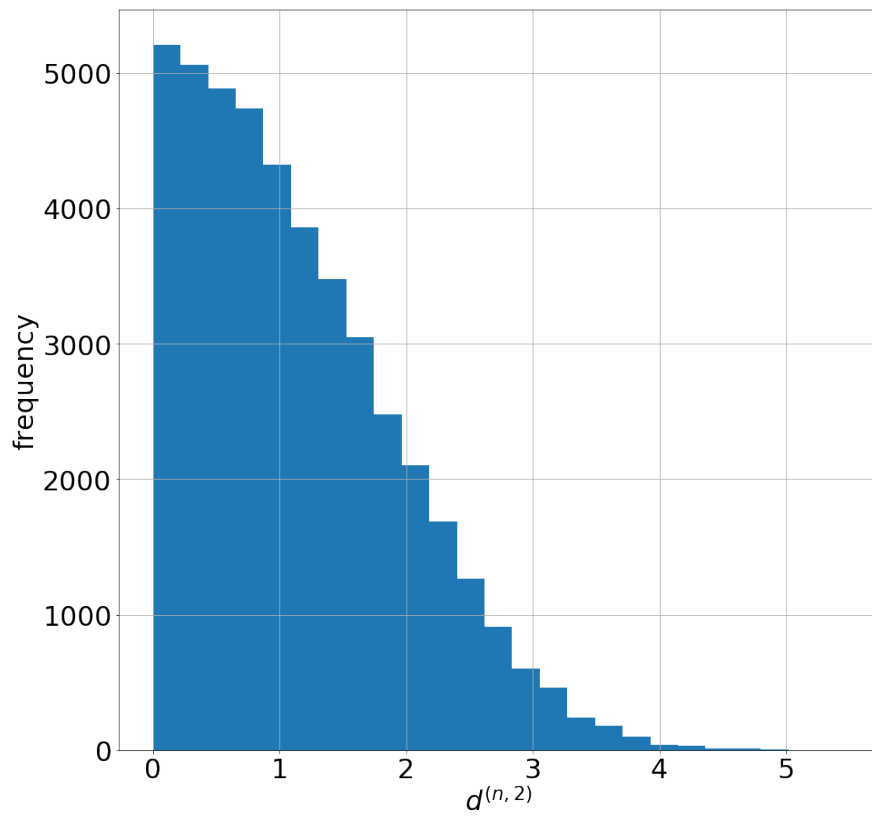


图 4.14  $d^{(n,2)}$ (潜在変数 : 500 次元)

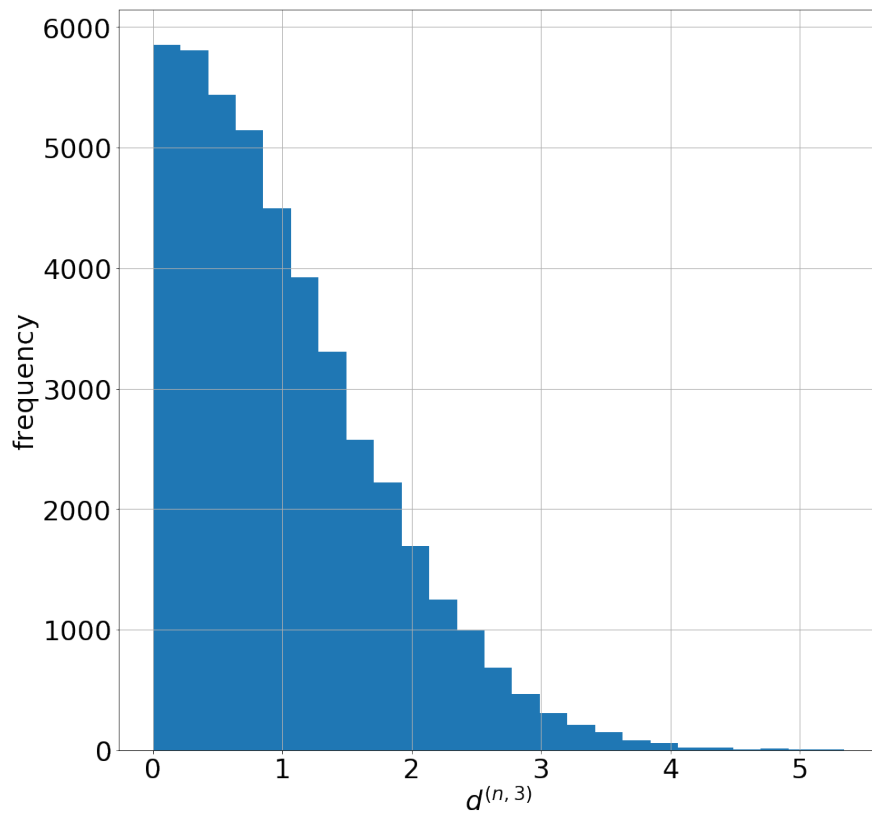


图 4.15  $d^{(n,3)}$ (潜在変数 : 500 次元)

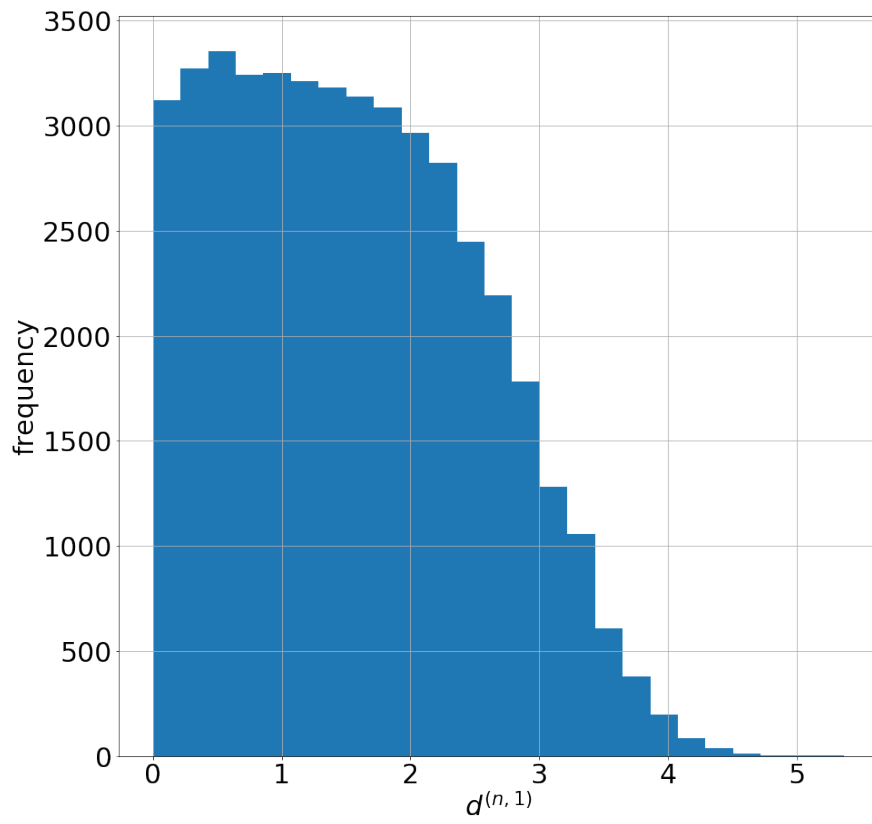


图 4.16  $d^{(n,1)}$ (潜在变数: 1000 次元)

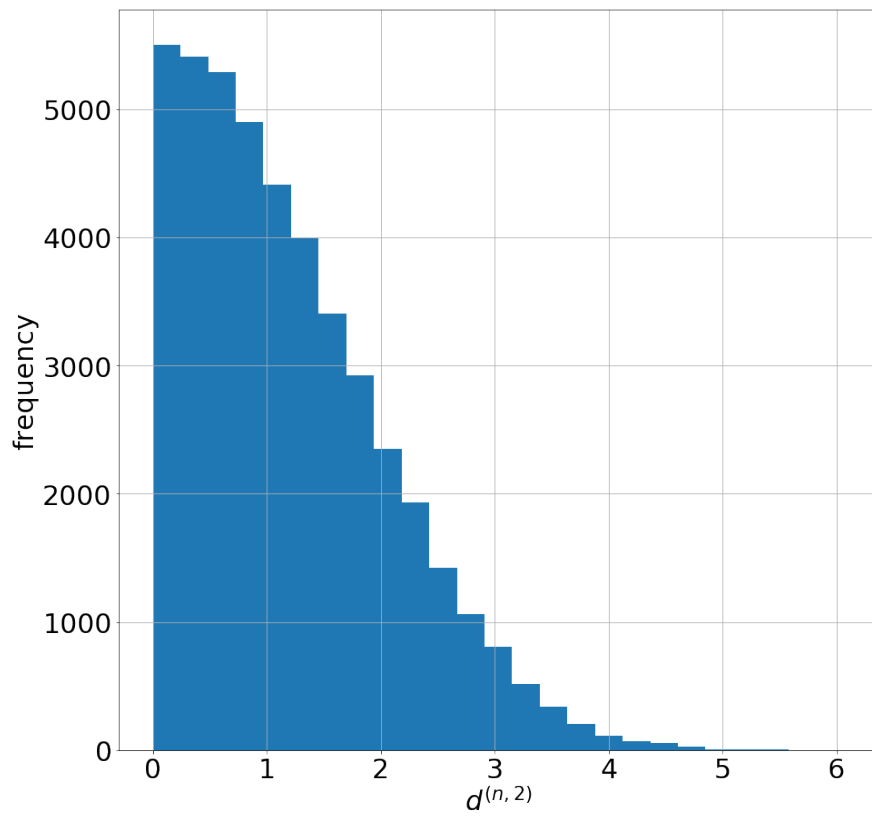


图 4.17  $d^{(n,2)}$  (潜在变数: 1000 次元)

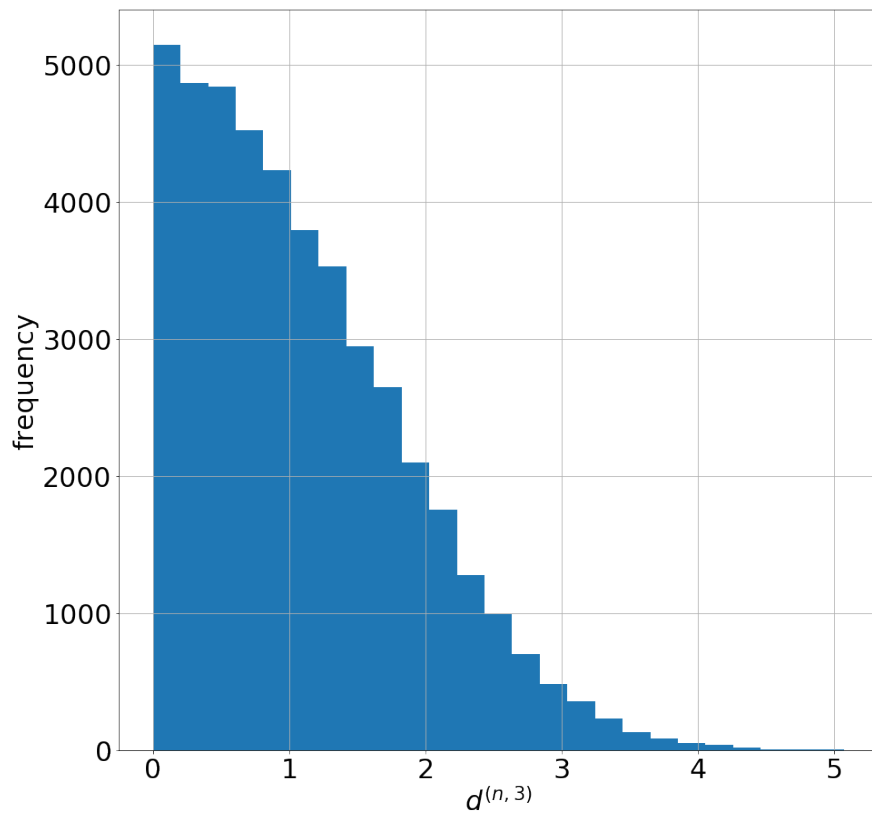


图 4.18  $d^{(n,3)}$  (潜在变数: 1000 次元)

## 4.6 分布の平均から離れたデータ

潜在変数の空間において基本的に多次元正規分布に近いサンプルの分布が見られたということは、分布の周辺にあるサンプルはなんらかの形で平均から大きく外れた画像であることを示唆していると考えられる。潜在変数の重心からの距離およびPCAで求めた主成分軸において、平均から最も離れたサンプルがどのような画像であるかを解析し、潜在変数空間における分布の構造と画像の特徴との関係を解析した。

ここでは、分布の平均から離れた画像をそれぞれ5枚ずつ示す。

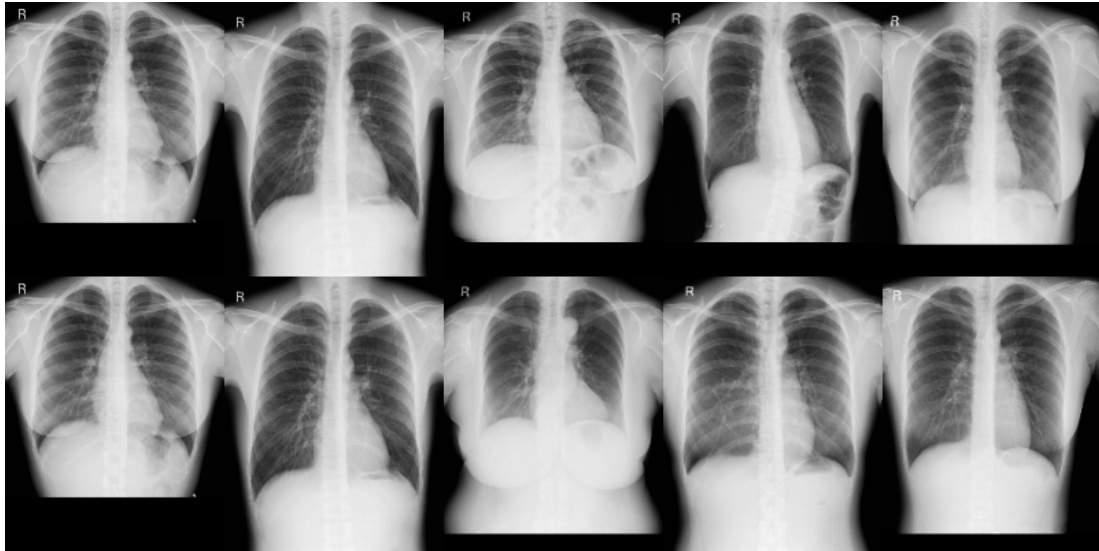


図 4.19  $d^{(n)}$  の値が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

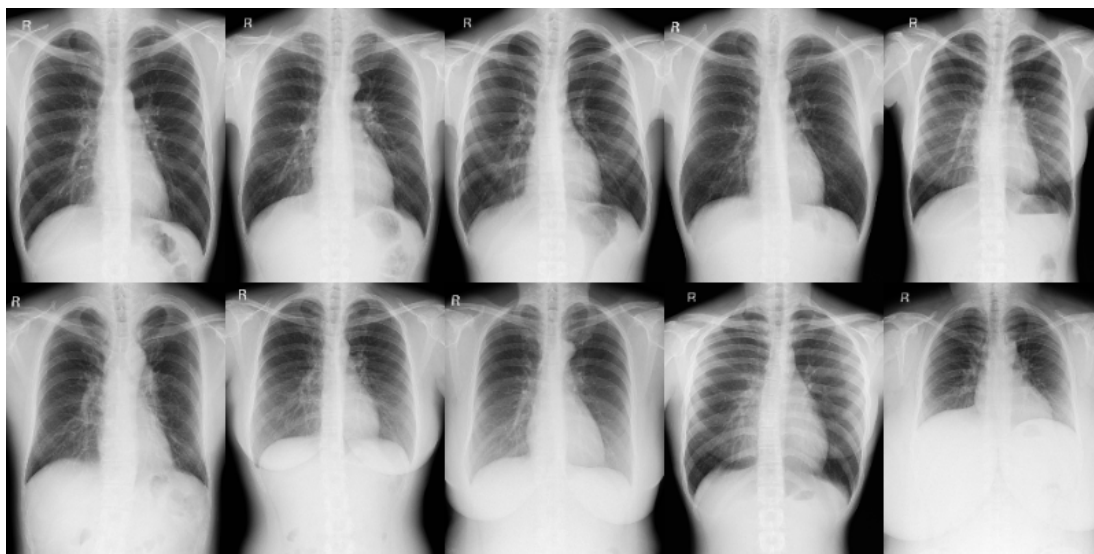


図 4.20 第1主成分の値が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

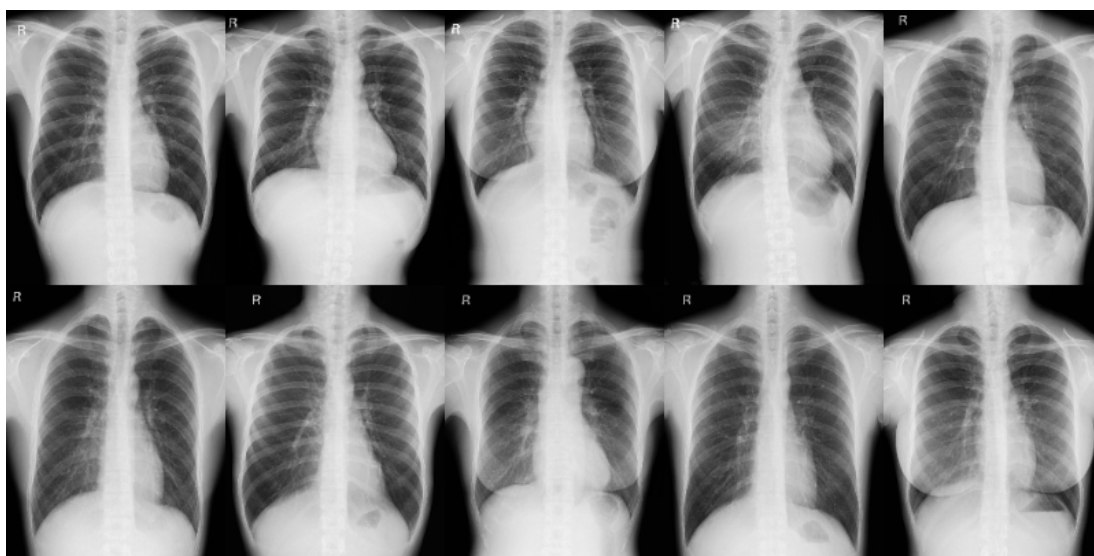


図 4.21 第2主成分の値が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

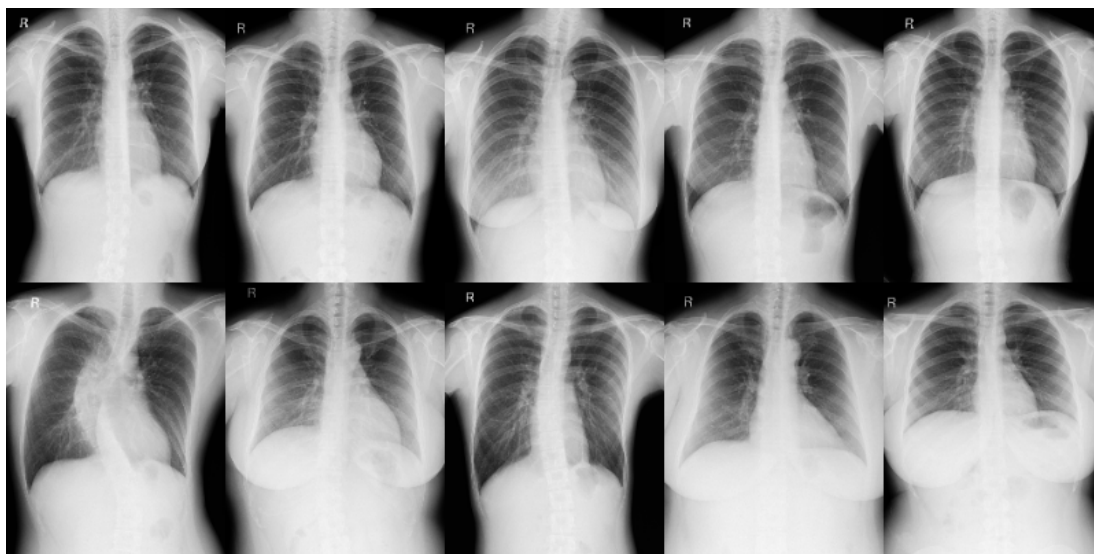


図 4.22 第3主成分の値が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

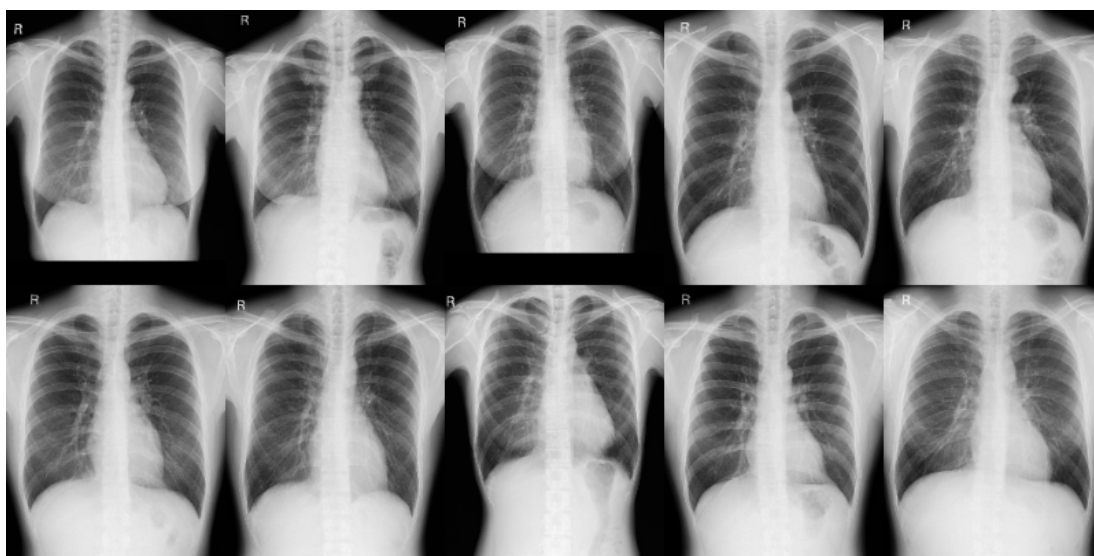


図 4.23  $d^{(n,1)}$  が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）



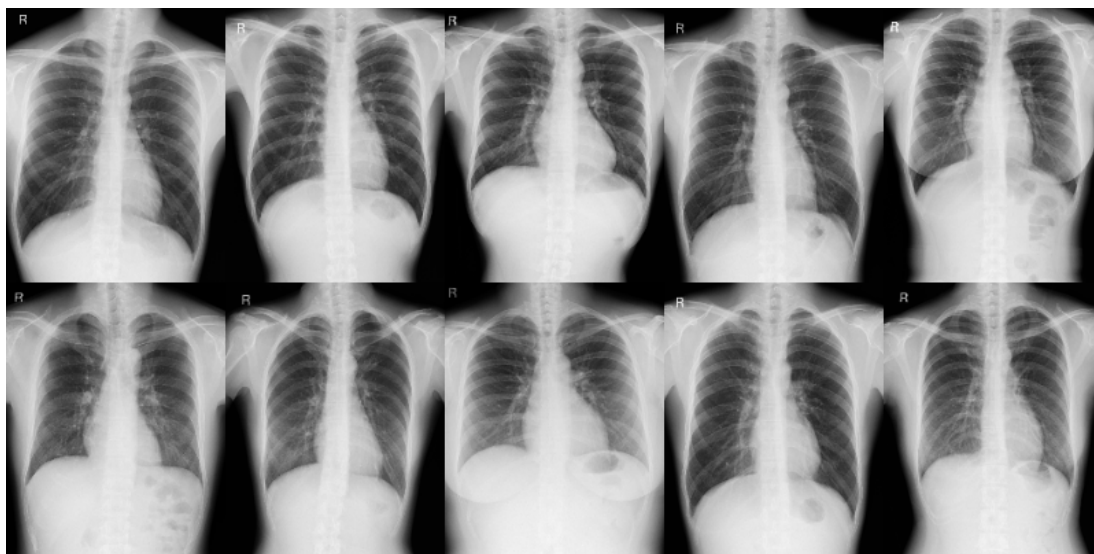


図 4.24  $d^{(n,2)}$  が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

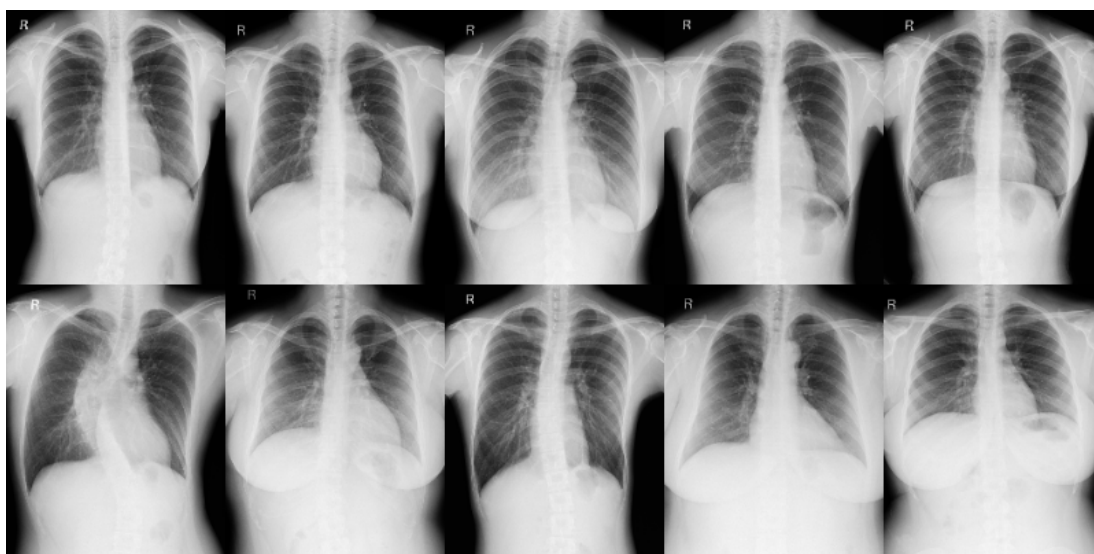


図 4.25  $d^{(n,3)}$  が大きい画像（上：潜在変数；500次元，下：潜在変数；1000次元）

ここで、注目すべきは  $d^{(n)}$  の大きい5枚は形状が近いものであるということである。これは、VAEによって肺の形状や大きさ等の大域的な特徴が抽出された結果であると考えられる。

#### 4.6.1 潜在変数空間と診断の関係

次に、潜在空間の重心からの距離および主成分の値が大きい100枚についてフィッシャーの正確確率検定によって診断結果との関係を調べた。なお、入力画像として用いた44725枚の画像のうち、正常でないものが4513枚あった。

表 4.1 異常性が確認された画像の枚数と p 値（潜在変数：500 次元）

	異常性が確認された画像の枚数	p 値
$\delta^{(n)}$	18	0.02
$\hat{z}^{(n,1)}$	14	0.19
$\hat{z}^{(n,2)}$	10	0.65
$\hat{z}^{(n,3)}$	10	0.65
$d^{(n,1)}$	10	0.65
$d^{(n,2)}$	5	0.99
$d^{(n,3)}$	12	0.40

表 4.2 異常性が確認された画像の枚数と p 値（潜在変数：1000 次元）

	異常性が確認された画像の枚数	p 値
$\delta^{(n)}$	16	0.07
$\hat{z}^{(n,1)}$	17	0.04
$\hat{z}^{(n,2)}$	11	0.52
$\hat{z}^{(n,3)}$	22	0.001
$d^{(n,1)}$	13	0.28
$d^{(n,2)}$	13	0.28
$d^{(n,3)}$	10	0.40

## 5. 考察

### 5.1 第1主成分について

主成分分析は入力データ空間の分散が最も大きい方向に第1主成分を設定する。よって、第1主成分は各入力画像が普遍的に持つ特徴（肺の大きさや形状等）が反映されたものであると考えられる。そこで、この仮説を検証するために、男女別の第1主成分を確認した。その分布を図5.1, 図5.3に示す。図4.7, 図4.10と図5.1, 図5.3をそれぞれ比較することにより、各図における2つの頂点がほぼ一致していることが確認できる。このことから、第1主成分は画像全体の大きさや形状等を反映した特徴量であることが考えられる。

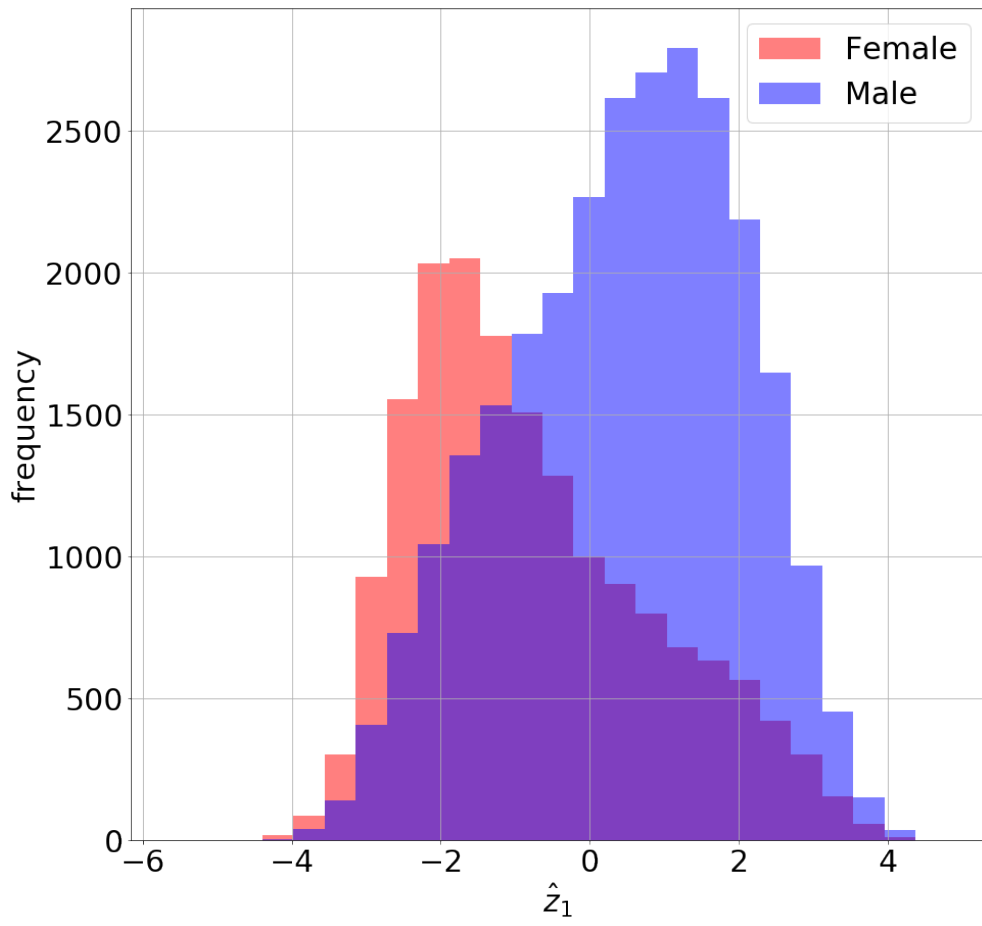


図 5.1 男女別の第 1 主成分（潜在変数:500 次元）

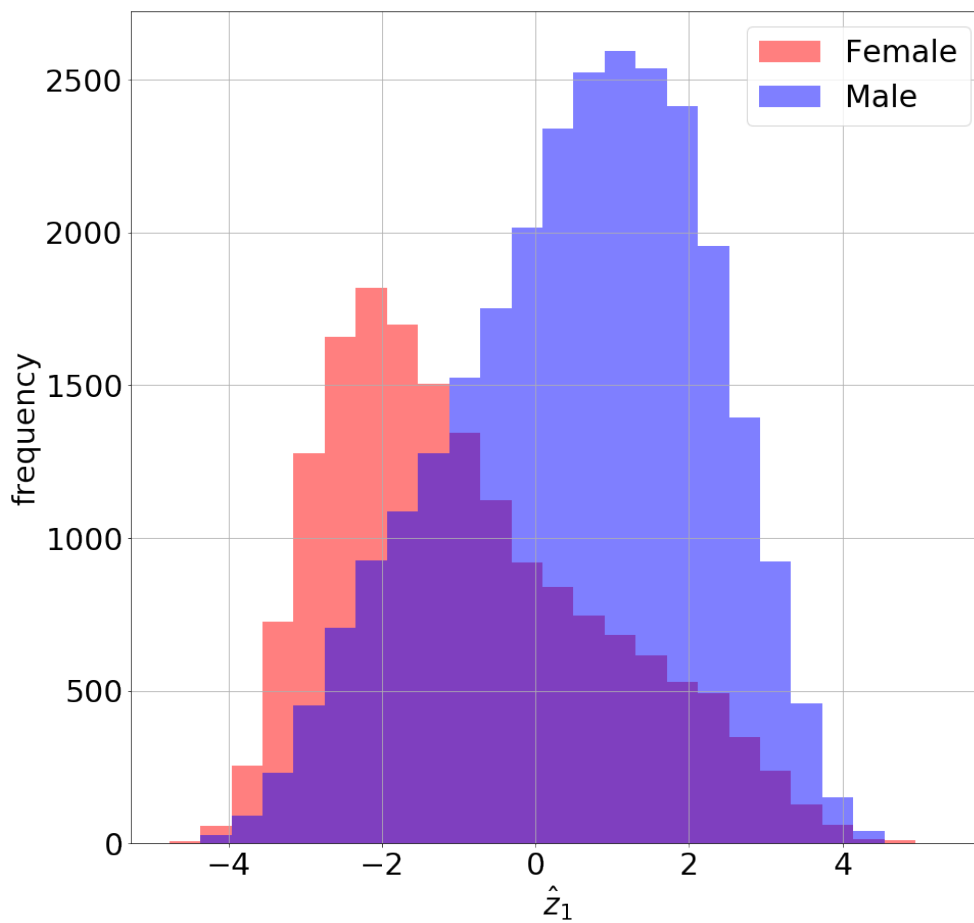


図 5.2 第 1 主成分

図 5.3 男女別の第 1 主成分 (潜在変数:1000 次元)

## 5.2 1000 次元潜在変数の第 3 主成分

本研究において，異常部位が確認される画像を最も抽出できた指標は，1000 次元潜在変数の第 3 主成分であった．本節では，これについての考察を行う．

## 5.3 抽出された画像について

抽出された画像の症例とその件数を表 5.1 に示す．

表 5.1 抽出された画像の症例とその件数

症例	件数
側弯	5
胸膜肥厚	4
吸気不十分	3
結節影	3
大動脈拡張	1
肺動脈拡張	1
心陰影の拡大	1
斑状影	1
肋骨の奇形・変形	1
腫瘤影	1
術後	1

### 5.3.1 背骨に関する所見

1000次元潜在変数の第3主成分によって、抽出された画像の所見で最も見受けられた所見は「側弯」である。側弯とは、背骨が左右に曲がってしまっている状態を指す。第3主成分において異常が確認された22枚の画像のうち5枚が側弯と診断されたものであった(図5.4)。また、同様に、背骨の特徴によって抽出されたと考えられる画像の所見として「術後」である。実際に、画像を見ると背骨の部分に、装具がついていることを確認できた(図5.5)。このことから、第3主成分においては、画像全体の形状や大きさだけでなく、画像中の中央に位置する背骨の形状を反映していると考えられる。

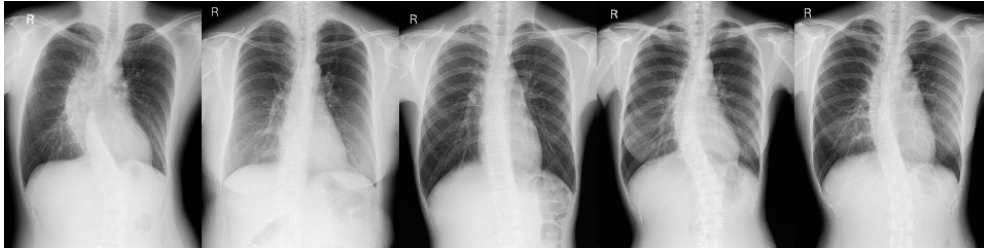


図 5.4 第 3 主成分（潜在変数：1000 次元）において側弯が確認された画像

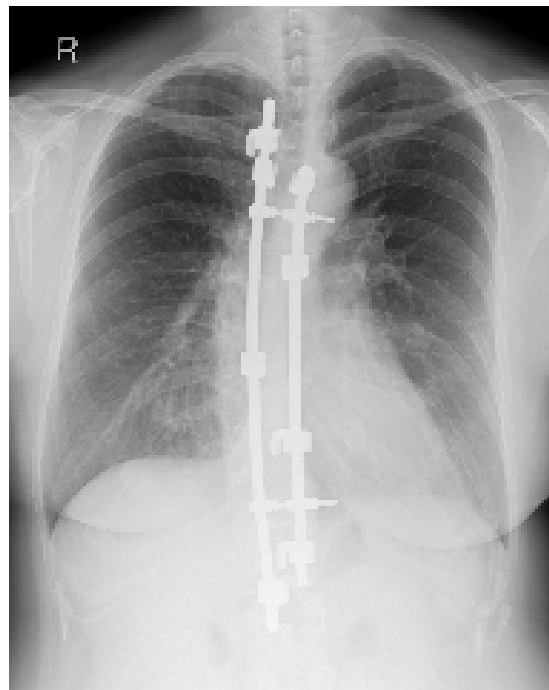


図 5.5 第 3 主成分（潜在変数：1000 次元）において側弯が確認された画像

## 6. 総括

本研究では、VAEによって抽出された特徴量をもとに、潜在空間における画像データセットの分布を確認した。また、その特徴量をPCAによって次元圧縮したものと実際の診断結果との対応をフィッシャーの正確確率検定によって確認した。VAEにおいては、潜在変数の次元が500次元の場合、1000次元の場合のいずれにしても同等のレベルでの再構成ができた。よって、より次元数を減らすことでより低次元の特徴量を得られると考えられる。また、畳み込みフィルタの数を増やすなど、ネットワークの構造を変更すれば、より詳細部分の再構成ができると考えられる。

また、500次元の潜在変数から算出された $\delta^{(n)}$ や1000次元の潜在変数の第3主成分から診断結果との有意な対応関係が確認できた。第1主成分や第2主成分から対応関係が確認されず、第3主成分から確認されたのは、第1主成分や第2主成分は肺の形状や大きさなどの画像の特徴としてばらつきの大きい大局的な特徴を表していて、第3主成分は背骨等の詳細部分の特徴を反映した指標となっているためであると考えられる。今回は、連続的な分布が得られるVAEを用いたが、他のモデルにおける中間層を解析すると、また異なる結果が得られる可能性がある。



## 謝辞

本研究は奈良先端科学技術大学院大学 計算システムズ生物学研究室のご指導，ご協力を頂き遂行することができました．本研究を行うにあたり研究のみならず，前向きな思考の重要性をご教授して下さった計算システムズ生物学研究室 金谷 重彦 教授に深く感謝申し上げます．

また，中間発表や修士論文作成において的確なご指摘をくださいました奈良先端科学技術大学院大学 生体医用画像研究室 佐藤 嘉伸 教授に深く感謝申し上げます．

また，深層学習の基礎知識や研究の方針やモデルの検討等を昼夜問わず，熱心に指導して下さった奈良先端科学技術大学院大学 計算システムズ生物学研究室 小野 直亮 准教授に深く感謝申し上げます．小野先生のご尽力のおかげで，多くの知見を得るとともに，研究を遂行することができました．

また，研究室のゼミ等で，様々なご意見をくださいました奈良先端科学技術大学院大学 計算システムズ生物学研究室 Md. Altaf-Ul-Amin 准教授，計算システムズ生物学研究室 黄 銘 助教に深く感謝申し上げます．

また，本研究で使用したデータセットをご提供くださった東京大学 本郷保健センター 袴田 千愛 様に深く感謝いたします．

また，研究に関することだけでなく，日々時間を共にする仲間として様々な面で支えていただきました計算システムズ生物学研究室に在籍されている皆様，および修了生の皆様に感謝いたします．

最後に，大学院修了までの24年間において数多の支援をして下さった家族に感謝し，ここに謝辞といたします．

## 参考文献

- [1] Matthew D. Zeiler and Rob Fergus. Visualizing and Understanding Convolutional Networks. *European Conference on Computer Vision*, pp. 818–833, 2014.
- [2] OECD. Health at Glance 2019, 2019.
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks Alex. *Neural Information Processing Systems*, 2012.
- [4] David G Low. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, pp. 91–110, 2004.
- [5] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 2008.
- [6] Navneet Dalal, Bill Triggs, Navneet Dalal, and Bill Triggs. Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 886–893, 2005.
- [7] Ren Shaoqing, He Kaiming, Girshick Ross, and Sun Jian. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 1137–1149, 2015.
- [8] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis*, Vol. 39, No. 12, pp. 2481–2495, 2017.
- [9] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Genera-

- tive Adversarial Nets. *Advances in Neural Information Processing Systems*, Vol. 3, pp. 2672–2680, 2014.
- [10] Ivo M. Baltruschat, Hannes Nickisch, Michael Grass, Tobias Knopp, and Axel Saalbach. Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. *Science*, pp. 1–10, 2019.
- [11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Symposium of BIOMEDICAL IMAGING*, Vol. 9351, pp. 234–241, 2015.
- [12] Tatsuya Yokota, Hidekata Hontani, Qibin Zhao, and Andrzej Cichocki. Manifold Modeling in Embedded Space: A Perspective for Interpreting ”Deep Image Prior”. *arXiv:1908.02995*, pp. 1–11, 2019.
- [13] G. E. Hinton and R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks. *Science*, Vol. 313, pp. 504–507, 2006.
- [14] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. *The International Conference on Learning Representations*, 2014.