

深層生成モデル VAE および VQ-VAE の潜在空間を使った須恵器のクラスタリング

河原和好 南雲彩花 山本亮 市川健太 藤田晴啓

先行研究では、疑似ラベルを教師とする畳み込みニューラルネットワーク (CNN) を用いて須恵器の特徴量を抽出し、その特徴量に基づき k-means 法によるクラスタ解析を行った。しかし、Grad-CAM 等による確認では、畳み込み層による特徴抽出では局所的な情報を抽出していると思われるが、全体的な構造を捉えきれているかは確認できていなかった。本研究では深層生成モデルにおける潜在空間から得られる特徴量表現を用いたクラスタリング手法の有効性を検討する。本手法は、潜在空間から須恵器の 3 次元データを再構成することで、大域的な情報の妥当性を評価することが可能であると考えられる。

VAE にて入力した 3D 形状を再構成の途中で取得される潜在変数を使ったクラスタリング手法を試みた。入力データは 128x128x128 ボクセルで 1000Epoch では須恵器形状をかなりの精度で再構成できた。潜在変数 1024 次元にてクラスタリングを試みたが明瞭にまとまったクラスターはみられなかった。(20231229 スライド)

VAE は潜在変数を正規分布にしたがう連続的な変数としているのに対し、VQ-VAE (ベクトル量子化変分オートエンコーダ) の潜在変数は離散的なベクトルになる。自然界の事象は離散的なものも多く、VQ-VAE を使い二つのモデルを開発した。

モデル 1 (20240227 スライド) は学習率 0.001、Epoch 300。モデル 2 (20240304 スライド) は学習率 0.0002、Epoch 1000 とした。ふたつのモデルは 16384 次元の潜在変数とともに使用し比較を行なった。結果はモデル 1 はクラスターのまとまりがよい結果となった。須恵器の型式を連続的なものではなく、離散的なものと考えれば、潜在変数を 16384 にした場合では明らかに 3 集団に分かれている。II-5 だけのクラスター、II-3,4,5 が混在したクラスター、II-1 および 2 が中心となるが、端の方に II-3 や 4 が混在するクラスターで形成された。モデル 2 はクラスターがまとまらず連続したかたちとなった。一方モデル 2 の方が再構成の精度が安定している。

モデル 1 で明瞭にクラスタが分かれたのは、おそらく蓋の稜線の有無が大きく作用したのではないかと推測されるが、他方 II-5 だけで固まったクラスターの説明がつかない。クラスターのまとまりを作った II-5 型式のものは、実物を観察すると底がすぼまる特徴的な形態のものが多い。

VQ-VAE は引き続き学習率および Epoch がクラスターのまとまりを変動させる原因の究明と VQ-VAE2 等の他のモデルによる解析を検討する。