

# 商標の類似性の定量的評価及び文字と図形とからなる商標の生成

ディープラーニング基礎講座2018 Team8

## 概要

本プロジェクトでは、商標の類似性の評価に係る負荷を軽減するために、商標の類似性の評価を深層学習を用いて定量的に行えるか調査する。また、侵害や模倣被害の予見性を向上させるために、類似度に応じた商標の分布を深層学習によって可視化できるか調査する。さらに、深層学習によって文字と図形とからなる商標を生成できるか調査する。

## 1. 序論

深層学習によって、画像の類似性の判断を行えるようになってきている。画像の類似性の判断が重要となる分野として商標がある。商標の類似性の判断は、商標権の効力発生や権利行使の可否を左右する。類似性の判断は経験を要する行為であり、特許庁や裁判所等において慎重に行われ、社会的に高いコストが払われている。

また、商標の利用者としては、他の利用者の商標との類似性に対して注意を払う必要があるが、類似度に応じた商標の分布を把握することは容易ではなく、侵害や模倣被害を予見するのが難しいという問題がある。

本プロジェクトでは、上記の課題に鑑み、商標の類似性の評価を深層学習を用いて定量的に行えるか調査する。また、類似度に応じた商標の分布を深層学習によって可視化できるか調査する。さらに、深層学習によって文字と図形とからなる商標を生成できるか調査する。

## 2. 方法

### 2.1 商標の類似性の定量的評価

VAEを用いることによって得られる潜在空間上の表現ベクトルのコサイン類似度を類似性の評価指標として用いた。エンコーダー及びデコーダーは、ボトルネック+3層のCNNとした<sup>1</sup>。エンコーダー側CNNのチャンネル数は、1層目から3層目まで順に64、128、256とした。デコーダー側CNNは、エンコーダーとは逆順にアップサンプリングした。潜在空間の次元数は512とした。学習は、EUIPOが公開しているデータセット<sup>2</sup>を使用して行った。学習データ数及びテストデータ数は、それぞれ40000及び1000とし、エポック数は200とした。画像サイズは128\*128である。

### 2.2 商標の分布の可視化

テストデータの商標画像を訓練済みVAEのエンコーダーに入力した。出力として得られた512次元の表現ベクトルの分布を、t-SNE<sup>3</sup> (t-distributed Stochastic Neighbor Embedding) を使って2次元空間に写像して可視化した。

### 2.3 商標の生成

GAN (WGAN-GP<sup>4</sup>) を用いて、商標の生成を試みた。ただし、WGAN-GPの構造は簡易的にした；潜在空間の次元は200、Generatorは5層のLinear層、Discriminatorは3層のLinear層を用いた<sup>5</sup>。訓練時は、使用データ: グレースケール画像15000枚、画像サイズ: 128\*128, エポック数: 2000とした。

## 3. 結果

### 3.1 商標の類似性の定量的評価

まず、VAEでの学習が行えているか確認するために、Reconstruction画像（200 epoch時）をFig.1に示す。Fig.1に示すように、Input画像（テストデータ）とReconstruction画像とは近似しており、VAEによる学習が行えているものと考えられる。

商標の類似性の定量的評価が行えるか確認するために、コサイン類似度の範囲ごとに、当該範囲にコサイン類似度が入る画像のペアをFig.2からFig.9に示す。また、それらの画像のペアの特徴をTable 1に示す。

### 3.2 商標の分布の可視化

512次元の潜在空間における表現ベクトルの分布をt-SNEによって2次元空間に写像したものをFig.10に示す。2次元空間上で距離が近い商標同士を比較してみたが、類似性は認められなかった。

### 3.3 商標の生成

WGAN-GPによる生成画像をFig.11に示す。丸い輪郭や白い背景を生成できているものもあるが、明確な図形・文字が生成されていない。比較のために、VAEによるRandom sampling画像（200 epoch時）も合わせて示している。VAEによるRandom sample画像は画像全体に白い部分が広がっている物が多い。

## 4. 考察

### 4.1 商標の類似性の定量的評価

色彩のみが異なるものについてはコサイン類似度が高く、文字や図形の違いが大きくなるにつれてコサイン類似度が低くなっており、人間による類似性の評価とコサイン類似度による類似性の評価がかなり近いものになることが分かった。商標の特徴が、潜在空間における表現ベクトルに埋め込まれているものと推察され、表現ベクトルの各次元がどういった特徴に対応しているかを調べることによって、さらに詳細な類似性の定量的評価が可能なのではないかと考えられる。

### 4.2 商標の分布の可視化

t-SNEを用いることによって、潜在空間におけるデータ点同士の距離を維持しつつ2次元空間への写像を試みたが、潜在空間の次元数が512と比較的高いために上手くいかなかったものと考えられる。今回用いた商標は、文字や図形が入り混じった複雑なものが多く、潜在空間における表現ベクトルの分布がかなり複雑になっていると予想され、そうしたことも上手くいかなかった要因の一つと考えられる。

### 4.3 商標の生成

定性的評価の結果、GANにより大まかな輪郭や背景はある程度生成できる場合が見られる。明確な原因は特定できていない。改善案として、ネットワーク構造を複雑にする・訓練を増やすなどが考えられる。いくつかモデルを作成した上で定量的指標で評価を行うことも必要である。



Fig.1 The left is a set of input images and the right is a set of reconstructed images



Fig.11 The left is a set of images generated by GAN and the right is a set of images generated by VAE

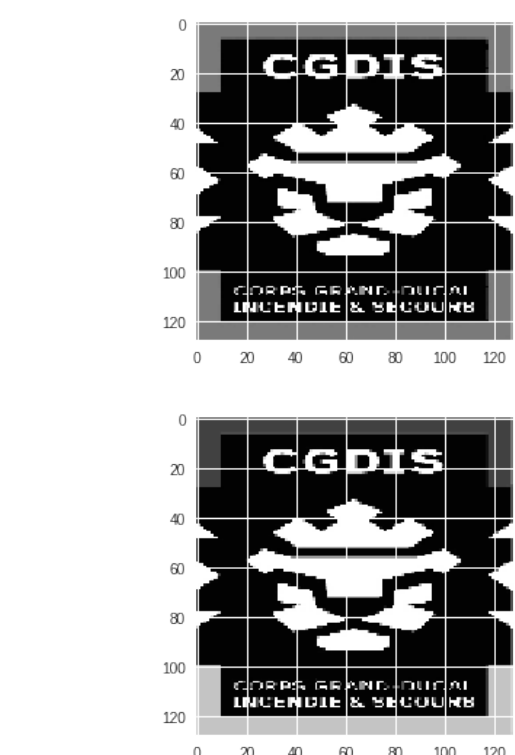


Fig.2 0.8≤cos

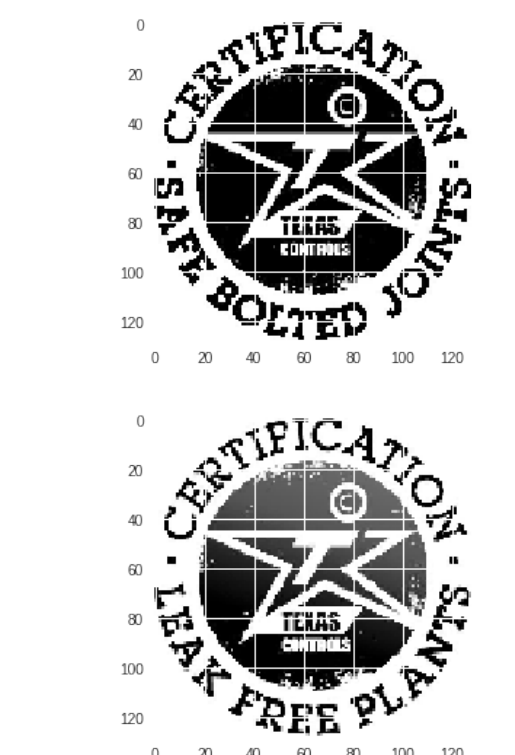


Fig.3 0.7≤cos<0.8

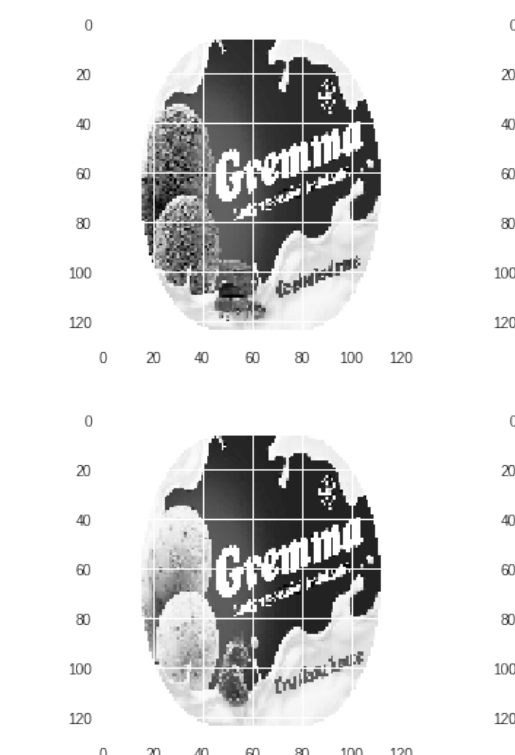


Fig.4 0.6≤cos<0.7

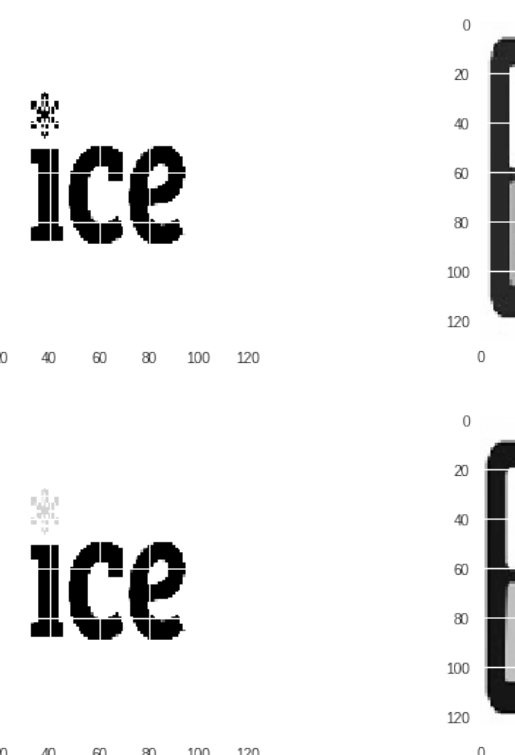


Fig.5 0.5≤cos<0.6



Fig.6 0.4≤cos<0.5

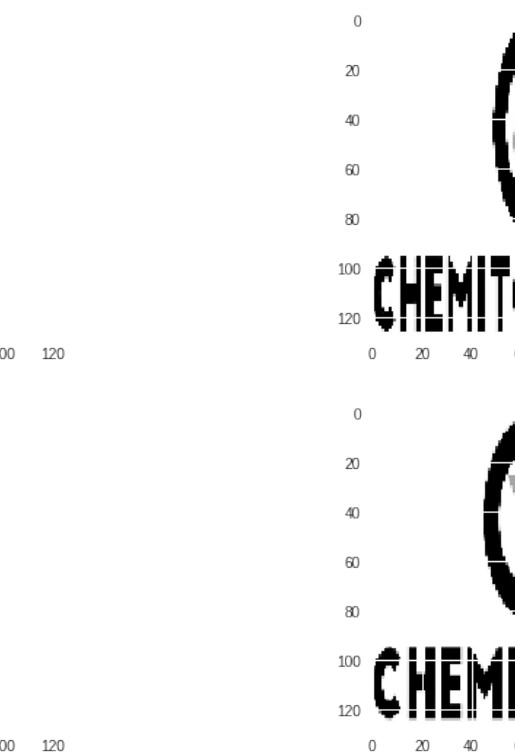


Fig.7 0.3≤cos<0.4

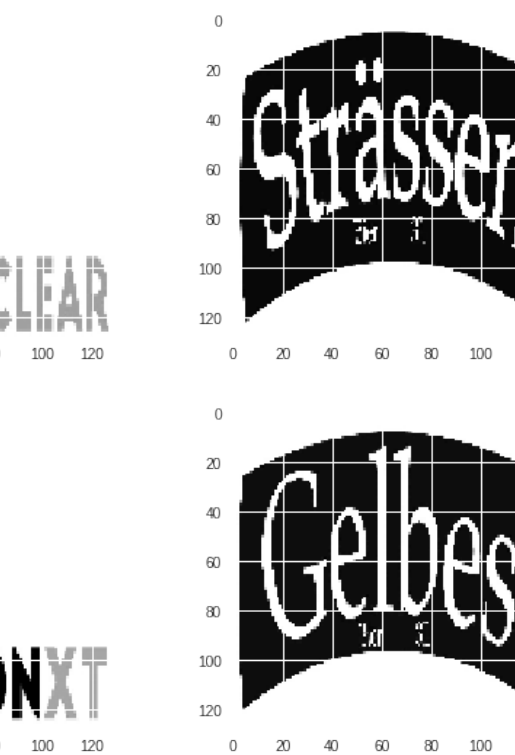


Fig.8 0.2≤cos<0.3

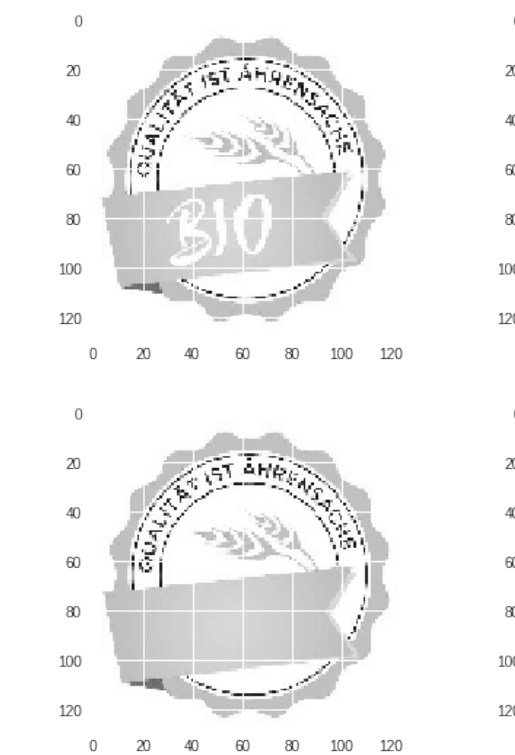


Fig.9 0.1≤cos<0.2

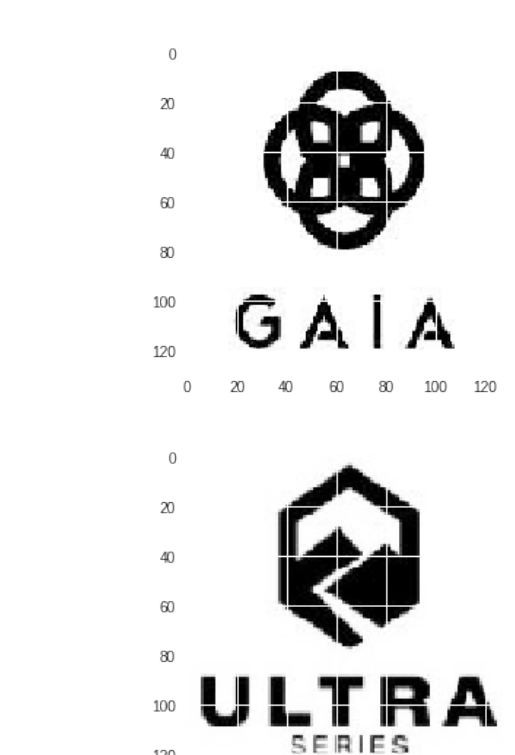


Fig.10 0.0≤cos<0.1

cos	特徴
0.8≤cos	色彩のみが若干異なる
0.7≤cos<0.8	0.8≤cosと比較して色彩の違いが若干大きい
0.6≤cos<0.7	0.8≤cosと比較して色彩の違いが大きい
0.5≤cos<0.6	色彩が大きく異なる。文字が一部異なるものもある。
0.4≤cos<0.5	一部の図形が少し異なる
0.3≤cos<0.4	一部の文字や図形が大きく異なる
0.2≤cos<0.3	外形はかろうじて似ているものの、内容はかなり異なる
0.1≤cos<0.2	印象は似ているが、類似している要素がほとんどない

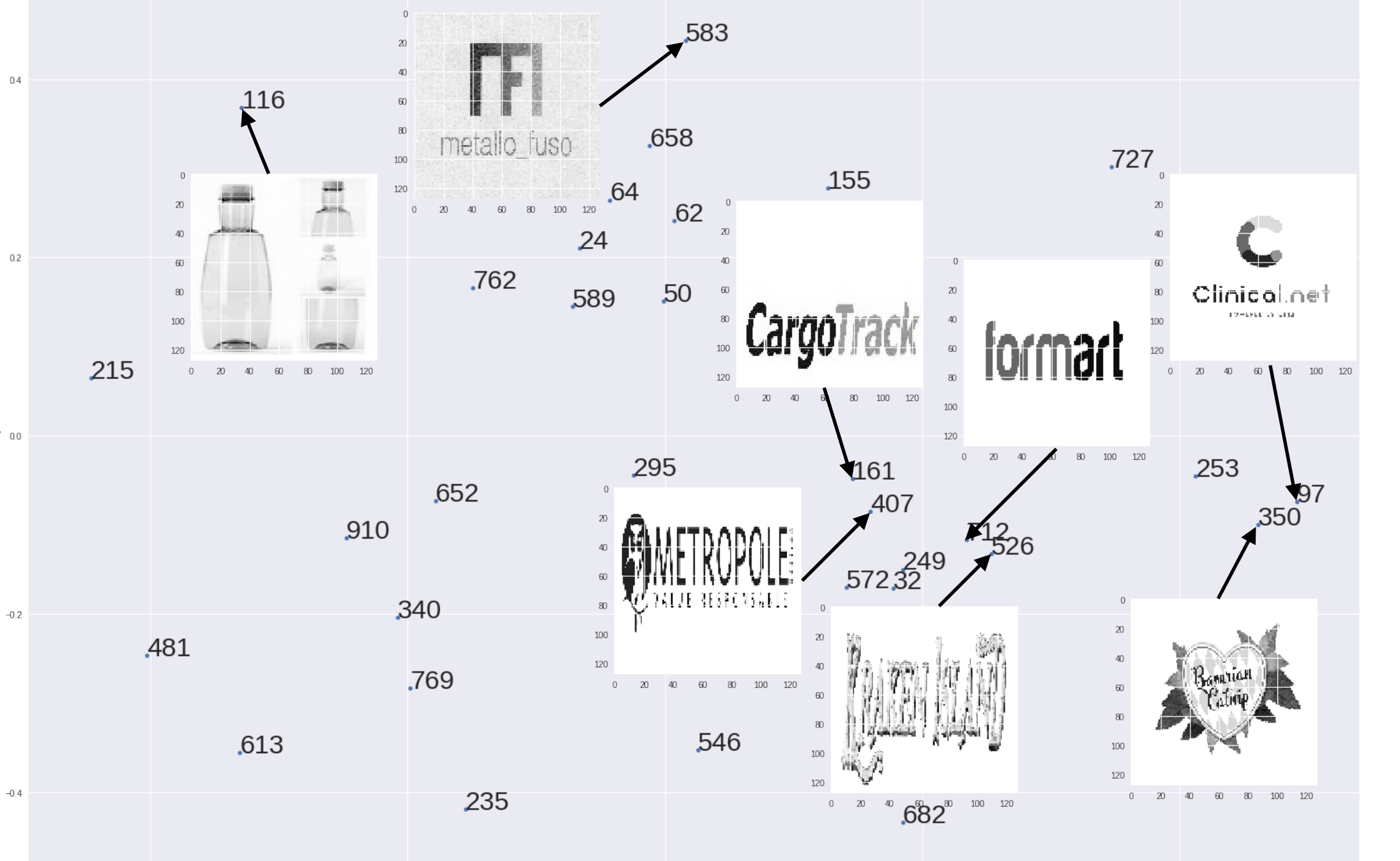


Fig.10 Distribution of trademarks obtained by t-SNE

## 5. 結論

### 5.1 商標の類似度の定量的評価

潜在空間における表現ベクトル同士のコサイン類似度を商標の類似性の定量的な評価指標として用いることができる可能性があることが分かった。

### 5.2 商標の分布の可視化

潜在空間における表現ベクトルを、表現ベクトル同士の類似度を維持しつつ低次元に写像する方法を検討する必要がある。

### 5.3 商標の生成

少なくとも輪郭や背景については、ある程度生成できる可能性があることが分かった。

## 6. 参考文献・実装

- VAEの実装: <https://gist.github.com/koshian2/64e92842bec58749826637e3860f11fa>
- 商標データ: <https://euiipo.europa.eu/ohimportal/en/open-data>
- L. van der Maaten, G. Hinton, Visualizing Data using t-SNE. <http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>
- I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved Training of Wasserstein GANs. <https://papers.nips.cc/paper/7159-improved-training-of-wasserstein-gans.pdf>
- WGAN-GPの実装: [https://github.com/eriklindernoren/PyTorch-GAN/blob/master/implementations/wgan\\_gp/wgan\\_gp.py](https://github.com/eriklindernoren/PyTorch-GAN/blob/master/implementations/wgan_gp/wgan_gp.py)