

感性語を考慮した敵対的生成ネットワーク を用いたロゴマーク生成

安藤 祐斗

富山県立大学 情報基盤工学講座
t815008@st.pu-toyama.ac.jp

July 16, 2021

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

背景

企業やブランドを象徴するロゴマークの現在主流の作成方法は、デザイナーによるもので、金銭面・時間面ともにコストが大きいという問題点がある。そこで、GAN(Generative Adversarial Networks) を用いてロゴマークの色情報をもとにロゴマークを自動生成する技術は存在するが、デザインを考える際に使いやすいものとは言えない。

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

目的

本論文では, ロゴマークのデザインを行う際にイメージを表すキーワードが重要であると考え, "かわいい" や "モダン" などの指定した感性語を元にロゴマークを生成するシステムを提案する. また, 学習する際に必要な, ロゴマーク画像とそれに対応した感性語の大量のデータを取得することは非常に困難であるため, 転移学習の考え方を採用し, 家具画像・感性語のデータセットを作成する.

はじめに

関連研究

ロゴマーク・感性語データセットの作成

ADDA による転移学習

ロゴマーク生成ネットワーク

生成例と評価実験

まとめ

GAN とは

生成器と識別器を同時に競わせて学習する.

損失関数 L_{adv} は, 識別器を D , 学習データ画像を x , 生成器を G , 乱数を z , 生成画像を $G(z)$ とすると下の式で表される.

$$\min_G \max_D L_{adv} = \log D(x) + \log(1 - D(G(z))) \quad (1)$$

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

Sage らの研究

Large Logo Dataset の作成

ロゴマーク画像のクラスタリングを行い, その結果得られたクラスターを教師ラベルとして扱うことで, これらラベルを入力としてロゴマーク生成を実現させた.

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

手順

- (1) デザイナーズ家具通販サイトより家具画像・感性語組を取得
- (2) Adversarial Discriminative Domain Adaptation(ADDA) を用いてロゴマーク画像に感性語ラベルを付与

ADDA

GAN の敵対的に学習させるという学習手法を用いた転移学習手法

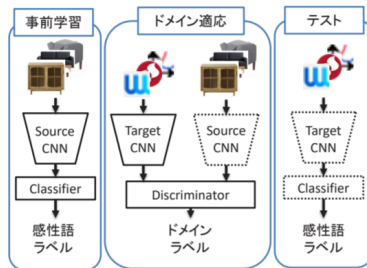


図 1: ADDA による転移学習の流れ

はじめに
関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

デザイナーズ家具通販サイト：FLYMEe

3 枚の画像と 1 文の説明文を取得し、1 枚の画像につき最大 5 個の感性語ラベルを付与した。

感性語の抽出

形態素解析ツール：MeCab

不適切である、単語となりえないもの (例: ザイン)、物体の名称を表す単語 (例: 商品) などを削除

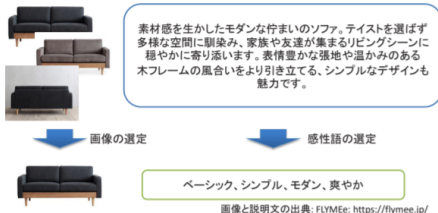


図 2: 家具画像・感性語データセット作成の流れ

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

(1) 事前学習：ソースドメインにおける X_s , 正解ラベル Y_s を畳み込みニューラルネットワーク (Convolutional Neural Networks, CNN) を用いて学習させる。

ソース CNN の最終層における特徴空間を $M_s(x_s)$, 最終出力を $C(M_s(x_s))$, クラス数を K とすると, 損失関数 L_{cls} は次式で表される。

$$\min_{M_s, C} L_{cls} = -\sum_{k=1}^K y_{s,k} \log C(M_s(x_s)) \quad (2)$$

(2) ドメイン適応：ソースドメインとターゲットドメインの CNN 最終層における特徴空間が類似するように学習させる。

ターゲットドメイン画像を X_t , ターゲット CNN の最終層における特徴空間を $M_t(x_t)$ とすると, 損失関数 L_{adv} は次式で表される。

$$\min_{M_t} \max_D L_{adv} = \log D(M_s(x_s)) + \log(1 - D(M_t(x_t))) \quad (3)$$

(3) テスト：ドメイン適応フェーズで学習させたターゲットドメインの CNN モデルを使用し, ターゲット画像にラベルを付与する。

マルチラベルに対応した CNN

1 枚の画像につき正解ラベルを最大 5 個付与したマルチラベルデータセットが使えるように, Binary Relevance 法という手法を使い, 従来の CNN をマルチラベルに対応するよう改良.

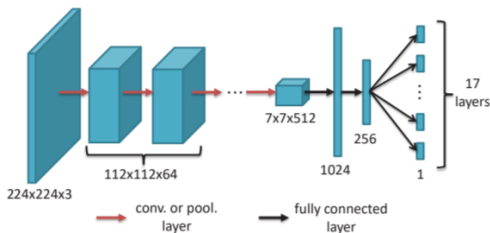


図 3: マルチラベルに対応させた CNN モデル

提案手法の ADDA の学習安定化

10/19

学習安定化のため, 以下 2 点を損失関数に変更を加えた.

- Wasserstein GAN モデル (WGAN) の使用
- Feature Matching の導入

WGAN は損失関数に Wasserstein 距離を用いたモデル

$$\min_G \max_D L_{adv} = D(x) - D(G(z)) \quad (4)$$

Feature Matching はソースターゲットドメイン画像を入力とした場合の識別器の中間層動詞の L2 ロスを生成器の損失項に加える手法. 識別器の中間層における出力を関数 f とすると次式で表される.

$$\min_G L_{fm} = \|f(x) - f(G(z))\| \quad (5)$$

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

提案されるロゴマーク生成ネットワーク

11/19

識別器が入力画像の真偽判定に加え、クラス分類も行うモデルである ACGAN 構造が用いられた。クラス分類における損失 L_{cls} を考慮し、また、通常の GAN の損失 L_{adv} に gradient penalty を導入する c を入力ラベル、係数 λ 、生成画像と学習画像の任意の中間画像を \hat{x} とすると、損失関数は以下の式で表される。

$$\min_D \max_G L_{adv} = D(G(\mathbf{z})) - D(\mathbf{x}) + \lambda [\|\nabla_{\hat{x}} D(\hat{x}) - 1\|^2] \quad (6)$$

$$\min_{G,D} L_{cls} = - \sum_{k=1}^K (y_k \log D(\mathbf{x}) + y_k \log D(G(\mathbf{c}, \mathbf{z}))) \quad (7)$$

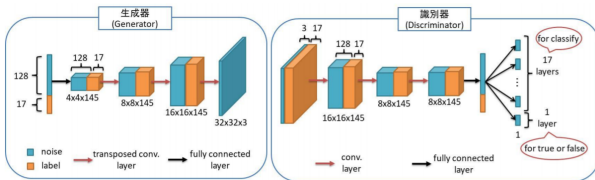


図 4: 提案するロゴマーク生成ネットワークの構成

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

ロゴマークに適した感性語抽出に関する実験

12/19

ロゴマークの感性語として適している感性語の選択を行うための予備実験

被験者 20 代の男女 17 名に家具画像・感性語データセットのうち画像枚数の多い上位 50 単語に対してロゴマークに適しているかどうかを「当てはまる」「どちらとも言えない」「当てはまらない」の 3 段階で評価を行った。

3 段階の選ばれる確率がともに 1/3 であるという帰無仮説に基づき、カイ二乗検定を行い、有意水準 5 %において有意差が得られたものが表 1

表1 予備実験において有意差が得られた感性語

感性語	スコア	感性語	スコア	感性語	スコア	感性語	スコア
モダン	16	シック	13	美しい	12	多彩	10
かわいい	16	ゴージャス	13	高級	12	自然	9
シンプル	15	エレガント	13	クラシック	12	楽しい	8
シャープ	15	ユニーク	13	綺麗	11	新鮮	5
明るい	15	爽やか	13	上品	11	心地よい	3
ポップ	15	クール	13	カジュアル	11	フリー	1
ダーク	14	スマート	13	力強い	11	丈夫	1
ダイナミック	14	奇抜	13	不思議	11		

上位 24 単語を選択し、今後の実験に用いた。

はじめに

関連研究

ロゴマーク・感性語データセットの作成

ADDA による転移学習

ロゴマーク生成ネットワーク

生成例と評価実験

まとめ

事前学習

家具画像・感性語データセットを用いて CNN の学習をする

感性語クラス：24 クラス, 全画像数：14050 枚

学習枚数：13500 枚, ミニバッチサイズ：32, 学習回数:50 回

top1 精度：66.3 %, top5 精度：90.6 %

ドメイン適応

家具ドメイン CNN モデルを用いてロゴマークドメイン CNN の学習をする

ロゴマーク画像データセット：LLD-icon データセット

学習枚数：13500 枚, ミニバッチサイズ：16, 学習回数：50

テスト

ドメイン適応フェーズにより学習したロゴドメイン CNN モデルを用いてロゴマーク画像に対して感性語のラベリングをする。

ロゴマーク全画像 485377 枚に対してテストを行い, 各画像に対して出力上位 5 単語を取得



	感性語	予測 確率	感性語	予測 確率	感性語	予測 確率
top1	モダン	0.999	かわいい	0.996	ユニーク	0.950
top2	シンプル	0.000	美しい	0.003	美しい	0.045
top3	シック	0.000	綺麗	0.000	明るい	0.001
top4	ダーク	0.000	ゴージャス	0.000	奇抜	0.000
top5	美しい	0.000	爽やか	0.000	モダン	0.000

図 5: 転移学習後のロゴマーク画像とその出力例

前スライドにおいて感性語ラベルを付与したロゴマークデータセットのうち
画像枚数が 1000 枚以上である 17 クラスを選択しロゴマーク生成を行った。
学習画像枚数：485377 枚, バッチサイズ：64, 学習総数：200000 イテレー
ション

表2 ロゴマーク・感性語データセットにおける感性語と画像枚数

感性語	枚数	感性語	枚数
シンプル	230,426	ポップ	11,035
美しい	133,485	ユニーク	9,901
モダン	57,513	ダーク	6,509
シック	45,299	上品	6,252
かわいい	37,689	高級	4,881
ゴージャス	23,184	シャープ	2,576
エレガント	22,379	カジュアル	2,210
綺麗	14,436	爽やか	1,709
明るい	12,458		



図 6: 作成したロゴマーク・感性語データセットの一部

ロゴマーク生成結果

15/19

上半分が ACGAN ネットワークを使用した場合, 下半分が Sage らのネットワークを使用した場合の感性語によるロゴマーク生成



図 7: 入力となる感性語ラベルと生成画像 (上: ACGAN ネットワーク, 下: Sage らのネットワーク)

はじめに
関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

MS-SSIM 指標を用いた Diversity スコアを評価指標として定量評価を行った。

Diversity スコアは低いほど多様性があるといえる→低いほど良い



図 8: Sage らの研究における生成画像（上：ACGAN ネットワーク, 下：Sage らのネットワーク）

表3 MS-SSIM指標を用いた Diversity スコア

モデル	Diversity スコア
Sageらの研究 [5]	0.056 ± 0.102
Sageらの研究 (ACGAN) [5]	0.057 ± 0.103
提案手法 (Sageらのネットワーク)	0.162 ± 0.183
提案手法 (ACGAN)	0.152 ± 0.174

ロゴマーク生成用データセットを用いて感性語予測の CNN モデルを構築し、生成したロゴマークについて感性語の推定を行い、識別率を算出した。また、感性語予測 CNN モデル学習に用いていないテスト用画像 5000 枚についても識別率を算出した。

ロゴマーク画像：480000 枚、ミニバッチサイズ：32、学習回数：5 回、評価指標：topK 精度

表4 CNNを用いた感性語識別率

画像	top1 [%]	top5 [%]
テスト用画像	59.6	94.1
提案手法 (Sageらのネットワーク)	29.7	71.4
提案手法 (ACGAN)	29.5	72.0

はじめに

関連研究

ロゴマーク・感性
語データセットの
作成

ADDA による転
移学習

ロゴマーク生成
ネットワーク

生成例と評価実験

まとめ

入力した感性語に即したロゴマーク画像を生成できているかどうかを検証するため主観評価実験を行った。

各クラス 84 枚の ACGAN モデルにおける生成画像について実験した。被験者に生成画像と入力感性語を提示し、以下の 2 項目について評価を行った

- (1) 生成画像が何らかの形・模様を表しているか
 - ・ 0 (当てはまらない) と 1 (当てはまる) の 2 段階
- (2) 感性語が表すイメージに適しているか
 - ・ 1 (当てはまらない) から 5 (当てはまる) の 5 段階

表5 主観評価実験の結果

評価項目		評価値平均
何らかの形・模様を表しているか (**)		0.85
感性語が表すイメージに適しているか	シンプル (**)	3.62
	美しい (**)	3.45
	モダン (**)	3.95
	シック (*)	3.19
	かわいい (**)	3.74
	ゴージャス	2.99
	エレガント (*)	3.35
	綺麗 (**)	3.51
	全感性語 (**)	3.48

(*) 有意水準 5% で有意差, (**) 有意水準 1% で有意差

本論文の貢献

- 家具通販サイトを用いた家具画像・感性語データセットの作成
- 転移学習手法 ADDA を用いたロゴマーク画像・感性語データセットの作成
- 感性語ラベルを考慮したロゴマーク生成

考察

- 感性語が与えるイメージを反映したロゴマークが生成可能である
- 既存研究には劣るものの多様性の画像を生成できている
- 画像・感性語データセットの学習方法の改良