

敵対的生成ネットワークを用いた、 3次元点群形状特徴量の教師なし学習

上西和樹^{†1} 古屋貴彦^{†1} 大淵竜太郎^{†1}

概要：3次元形状データを効率的に管理するため、形の類似性に基づいてこれら3次元形状を比較、検索、分類する手法への要求が高まっている。近年提案された3次元形状を処理する深層ニューラルネットワーク(DNN)の多くは、3次元形状とこれに付与されたラベルの対応を教師付き学習することで、高精度な形状特徴量を獲得する。しかし、人手によるラベル付けには手間がかかるため、実際の3次元形状データの多くはラベルを持たない。このため、ラベルなしの3次元形状データから高精度な形状特徴量を獲得する手法が望まれている。本研究では、教師なし学習の一種である敵対的生成ネットワーク(GAN)を用いた3次元形状特徴量を提案する。提案する点群GAN法は、3次元点群を生成するDNN(生成DNN)と、3次元点群の本物らしさを予測するDNN(判別DNN)とを互いに競わせながら訓練する過程で、3次元点群の形状特徴を捉える能力を判別DNNに獲得させる。3次元モデルの形状類似検索のシナリオで評価した結果、判別DNNが獲得した形状特徴量は、既存の手作り形状特徴量よりも高精度に3次元形状を比較できることが分かった。

キーワード：3次元形状特徴量、教師なし学習、深層学習、敵対的生成ネットワーク、3次元形状類似検索

1. はじめに

近年、3次元形状モデル(3Dモデル)の数が急激に増加している。増加の背景には、仮想現実(VR)や拡張現実(AR)の技術の発展、安価な3Dスキャナや3Dプリンタの普及等がある。多数の3Dモデルを効率的に管理するため、形の類似性に基づいてこれら3次元形状を比較、検索、分類する手法への要求が高まっている。

一般的に、3Dモデル間の形状類似度はそれら3Dモデルから抽出された3D形状特徴量を比較することで計算される。形状特徴量は、人間により設計された「手作り形状特徴量」と3D形状データを学習して得た「学習形状特徴量」の2つに大別できる。研究者の経験と知識に基づいて設計された手作り形状特徴量(例えば、[9][10][11][12])は、一定の形状比較精度を示すものの、必ずしも形状比較に最適であるとは限らない。一方で、3Dモデルの形を解析して得た学習形状特徴量は、手作り形状特徴量よりもそれら3Dモデルの形状比較に適応していると期待できる。

学習形状特徴量を得るアプローチには、教師あり学習と教師なし学習がある。近年提案された高精度な学習形状特徴量の多くは教師あり学習を用いる([3][5][13][14]など)。これら既存研究では、3Dモデル形状特徴量抽出向けに設計した深層ニューラルネットワーク(DNN)をできるだけ多くのラベル付き3D形状データで訓練する。しかし、データベースに日々蓄積される大量の3Dモデルの殆どはラベルを持たない。人手による3Dモデルへのラベル付与の手間が大きいためである。そのため、教師あり学習に使える3Dモデルの数は限られてしまう。一方で、教師なし学習を用いれば、ラベル付与の手間なしに多量の3Dモデルを学習できる。しかし、3D形状特徴量の教師なし学習はこれま

でほとんど研究されていない。これは、教師なし学習で3D形状特徴量を獲得するのが困難なためである。

本研究では、ラベルを持たない3Dモデル群から高精度な学習形状特徴量を獲得することをねらう。そのために、教師なし学習の一種である敵対的生成ネットワーク(GAN)を用いる。対象とする3次元形状表現は3D点群である。3D点群は汎用性が高く、また近年、距離画像センサの普及等により用途が拡大している。

本研究で提案する点群GANは、3D点群を生成するDNN(生成DNN)と、3次元点群の真贋を判別するDNN(判別DNN)の、2つのDNNで構成される。生成DNNは、入力されたランダムな潜在ベクトルを3D点群に変換する。判別DNNは、入力された3D点群について、それが人手で製作された“本物”であるか、生成DNNが出力した“偽物”であるかを判別する。生成DNNと判別DNNが互いに敵対して学習することで、生成DNNは本物らしい3D点群の生成を学習し、また判別DNNは3D点群の判別能力、ひいては形状特徴を捉える能力を学習する。

本来のGAN[7]は2次元画像の生成および判別のためのDNN構造を持ち、3D点群の生成および判別に用いることはできない。そこで本研究では、3D点群を処理するための点群GANを新たに設計する。点群GANの判別DNNが獲得する形状特徴量の要件の1つは、3D点群を構成する点の順序に対する不変性である。そこで我々は、点の順序変化に影響を受けない3D点群解析用DNNを用いて点群GANを構築する。具体的には、QiらのPointNet[3]を判別DNNに、またYangらのFoldingNet[4]を生成DNNに用いる。

また、複数の研究([7][8]など)において、GANの学習は不安定であることが示されている。点群GANの学習を安定化させるため、我々は、Gulrajaniら[1]によって提案されたGANの学習則を用いる。具体的には、3D点群データ

^{†1} 山梨大学
University of Yamanashi

個々の真贋 [7]の代わりに、本物の点群データ集合と偽物の点群データ集合の集合間距離を損失として利用する。

点群 GAN の学習の後、判別 DNN を 3D 点群の特徴抽出器として用いる。判別 DNN の中間層から抽出された形状特徴量を、3D 形状の比較、検索、分類等に用いる。

提案手法の点群 GAN が獲得した形状特徴量を、3D モデルの形状類似検索のシナリオで実験的に評価する。実験の結果、提案する点群 GAN により教師なし学習で得た学習形状特徴量は、既存の手作り形状特徴量を上回る検索精度を示すことが分かった。また、点の順序変化に影響を受けない 3D 点群解析用 DNN ([3][4]) の採用、および、GAN の安定な学習則 [1]の利用が、それぞれ形状特徴量の精度の改善に効果的であることを確かめた。

2. 関連研究

2.1 DNN を用いた 3 次元形状特徴量の学習

DNN を用いた学習形状特徴量の多く（例えば、[3][5][13][14]）は、3次元形状とこれに付与されたラベルを教師データとした教師あり学習により獲得される。これまでに、ボクセル表現、3D 点群、多視点見かけ画像等の形状表現を対象とした学習形状特徴量が提案された。Wu らが提案した ShapeNet [5] はボクセル表現を対象とし、3次元の畳み込みニューラルネットワーク（CNN）を用いて 3D 形状特徴量を抽出する。Su らが提案した MV-CNN [13]は、3D モデルの多視点の見かけ画像を対象とする。MV-CNN は 2D の CNN を用いて見かけ画像の各々から画像特徴量を抽出し、全視点の画像特徴量を 3D モデル当たり 1 つの特徴量に統合することで 3D 形状特徴量を得る。

本研究と同様、3D 点群を対象とした研究には[3]、[4]、[14]などがある。Furuya らが提案した DLAN [14]は、3D モ

デルの局所領域から抽出した多数の手作り幾何特徴量を DNN で精製、統合することで 3D モデルの形状特徴量を得る。Qi らが提案した PointNet [3]は、DNN を用いて 3D 点群の各点を高次元の特徴量に変換し、それらを統合した後、カテゴリ分類や領域分割を行う。PointNet は点群データを直接入力でき、かつ、点の順序変化に影響を受けない利点がある。[3]、[14]はいずれも、DNN を効果的に訓練するために多数のラベル付き 3D 点群データが必須となる。

近年、Yang ら [4]は 3D 点群を対象とした教師なし形状特徴量学習法 FoldingNet を提案した。FoldingNet は 3D 点群向けに設計された Autoencoder であり、PointNet 同様、点の順序変化に対する不変性を持つ。FoldingNet では、3D 点群の符号化器として PointNet に類似した DNN が用いられる。3次元点群の復号化では、符号化により得た形状特徴量を手がかりに、格子状配置を持つ 2D 点群を「折り畳む」ことで 3D 形状を復元する。Yang らは、2D 格子点群の折畳みによる復号化が高精度な形状特徴量の学習に効果的であることを示した。

本研究で提案する点群 GAN は、2D 格子点群の折畳みによる 3D 点群生成を行うことで、生成される点群の質を改善し、特徴量学習の効果を高める。

2.2 敵対的生成ネットワーク

敵対的生成ネットワーク（GAN）[7]は本来、2D 画像の生成を学習する DNN として提案された。GAN は生成 DNN と判別 DNN から成る。生成 DNN はランダムな潜在ベクトルを入力として受け取り、2D 画像を出力する。一方で判別 DNN は、2D 画像を入力として受け取り、その画像が本物である確率を出力する。GAN の学習では、生成 DNN と判別 DNN が敵対する。即ち、判別 DNN は画像が本物であるか、生成 DNN が出力した偽物であるかを判別するように

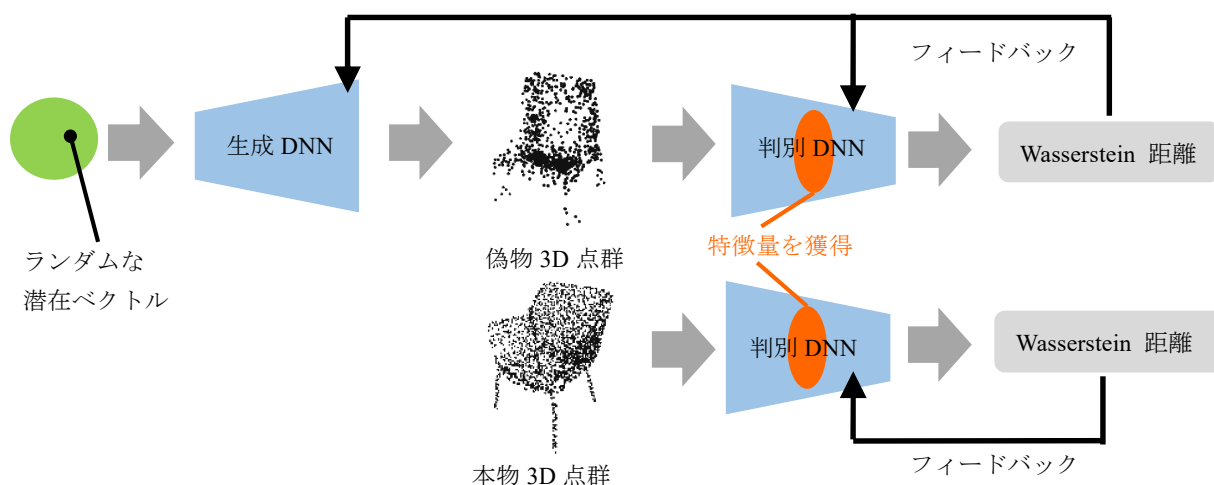


図 1 点群 GAN を用いた教師なし形状特徴量学習。潜在ベクトルを 3D 点群に変換する生成 DNN と、3D 点群の本物らしさを測る判別 DNN とが敵対的に学習する過程で、判別 DNN が 3D 点群の形状特徴量を獲得する。

訓練される。対して生成 DNN は、判別 DNN が誤判別するような「本物そっくりの偽物」を生成するように訓練される。敵対的な学習の結果、判別 DNN が画像分類に有用な画像特徴量を獲得することが示されている[15]。

複数の研究 ([7][8]など)において、GAN の学習が不安定であることが示されている。例えば、生成 DNN と判別 DNN の学習の進捗のバランスが崩れると学習が進まなくなる。GAN の学習を安定化させるために、DNN 構造の改良[15]や正則化[8]等が提案された。Arjovsky ら[2] は主に損失関数の改良により学習を安定化した Wasserstein GAN (WGAN) を提案した。従来の GAN の生成 DNN が偽物の画像個々の本物らしさを高めるように学習するのにに対し、WGAN の生成 DNN は偽物の画像集合の本物らしさを高めるように学習する。画像集合の本物らしさの尺度として、本物の画像集合と偽物の画像集合との集合間距離 (Wasserstein 距離) が用いられる。Gulrajani ら [1]は、WGAN の学習をさらに安定化させた WGAN-GP 法を提案した。WGAN-GP では判別 DNN の勾配ノルムに対して制約を加えること (Gradient Penalty) で、より正確な Wasserstein 距離が得られるようになり、生成される画像の質が向上する。

本研究では、点群 GAN の学習に WGAN-GP 法を用いることで、学習の安定性を高める。

3. 提案手法

3.1 手法の概要

点群 GAN を用いた 3D 形状特徴量の教師なし学習法を提案する。図 1 に点群 GAN の概要を示す。点群 GAN は、3D 点群の生成を学習する DNN (生成 DNN) と、3D 点群の本物らしさを計測する DNN (判別 DNN) が敵対して学習する。敵対的な学習の後、判別 DNN の中間層から 3D 点群の形状特徴量を抽出し、この特徴量を 3D 形状の比較や検索に用いる。

点群 GAN が高精度な形状特徴量を学習するために、(1) 3D 点群を構成する点の順序変化に不変な DNN 構造、および (2) 学習が安定する損失関数が要求される。点の順序に対する不変性を得るために、PointNet [3]を判別 DNN として用い、また FoldingNet [4]の復号化器を生成 DNN として用いる。また、学習の安定性を高めるために、WGAN-GP 法 [1]に基づく損失関数を利用する。

3.2 ネットワーク構造

3.2.1 判別 DNN

図 2 に判別 DNN の構造を示す。判別 DNN は、入力された 3D 点群の各点を独立に全結合層へ入力し、点ごとの特徴量を得る。全結合層の数は 4 であり、各層のニューロン数は 64, 256, 512, 1024 である。次いで、点ごとの特徴量群

を最大値プーリングにより 3D 点群モデル当たり 1 つの特徴量に統合する。統合特徴量は後続の N 個の全結合層から成る特徴抽出部へ入力され、精製される。特徴抽出部の構造は、形状特徴量の精度に影響する重要な要素だと考えられる。そこで本研究では、特徴抽出部の構造が検索精度に与える影響を調査する。具体的には、特徴抽出部の層数を $N = 4, 6, 8$ とした場合に得られる形状特徴量の精度を比較する。この際、全結合層のニューロン数は、始めの $N-3$ 層が 1024, 最後の 3 層が 512, 64, 1 である。判別 DNN の全ての全結合層において、Leaky ReLU [16]を活性化関数として用いる。点群 GAN の学習時に限り、特徴抽出部の最終層から Wasserstein 距離が出力される。この Wasserstein 距離は、本物点群データの集合 (ミニバッチ) と偽物点群データの集合 (ミニバッチ) の間で計算された集合間距離である。Wasserstein 距離が小さいほど、偽物点群データがより本物らしいことを示す。

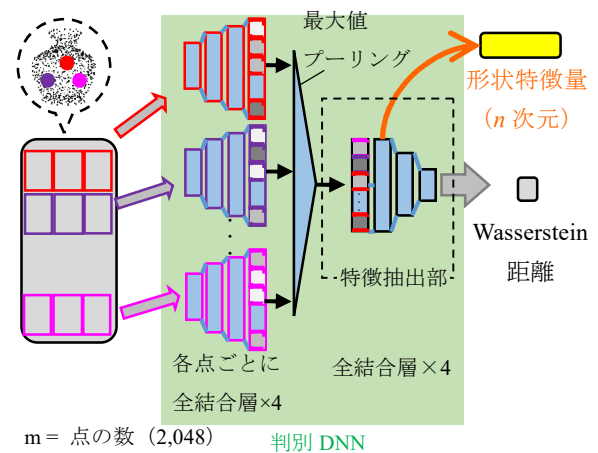


図 2 判別 DNN のネットワーク構造

3.2.2 生成 DNN

点群 GAN の生成 DNN には次の 2 つのいずれかを用い、評価実験において比較する。

- 全結合生成 DNN: 全結合層によって潜在ベクトルを 3 次元点群へ変換する DNN (図 3)。
- 折畳み生成 DNN: FoldingNet の復号化器同様、潜在ベクトルと 2D 格子点群を 3D 点群へ変換する DNN (図 4)。

上記いずれの生成 DNN においても、潜在ベクトルには平均 0.0, 分散 0.2 の正規分布からランダムにサンプルされた 32 次元ベクトルを用いる。また、生成される 3D 点群の点数は 2,048 とする。生成された 3D 点群は、判別 DNN へ入力する前に位置と大きさの正規化を施す。具体的には、点群の重心を原点に移動し、点群を半径 0.5 の球に収める。

全結合生成 DNN (図 3) は、3 層の全結合層から成り、各層のニューロン数は 1000, 2000, 6144 である。最終層の

6,144 は, 2,048 個の 3D 点群 (3×2,048) を表す. 活性化関数は, 中間層では Leaky ReLU を用い, 最終層では hyperbolic tangent を用いる.

折畳み生成 DNN (図 4) は, 潜在ベクトルで条件付けられた 2D 格子点群を, 2つの折畳み部を経て, 3D 点群へ変換する. 具体的には, まず, 2,048 個の 2D 格子点群の各々の座標ベクトルに潜在ベクトルを接続することで, 2,048 個の 34 次元点群を得る. この点群を 1つ目の折畳み部を用いて「中間的な」3D 点群へ変換する. 折畳み部は 3つ全結合層で構成され, 各層のニューロン数 1024, 1024, 3 である. 中間的な 3D 点群に潜在ベクトルを再度接続し, これを 2つ目の折畳み部に入力することで, 3D 形状を表す点群を得る. 2つ目の折畳み部の構造は 1つ目と同じである.

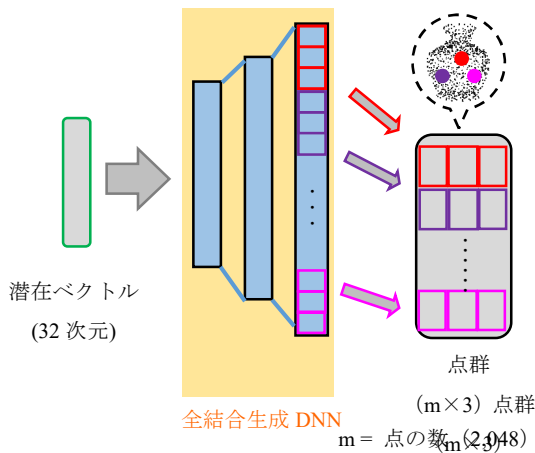


図 3 全結合生成 DNN のネットワーク構造

3.3 ネットワークの学習

3.3.1 学習の損失関数

点群 GAN の損失関数には, Gulrajani らによって提案された WGAN-GP の損失関数を用いる. 式(1)に生成 DNN の損失関数を示す. 式(1)において, \mathbf{z} は潜在ベクトル, $G(\cdot)$

は生成 DNN が出力した偽物点群, $D(\cdot)$ は判別 DNN が出力した Wasserstein 距離を示す. 式(2)は判別 DNN の損失関数である. 式(2)において, 係数 λ の項は Gradient Penalty, $\hat{\mathbf{x}}$ は本物点群 \mathbf{x} と偽物点群 $G(\mathbf{z})$ の座標値を内分して得た 3 次元点群を示す. 実験では $\lambda=1$ を用いる. 最適化には Adam [6]を用いる. 生成 DNN および判別 DNN の初期学習係数は共に 0.0001 を用いる.

本研究では, 上記の損失関数の比較対象として, Feature matching [8]を用いた損失関数 (数式(3)と(4))を用いる.

Feature matching は, 判別 DNN の中間層に現れる特徴量 f を, 本物点群と偽物点群の間で類似させる制約である. 式(3)における α の項が Feature matching であり, 実験では $\alpha=1$ を用いる.

$$L_G = -D(G(\mathbf{z})) \quad (1)$$

$$L_D = D(G(\mathbf{z})) - D(\mathbf{x}) + \lambda (\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|_2 - 1)^2 \quad (2)$$

$$L_G = -\log(D(G(\mathbf{z}))) + \alpha \|f(\mathbf{x}) - f(\mathbf{z})\| \quad (3)$$

$$L_D = -(\log(D(\mathbf{x})) + \log(1 - D(G(\mathbf{z})))) \quad (4)$$

3.3.2 学習データ

点群 GAN の学習データ (本物点群) には, 3D ポリゴンモデルを点群化したものを用いる. Ohbuchi らの手法 [9]を用い, 1 個のポリゴンモデル当たり 2,048 個の点群へ変換する. 3D ポリゴンモデルには, 椅子, 飛行機, 人間など 40 個のカテゴリに分類された 9,843 個の剛体 3D モデルからなる ModelNet40 [5] の学習用データセットを用いる. 点群 GAN は教師なし特徴量学習手法であるため, 3D モデル

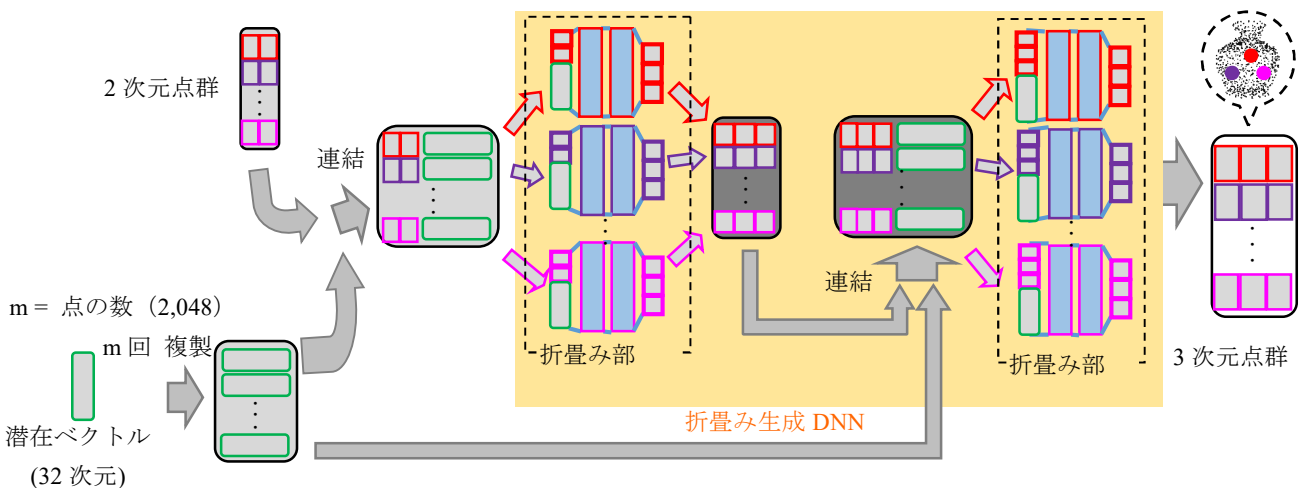


図 4 折畳み生成 DNN のネットワーク構造

に付与されたカテゴリラベルは用いない。検索精度の評価には、ModelNet40 の評価用データセットを用いる。

3.4 形状特徴量の抽出

点群 GAN 学習の後、判別 DNN を 3D 点群モデルからの形状特徴量抽出に用いる。特徴抽出対象の 3D 点群モデルを判別 DNN へ入力し、判別 DNN の特徴抽出部のニューロン活性がその 3D 点群モデルの形状特徴量である。特徴抽出部は N 個の全結合層を持つため、我々は N 個の層の各々から形状特徴量を抽出する。各特徴量の次元数はその特徴量を抽出した層のニューロン数と同じ 1,024 (ただし、判別 DNN(4)の 3 層目、4 層目のみ、それぞれ 512, 64) である。

2 つの 3D モデルの形状類似度は、それら 3D モデルの形状特徴量間のコサイン類似度を用いて計算される。

4. 実験と結果

4.1 実験条件

点群 GAN が学習した形状特徴量の精度を 3D モデルの形状類似検索のシナリオで評価する。実験ではまず、点群 GAN の構造および損失関数が形状特徴量の精度に与える影響を調査する。次に、既存の手作り形状特徴量と検索精度の比較を行う。

データセット: 評価実験には ModelNet40 データセット [5]を用いる。3.3.2 項で述べた通り、点群 GAN の学習には ModelNet40 の学習用データセットに含まれる 9,843 個の 3D モデルを用いる。検索精度の評価には ModelNet40 のテストデータセットに含まれる 2,468 個の 3D モデルを用いる。検索精度の評価尺度には Mean Average Precision (MAP) を用いる。各実験を 3 回ずつ行い、3 つの MAP 値の平均と標準偏差を報告する。

比較対象: 既存の 3D 点群向け手作り形状特徴量である AAD [9], SPRH [10], BF-LSF [11], DkSA-POD [12]を比較対象として用いる。AAD および SPRH は有向 3D 点群向けの全体形状特徴量である。AAD と SPRH は、有向点の対ごとに距離、向きベクトルの内積等の統計量を計算し、それら統計量をヒストグラム化して得た形状特徴量である。BF-LSF と DkSA-POD は、有向 3D 点群から局所形状特徴量を多数抽出し、これら局所特徴量を 3D モデル当たり 1 つの特徴ベクトルに統合して得た形状特徴量である。

4.2 実験結果

4.2.1 判別 DNN の比較

判別 DNN の特徴抽出部のネットワーク構造について、層数 N を 4, 6, 8 とした場合に得られる形状特徴量の精度を比較する。本節では、層数 N の特徴抽出部を持つ判別 DNN を「判別 DNN(N)」と表記する。例えば、 $N=4$ のとき、判

別 DNN(4)と表記する。生成 DNN には 3.2.2 節で述べた全結合生成 DNN を用い、WGAN-GP の損失関数を用いる。

図 5, 図 6, 図 7 にそれぞれ、判別 DNN(4), 判別 DNN(6), 判別 DNN(8)から抽出された形状特徴量の検索精度を示す。図中の「特徴抽出層」は特徴抽出部の入力側から何番目の層を特徴抽出に用いたかを示す。図 5, 図 6, 図 7 に示した結果より、実験したいずれの層数 N においても特徴抽出部の 2 層目または 3 層目から抽出した形状特徴量が高い検索精度を示した。一方で、出力に近い層から抽出した特徴量は精度が低く、ばらつきが大きい。出力層に近くなるにつれ形状特徴量が Wasserstein 距離に変換されるためと推察される。

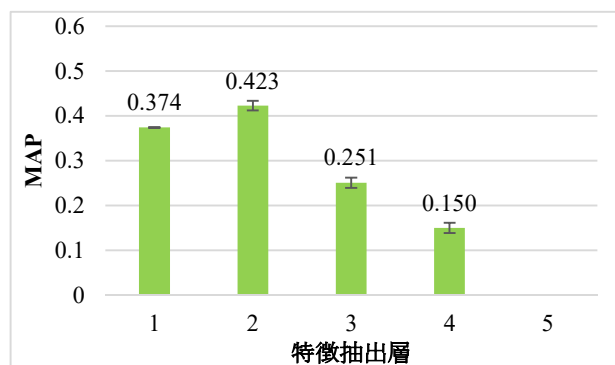


図 5. 判別 DNN(4)から抽出した特徴量の検索精度。

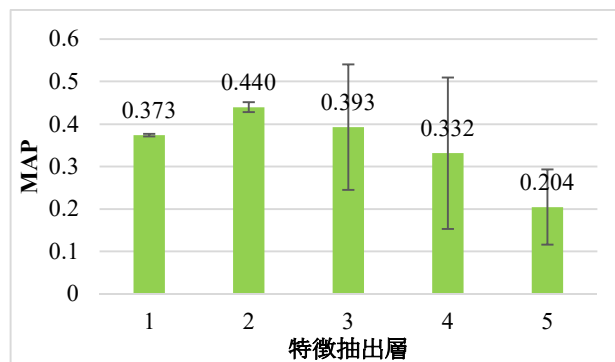


図 6. 判別 DNN(6)から抽出した特徴量の検索精度。

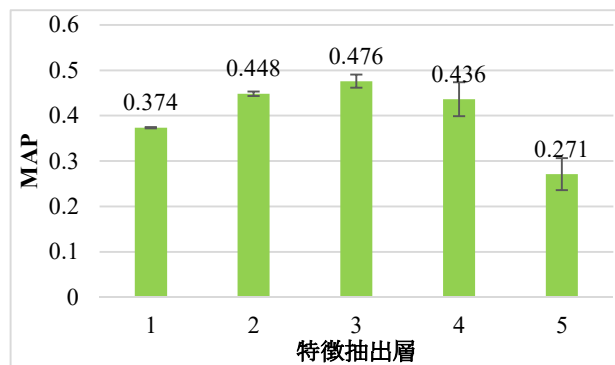


図 7. 判別 DNN(8)から抽出した特徴量の検索精度。

4.2.2 生成 DNN の比較

図 8 に全結合生成 DNN と折畳み DNN の比較結果を示す。判別 DNN には 6 層の特徴抽出部を持つ判別 DNN(6)を用い、WGAN-GP の損失関数を用いる。図 8 より、折畳み生成 DNN の採用により特徴量学習の効果が高まる事が分かる。折畳み生成 DNN を用いて学習する場合は、全結合生成

DNN を用いて学習する場合よりも精度の平均が高く、かつ精度のばらつきが小さい形状特徴量が獲得された。特に、特徴抽出部の 3 層目から抽出された形状特徴量の MAP 値が 0.505 と高い。4.2.5 節で示すように、折畳み生成 DNN は本物らしい 3D 点群を生成する。その結果、判別 DNN の学習が効果的に進み、形状特徴量が改善したと考えられる。

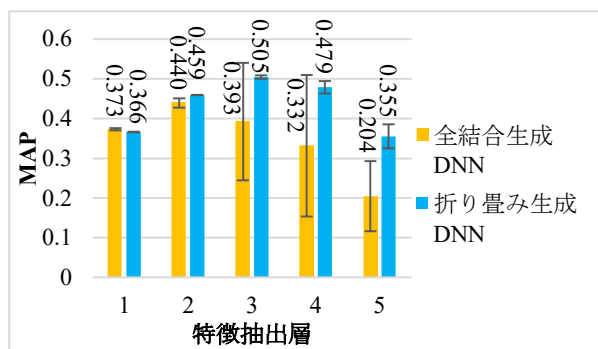


図 8 全結合生成 DNN と折畳み生成 DNN の比較。

4.2.3 損失関数の比較

点群 GAN の学習に用いる損失関数について、Gulrajani らの WGAN-GP 法と、Salimans らの Feature matching 法を比較する。生成 DNN には全結合生成 DNN、判別 DNN には判別 DNN(4)を用いる。本実験でのみ、点群モデル当たりの点数を 1,024 とする。図 9 に実験結果を示す。図 9 のグラフの横軸は形状特徴量の次元数である。特徴量の次元数は、判別 DNN の特徴抽出部のニューロン数を変化させ

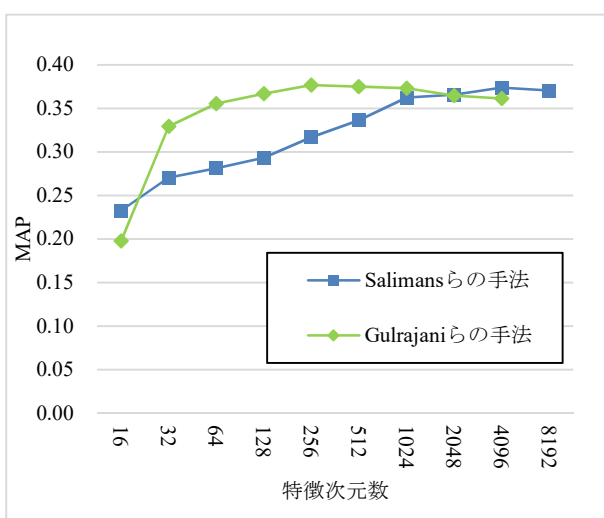


図 9 点群 GAN の学習に用いる損失関数の比較。

ることで決定する。

図 9 より、1,000 次元以上の形状特徴量を用いる場合には、どちらの損失関数も同等の検索精度を示す。一方で、数十～数百次元の形状特徴量を用いる場合は、WGAN-GP が高い検索精度を示す。WGAN-GP の利用により点群 GAN の学習が安定し、数十次元以上のいずれの次元数においても高精度な特徴量が学習された。

4.2.4 既存の手作り形状特徴量との比較

表 1 に、点群 GAN により学習した形状特徴量と、既存の手作り形状特徴量との検索精度比較を示す。点群 GAN は 2通りの DNN 構造を用いた場合の精度を示す。点群 GAN を用いて学習した形状特徴量は、比較に用いた 4 つの手作り特徴量を上回る検索精度を示す。比較対象の中でも DkSA-POD は、ラベルなしの 3D 点群モデルの形を比較するための state-of-the-art な形状特徴量である。点群 GAN が DkSA-POD を上回ることから、本研究のアプローチである GAN を用いた教師なし 3D 形状特徴量の学習が効果的であることが分かる。

表 1. 既存の手作り形状特徴量との比較。

3D 点群形状特徴量	MAP
AAD [9]	0.272
SPRH [10]	0.329
BF-LSF [11]	0.323
DkSA-POD [12]	0.469
点群 GAN (全結合生成 DNN + 判別 DNN(8))	0.476
点群 GAN (折畳み生成 DNN + 判別 DNN(6))	0.505

4.2.5 生成された 3D 点群

図 10 に、折畳み生成 DNN によって生成された偽物 3D 点群の例を示す。図 10 の点群はいずれも、点群 GAN の学習後の生成 DNN にランダム潜在ベクトルを入力することで得た。椅子、テーブル、ベッドなど、多様で、かつ、本物らしい 3D 点群が生成されたことが見て取れる。

5. まとめと今後の課題

ラベルを持たない多量の 3 次元形状データを効率的に管理するため、これらラベルなしの 3D 形状データから高精度な形状特徴量を獲得する手法が望まれている。本研究では、3D 点群向けの敵対的生成ネットワーク (GAN) を構築し、これを用いた 3D 形状特徴量の教師なし学習手法を提案した。提案した点群 GAN は、3D 点群を生成する DNN (生成 DNN) と、3 次元点群の本物らしさを予測する DNN (判別 DNN) とを互いに競わせながら訓練する過程で、3D 点群の形状特徴を捉える能力を判別 DNN に獲得させる。



図 10 点群 GAN の折畳み生成 DNN が生成した 3D 点群の例。

点群 GAN を効果的に、かつ、安定的に学習させるため、ネットワーク構造と損失関数の双方に対して工夫を施した。具体的には、点の順序変化に影響を受けない 3D 点群解析用 DNN を用いて点群 GAN を構築した。また、本物の点群データ集合と偽物の点群データ集合の集合間距離を学習の損失として用いた。

3D モデルの形状類似検索のシナリオを用いた実験的評価の結果、点群 GAN が獲得した形状特徴量は、既存の手作り形状特徴量を上回る検索精度を示すことが分かった。また、ネットワーク構造と損失関数の工夫点のそれぞれが、いずれも、形状特徴量の精度の改善に効果的であることを確かめた。

今後の課題として、(1) 形状特徴量のさらなる高精度化、(2) 3D 形状の幾何変換に対する頑強性の獲得、(3) 評価実験の強化を挙げる。現状の点群 GAN のネットワーク構造と損失関数が必ずしも形状特徴量の学習に最適であるとは限らない。ネットワーク構造または損失関数を改良することで特徴量の精度が改善する可能性がある。また、現状の点群 GAN は位置、大きさ、向きが一貫して揃った 3D 点群データを学習に用い、これらを生成・判別するように訓練される。そのため、獲得される形状特徴量がこれらの幾何変換に対する頑強性を持たないと予想される。今後は 3D 形状の回転、スケーリング、平行移動などの幾何変換に対する頑強性を点群 GAN へ付与する手法について検討する。さらには、教師なし形状特徴量学習の既存手法 (例えば、[4])との精度比較等を行うことで、評価実験の充実を図る。

参考文献

- [1] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C.: "Improved Training of Wasserstein GANs", In Advances in Neural Information Processing Systems, pp. 5769-5779 (2017).
- [2] Arjovsky, M., Chintala, S., and Bottou, L.: "Wasserstein generative adversarial networks", In International Conference on Machine Learning, pp. 214-223 (2017)
- [3] Charles, R. Q., Su, H., Kaichun, M., and Guibas, L. J.: "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation", Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference, pp77-85, (2017).
- [4] Yang, Y., Feng, C., Shen, Y., and Tian, D.: "FoldingNet: Point Cloud Auto-encoder via Deep Grid Deformation", In Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Vol.3, (2018).
- [5] Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J.: "3D ShapeNets: A Deep Representation for Volumetric Shapes". In Processing of the IEEE conference on computer vision and pattern recognition, pp. 1912-1920 (2015).
- [6] Kingma, D. P., and Ba, J.: "Adam: A method for stochastic optimization." International Conference on Learning Representations (ICLR) (2015).
- [7] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: "Generative Adversarial Nets", In Advances in neural information processing systems, pp. 2672-2680 (2014).
- [8] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). "Improved techniques for training gans", In Advances in Neural Information Processing Systems (pp. 2234-2242).
- [9] Ohbuchi, R., Minamitani, T., Takei T.: "Shape-Similarity Search of 3D Models by using Enhanced Shape Functions", IJCAT, Vol.23, No.2/3/4, pp70-85 (2005).
- [10] Wahl, E., Hillenbrand, U., and Hirzinger, G.: "Surflet-pair-relation histograms: a statistical 3D-shape representation for rapid classification," In 3-D Digital Imaging and Modeling, 3DIM 2003, Proceedings. Fourth International Conference, pp. 474-481 (2003).
- [11] Ohkita, Y., Ohishi, Y., Furuya, T., and Ohbuchi, R.: "Non-rigid 3d model retrieval using set of local statistical features", In Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference, pp. 593-598 (2012)
- [12] Furuya, T., and Ohbuchi, R.: "Accurate Aggregation of Local Features by using K-sparse Autoencoder for 3D Model Retrieval", In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, pp293-297, (2016)
- [13] Su, H., Maji, S., Kalogerakis, E., Learned-Miller, E.: "Multi-view Convolutional Neural Networks for 3D Shape Recognition," In Proc. ICCV 2015, (2015).
- [14] Furuya, T., Ohbuchi, R.: "Deep Aggregation of Local 3D Geometric Features for 3D Model Retrieval," In Proc. BMVC 2016, pp. 121.1-121.12, (2016).
- [15] Radford, A., Metz, L., & Chintala, S. (2015). "Unsupervised representation learning with deep convolutional generative adversarial networks", arXiv preprint arXiv:1511.06434.
- [16] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). "Rectifier nonlinearities improve neural network acoustic models", In Proc. icml (Vol. 30, No. 1, p. 3).