

縄文土器3D計測データによるクラスタ解析モデルの検討

板垣 正敏
Masatoshi Itagaki

板垣正敏中小企業診断士事務所
itagakim@gmail.com

市川 賢太
Kenta Ichikawa

株式会社BSN アイネット
ichikawa@bsnnet.co.jp, <http://www.bsnnet.co.jp>

藤田 晴啓
Haruhiro Fujita

新潟国際情報大学
fujita@nuis.ac.jp, <https://www.nuis.ac.jp>

宮尾 亨
Toru Miyao

新潟県歴史博物館
miyao.toru@pref.niigata.lg.jp, <http://nbz.or.jp>

1. はじめに

縄文土器の3次元スキャンデータをもとに、その形状や色彩を使用した深層学習クラスタ解析の手法を開発する。今回は特徴抽出器としてとして3次元畳み込みニューラルネットワーク(3D-CNN)を教師なし学習で学習させる手法の有効性を確認した。

2. 従来の手法の問題点

縄文土器の分類はこれまで出土地や出土地層などの情報を参考に、専門家が形状や様式、製造方法等に基づいて行ってきた。このため主観的要素が入る余地があり、専門家によって判断が異なる場合がある。また、縄文土器が地域をこえて運搬・流通していたかどうかや、製造技術を持った者の交流があったかどうかについては、一部材料となる土の組成分析などが行われた例を除けば、仮説の領域を出ず裏付けがなかった。

3. 今回の改良点

今回我々は、専門家による分類や出土地、出土地層などの情報によらず、縄文土器を3Dスキャナでスキャンして得られたデータのみをもとに、深層学習を用いたクラスタ解析を行い、類似性による分類とその類似点の解釈を行うことで、縄文土器の分類に客観性を持たせるとともに、離れた地域間での類似性を客観的に評価する手法の開発を目指す。

4. 手法の検討

4.1 先行事例

3Dデータのクラスタ解析では、従来[Zou 16]のようにphysical descriptorと呼ばれる物理特性記述子を用いた手法が使われてきた。代表的なphysical descriptorとしては、断面の輪郭曲線を展開しフーリエ解析により記述子を得る方法などがある。こうした手法は単純な構造物では容易だが、縄文土器のような複雑な形状を持つ物体では、特徴量を得るまでの過程が複雑であるほか、断面の方向や数などで結果が左右されてしまい、比較が困難であるなどの欠点がある。

4.2 特徴抽出器としての深層学習

今回我々は、3Dデータの特徴量を抽出する手法として、近年画像解析などの分野で主流になっている深層学習の手法を用いることとした。3Dデータ全体を用いることで対象物の特徴を偏りなく抽出できると考えたからである。

3Dデータの分類モデルとしては、3Dデータを表すデータの形態により、点群データをそのまま用いるもの、Voxelデータを用いるもの、投射による複数の2D画像を用いるものなどがある。[Su 18]これらの分類モデルの出力直前の特徴ベクトルを取り出すことによって、k-means法などの既存のクラスタ解析手法を用いることが可能になるが、深層学習モデルが特徴量を抽出できるようになるには、なんらかのデータによって事前学習を行う必要がある。

2D画像の場合には、ImageNetなどの大量かつ多量の画像を含むデータセットが存在し、これを使用して事前学習を行なったモデルを使用することが可能であるが、3D

データの場合、こうした巨大なデータセットが存在せず、また対象となるデータも、人為的に作成されたCADモデルのようなデータが主であることから、縄文土器のような複雑な形状を持つ対象物の特徴量抽出ができるかどうかは不明である。

このため、我々は2D画像において提案された[Fard 18, Caron 18, Caron 19]のように、深層学習とk-means法を組み合わせた深層学習クラスター解析の手法を用いることとした。(図1)この手法は、深層学習モデルから得た特徴量を元にクラスター解析を行い、出力されたクラスター割り当て情報を疑似ラベルとして深層学習モデルの学習を行うことを繰り返す手法である。この手法を取ることで、純粋に対象物群から得られた情報のみを元にしたクラスター解析が行えることになる。

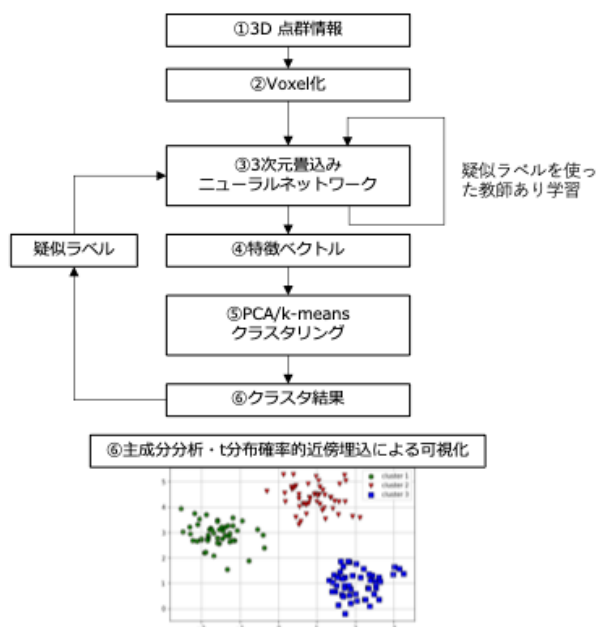


図1 深層学習クラスター解析の流れ

4.3 比較手法

純粋な教師なし学習により対象物の特徴量を得る代表的な手法として自己符号化器(autoencoder)がある。これは対象物の特徴量を抽出するencoderと得られた特徴量から元の対象物の復元を行うdecoderを接続したモデル(図2)であり、対象物そのものを教師データとして扱う自己教師あり学習とも呼ばれる。自己符号化器が集約した特徴量を隠れベクトル(latent vector)と呼ぶが、学習の終わった自己符号化器のencoderを特徴量抽出器として使い、得られた特徴量を使用してクラスター解析を行うことができる。

しかしながら、元の対象物を高い精度で復元できる自己符号化器の学習には、大量のデータが必要である。これに対し深層学習クラスター解析では、対象物を識別可能

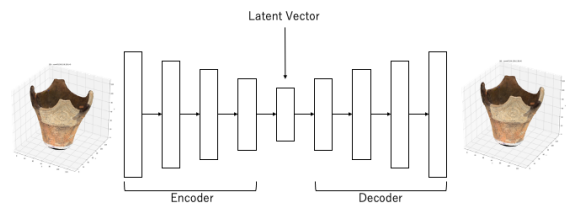


図2 自己符号化器

な分類器が学習できれば、十分特徴量抽出が可能であると考えられるほか、学習の過程でクラスター解析の結果と、そのクラスターに対象物を分類する分類器が自動的に得られることになる。この分類器を使用すれば、[Yang 18]のようにCAM(Class Activation Map)などの手法で、クラスターの特徴が対象物のどの部分に見られるかなどの解析も可能になると考えた。

5. 実験

5.1 実験手法

現状では、3Dスキャンが行われている縄文土器はわずかであるため、実際の縄文土器のデータを用いた手法の開発はできない。このため、3Dデータの処理性能評価等の目的で公開されているデータセットModelNet10[Wu 15]を使用し、いくつかの手法を試みることで、縄文土器の解析に適したモデルを開発することとした。

ModelNet10は点群データの形式で提供されているため、Pythonライブラリtrimesh[Dawson-Haggerty et al.]を使用してVoxelに変換を行なった。ただし、元のオブジェクトはサイズが異なるため、指定されたVoxel立方体の内側に一定のマージンを残した上で収まるようにリサイズを行なっている。今回は実験の際のメモリおよび計算時間の短縮のため、立方体のサイズを64x64x64にして変換を実施した。

モデルの構築には、深層学習部分にTensorFlow[Abadi 15]、主成分分析およびk-means++法の組み込みにscikit-learn[Pedregosa 11]、可視化にmatplotlib[Hunter 07]およびplotly[Plotly Technologies Inc. 15]を使用した。

5.2 実験結果

§1 実験1: Voxelデータのクラスタリング

3次元畳み込みによる特徴抽出の効果を確認するため、ベースラインとしてVoxelデータを展開したベクトルをそのまま使用し、主成分分析ならびにk-means++法によるクラスター解析を試みた。主成分(10次元)による2次元散布図(図3)およびt-SNEによる2次元散布図(図4)を示す。なお、凡例は各データの分類クラスを表示し

ている。椅子やモニターなどある程度分離ができているクラスもあるが、領域が重なり合っているクラスも多い。



図 3 主成分正規化後の特徴量の 2 次元散布図

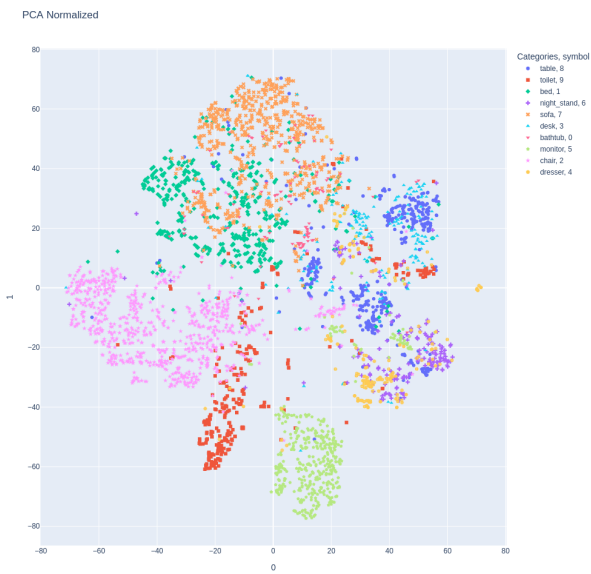


図 4 主成分正規化後の特徴量の t-SNE2 次元散布図

正規化した主成分を使用し、k-means++法のクラスター数 k を 2 から 20 まで変えてクラスター内残差平方和 (SSE) をグラフ化したのが図 5 である。明確な変曲点 (エルボー) はみられず、SSE はなだらかに低下している。この傾向は、主成分抽出前のベクトルをそのまま使用したもの (図 6) でも変わらなかった。

§ 2 実験 2: 深層学習クラスター分析 (クラスター数 10)

最初のモデルは、3次元の畳み込みと Max Pooling, その後バッチ正規化のユニットを 4 回繰り返した後、Global

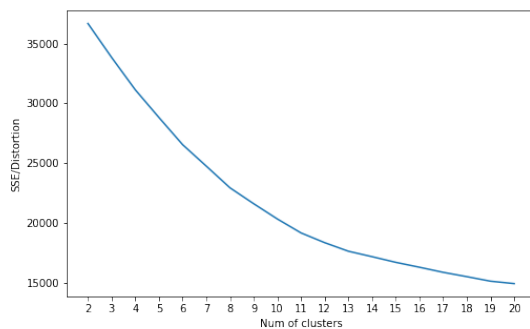


図 5 主成分正規化後の特徴量による k-means++法の SSE

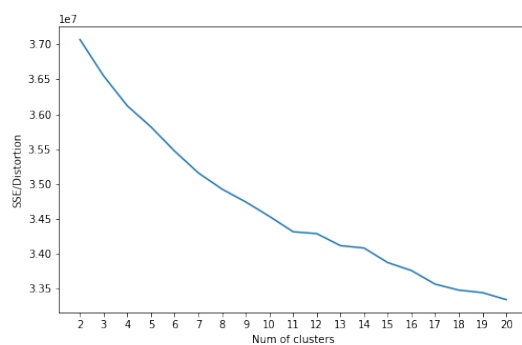


図 6 未処理の特徴量の k-means++法の SSE

Average Pooling 層をへて 128 ユニットの全結合層に至る特徴抽出器 (図 7) をベースに、活性化層と 10 クラス分類の出力層を加えた分類器を用いたもの (図 8) で、特徴抽出器の出力を元に主成分分析および平均 0 分散 1 の正規化を行なった特徴ベクトルを使った k-means++法によるクラスター疑似ラベルを生成しながら、学習を行う。

k-means++法による疑似ラベルの生成の都度、深層学習モデルの学習エポックを何回繰り返すかによって性能が異なることが予想されるため、これを 1 回、3 回、5 回、10 回と変えて実験を行なった。その際の損失関数および SSE の変化を示す。(図 9, 図 10)

100 エポック学習後のモデルを使用して特徴量を抽出し、正規化した主成分によるプロットを行なったのが図 11 である。

また、t-SNE によってクラスター表示を行なったのが図 12 である。特徴ベクトルそのもの (図 12), 10 次元の主成分 (図 13), その主成分を正規化したもの (図 14) を比較しても、大きな違いは見られない。これは、3次元畳み込みによる深層学習モデルによって十分特徴量の抽出が行われていることを示唆している。

一方、前述の図でもわかる通り、得られたクラスターが学習データのクラスと 1 体 1 に対応しているわけではない、これを確認するため元データのクラスと、得られたクラスターの ID の対応状況を図 15 で示す。ベッド

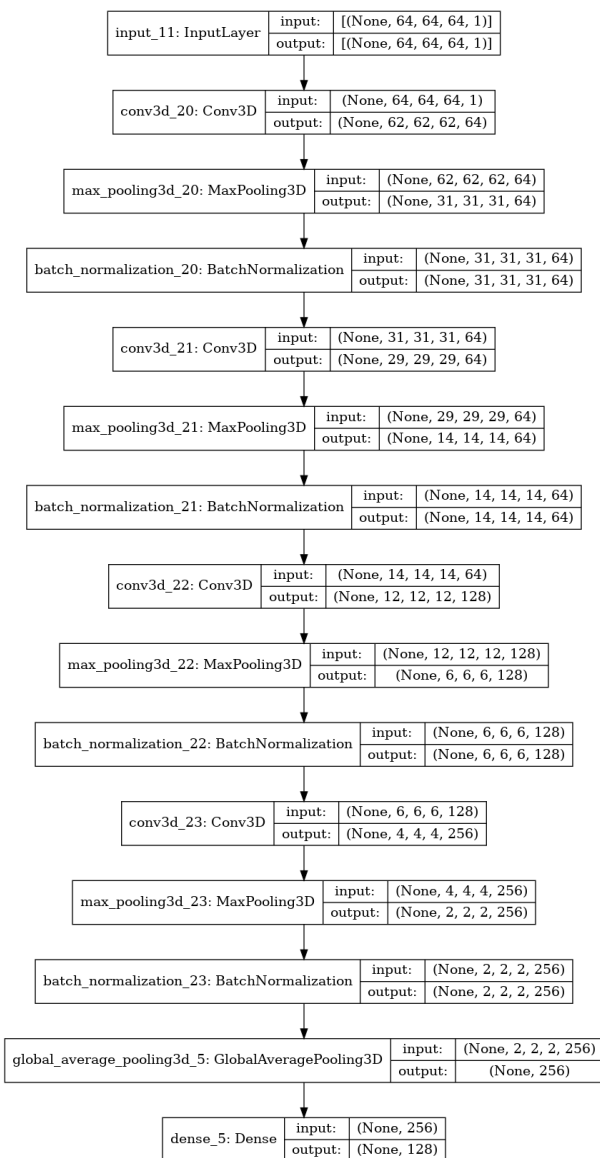


図 7 特徴抽出器の構造

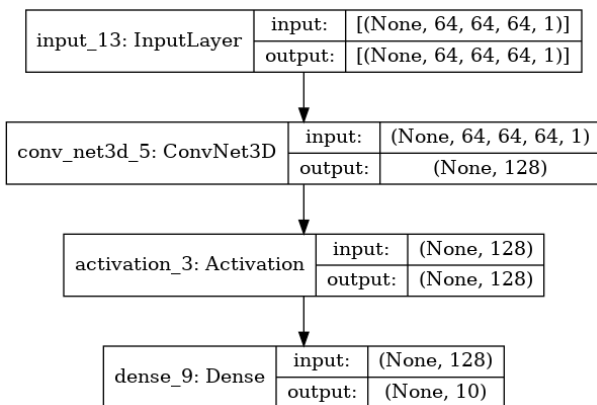


図 8 分類モデル全体の構造

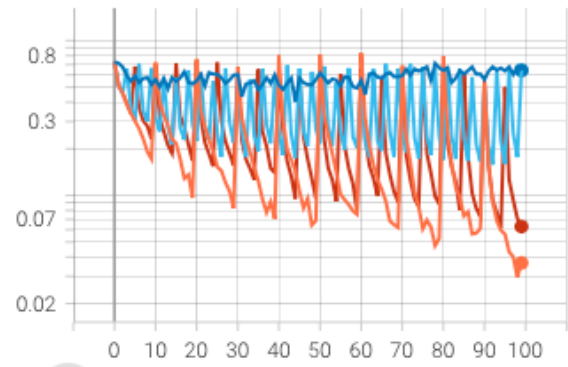


図 9 損失関数の推移

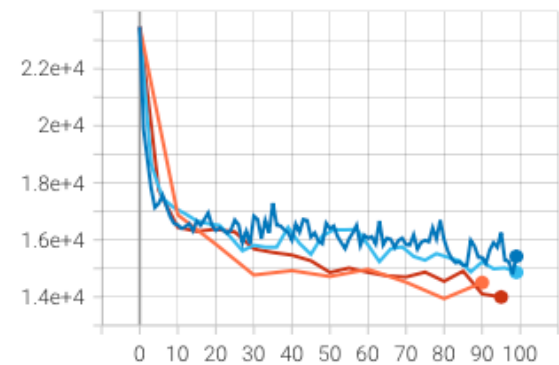


図 10 SSEの推移

やモニターではほぼ対応するクラスターがみられるが、クラスター内に複数のクラスが混在しているものが多い。これは、学習後のモデルが物理形状の特徴を基準にクラスタリングを行なっているからであると考えられる。

クラスター数を2から20まで変えて実施したk-means++

PCA Normalized (w/o t-SNE)

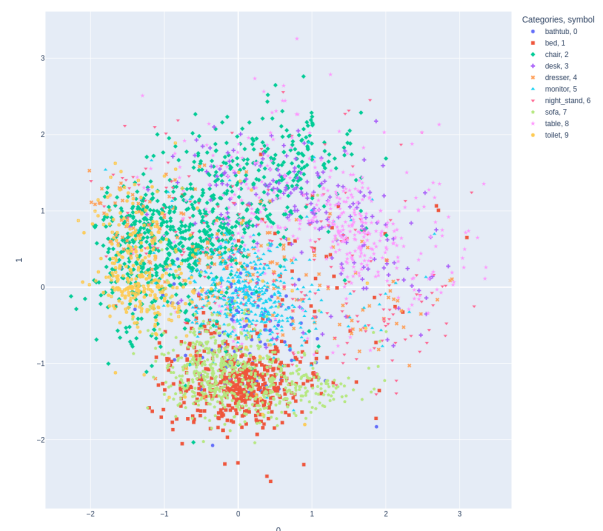


図 11 主成分正規化後の特徴量の2次元散布図

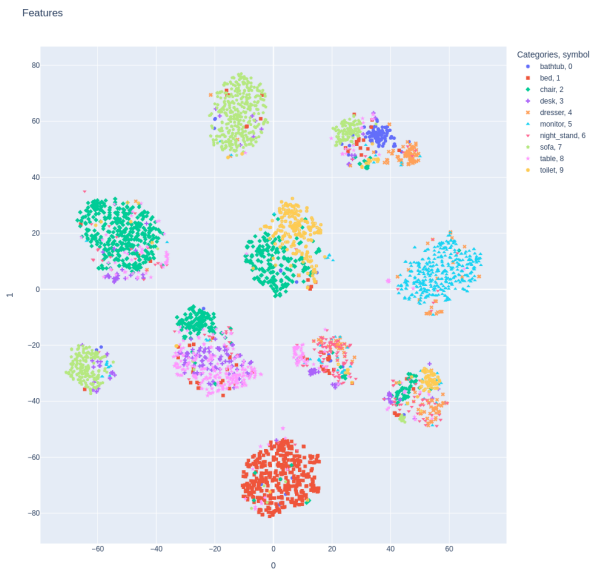


図 12 特徴量の t-SNE2 次元散布図

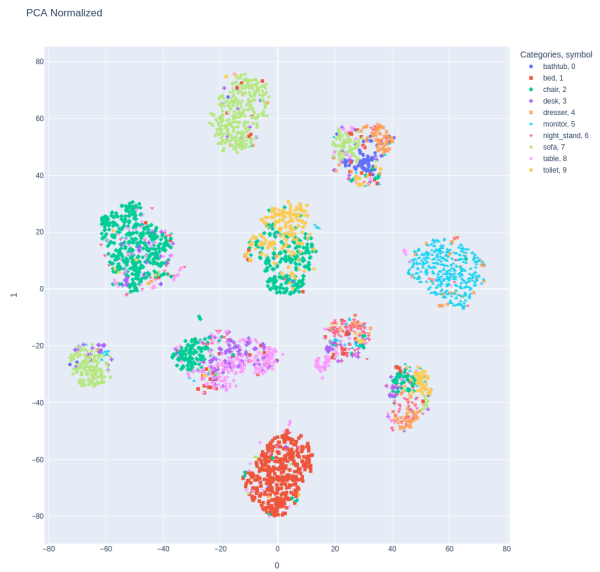


図 14 特徴量の主成分を正規化した t-SNE2 次元散布図

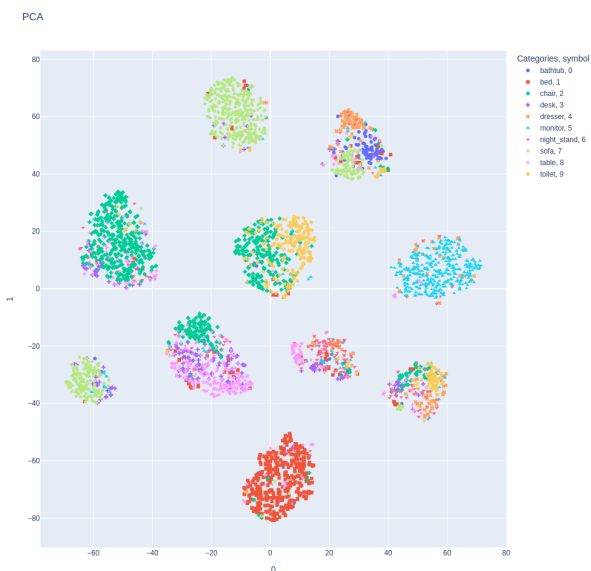


図 13 特徴量の主成分の t-SNE2 次元散布図

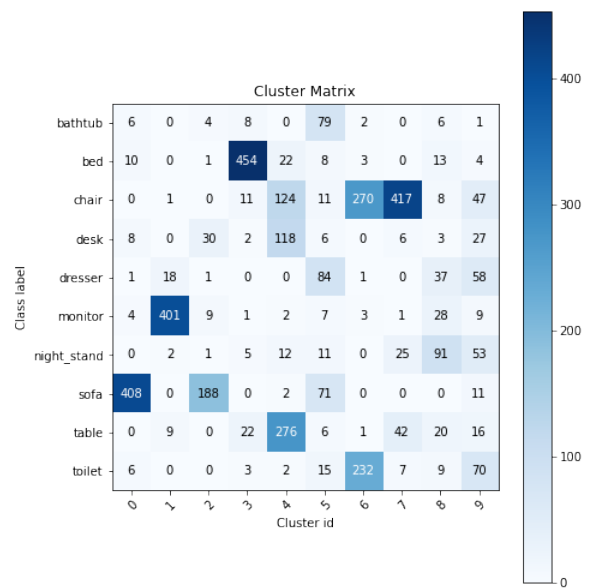


図 15 クラスラベルとクラスターの対応

法の SSE の変化を図 16 に示す。クラスター数 10 が変曲点となっているが、これは学習時に設定したクラスター数が 10 であったためと考えられる。

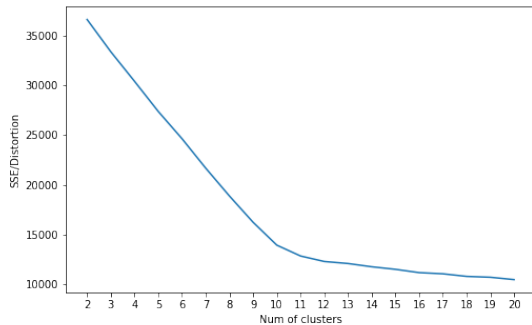


図 16 主成分正規化後の特徴量の t-SNE2 次元散布図

§ 3 実験 3: 深層学習クラスター解析 (クラスター数 20)

比較のためにクラスター数を $k=20$ に設定し実験を行った。疑似ラベル更新ごとの深層学習モデルの学習回数は 5 回に設定した。主成分 (図 17) および t-SNE 法の 2 次元散布図 (図 18) でわかるように、クラスター内に複数のクラスが混在することが少なくなっている一方、同じクラスでも複数のクラスターに分離が見られる。なお、t-SNE 法の場合、クラスター間の距離や位置関係は意味を持たないことに注意が必要である。

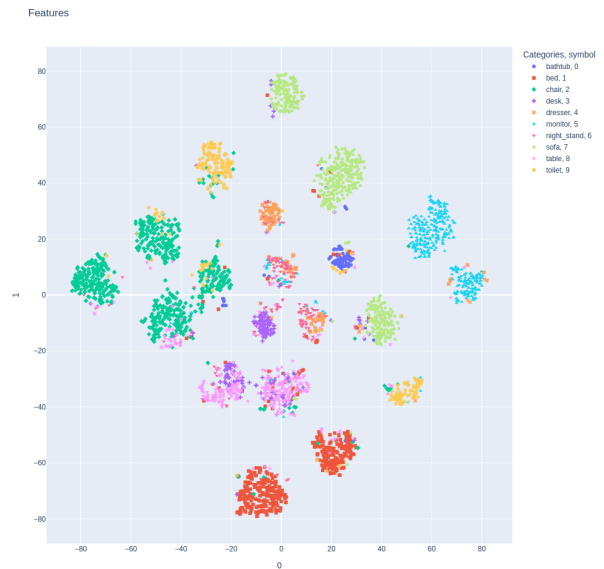


図 18 主成分正規化後の特徴量の t-SNE2 次元散布図



図 17 主成分正規化後の特徴量の 2 次元散布図

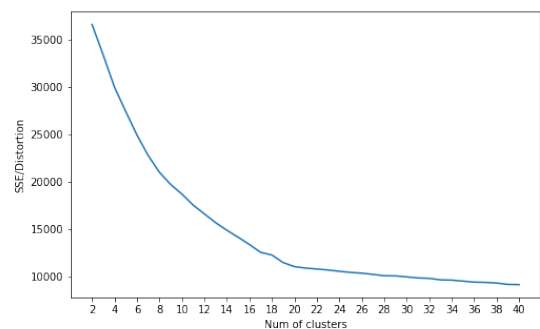


図 19 k-means++法の $k=2$ から 40 までの SSE 推移

クラスター数を 2 から 40 まで変えて実施した k-means++ 法の SSE の変化 (図 19) を見ると、変曲点は $k=20$ の付近と見られ、学習の際に設定したクラスター数が学習結果に反映していることが確認された。

6. ま と め

k-means++法による疑似ラベル生成と3次元畳み込みニューラルネットワークを用いた深層学習クラスター解析が対象物の特徴抽出に有効であることが確認できた。ただし、クラスター数は疑似ラベルを生成するためにk-means++法で指定したクラスター数によって影響を受けるほか、k-means++法の初期化のランダム性により、最初に割り当てられたクラスターの影響を受ける可能性がある。このため、本研究が想定する縄文土器の事前情報によらないクラスター解析においては、想定されるクラスター数よりも大きな値をクラスター数に設定し、より細かい特徴分類を数回試行する必要があると考える。

また、t-SNE法によるクラスター可視化の手法は、クラスタリングの状況を確認するには有効だが、非線形の次元削減を行なっているため、クラスターの位置や相互関係は意味を持たない。このため、対象物の他の対象物群との関係を考察する意味では、主成分分析による次元削減後の3次元散布図などを使用した方が良いと考えられる。

今後は、事前に設定したクラスター数やクラスターの初期値によって影響を受けることのない、自己符号化器による特徴抽出や、距離学習の手法を使って3次元畳み込みニューラルネットワークの学習を行う手法も試行する予定である。

◇ 参 考 文 献 ◇

- [Abadi 15] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015), Software available from tensorflow.org
- [Caron 18] Caron, M., Bojanowski, P., Joulin, A., and Douze, M.: Deep Clustering for Unsupervised Learning of Visual Features, in *Proceedings of the European Conference on Computer Vision (ECCV)* (2018)
- [Caron 19] Caron, M., Bojanowski, P., Mairal, J., and Joulin, A.: Unsupervised Pre-Training of Image Features on Non-Curated Data, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019)
- [Dawson-Haggerty et al.] Dawson-Haggerty et al.: trimesh
- [Fard 18] Fard, M. M., Thonet, T., and Gaussier, E.: Deep k -Means: Jointly clustering with k -Means and learning representations (2018)
- [Hunter 07] Hunter, J. D.: Matplotlib: A 2D graphics environment, *Computing in Science & Engineering*, Vol. 9, No. 3, pp. 90–95 (2007)
- [Pedregosa 11] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in Python, *Journal of machine learning research*, Vol. 12, No. Oct, pp. 2825–2830 (2011)
- [Plotly Technologies Inc. 15] Plotly Technologies Inc.: Collaborative data science (2015)
- [Su 18] Su, J.-C., Gadelha, M., Wang, R., and Maji, S.: A Deeper Look at 3D Shape Classifiers, in *Second Workshop on 3D Recon-*

struction Meets Semantics, ECCV (2018)

- [Wu 15] Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J.: 3D ShapeNets: A Deep Representation for Volumetric Shapes (2015)
- [Yang 18] Yang, C., Rangarajan, A., and Ranka, S.: Visual Explanations From Deep 3D Convolutional Neural Networks for Alzheimer’s Disease Classification (2018)
- [Zou 16] Zou, K., Zhang, Z., Zhang, J., Liu, J., and Gan, L.: 3D model retrieval scheme based on fuzzy clustering for physical descriptors, *Journal of Algorithms & Computational Technology*, Vol. 10, No. 1, pp. 12–22 (2016)