

# 畳み込みニューラルネットワークを用いたビール缶パッケージの 好意度予測及び要因の可視化

## Prediction of Favorability Rating on Beer-Can Package Designs Using Convolution Neural Network and Visualization by Class Activation Mapping.

篠原裕之<sup>\*1</sup>  
Hiroyuki Shinohara

石黒達治<sup>\*1</sup>  
Tatsuji Ishiguro

中村遵介<sup>\*2</sup>  
Shunsuke Nakamura

山崎俊彦<sup>\*2</sup>  
Toshihiko Yamasaki

<sup>\*1</sup> キリン株式会社  
Kirin Company, Limited

<sup>\*2</sup> 東京大学  
The University of Tokyo

A quantitative survey of favorability rating by multiple panels is the main method to decide package designs. However, it is expensive and there is a limitation in the number of designs that can be subjected to a single survey. Therefore, this study aims at predicting the panel evaluation from the past survey results by a convolution neural network and visualize the important features by Grad-CAM. As a result, it has been made possible to give prescreening test to package design and suggestions of important features to the designers.

### 1. 背景と目的

最終消費者の購買行動において商品のパッケージデザインは重要である [石井 10]。その為、企業において商品を上市する際には、パッケージデザインに関する消費者調査を行い、良好な結果が得られたデザインを採用する事が多い。しかし、十分な消費者パネルを集めた調査には多額の費用が必要である、一度の調査にかけられるパッケージデザインの数に上限があるという課題がある。

一方で、近年はディープラーニングを用いて審美性評価や関する研究が盛んにおこなわれている [Talebi 17] [Kong 16]。これらの例から適切なディープラーニングのネットワークモデルを用いればパッケージデザインに関する消費者調査結果を予測する事が可能であると考えられた。但し、これらの研究の多くは豊富なデータを用いてモデルを構築しており、単独の企業で有する限られたデータのみを用いて消費者調査結果を高精度で予測可能なモデルを構築する事は困難であることが予想された。

そこで本研究は以下の 2 点を実現する事を目的とした。即ち、未調査のパッケージデザインの中で好意度に対して好影響を及ぼすと予測される要素と悪影響を及ぼすと予測される要素を可視化する事、未調査のパッケージデザインを好意度調査に供した場合の結果を予測する事である。

### 2. 対象データ

本研究で対象としたデータはキリンビール株式会社が 2001 年から 2017 年にかけて実施した、350mL ビール缶パッケージの好意度調査である。

調査は新規製品の上市あるいは既存製品のリニューアルに際し、製品候補の 350mL ビール缶パッケージを非明示で 64 人から 300 人の消費者パネルに提示し、それに対する好意度を表 1 に示した 5 段階もしくは 7 段階で回答を得たものである。この調査結果について、回答結果を最高 10 点、最低 1 点となるように点数化し、全回答者の点数の平均好意度をパッケージデザインの好意度とした。好意度のヒストグラムを図 1 に示す。

表1 回答方法及び点数化方法

| 7段階評価     |      | 5段階評価   |       |
|-----------|------|---------|-------|
| 回答        | 点数   | 回答      | 点数    |
| 非常に好き     | 10点  | 非常に好き   | 10点   |
| 好き        | 8.5点 | やや好き    | 7.75点 |
| やや好き      | 7点   | どちらでもない | 5.5点  |
| どちらでもない   | 5.5点 | やや嫌い    | 3.25点 |
| あまり好きではない | 4点   | 非常に嫌い   | 1点    |
| 好きではない    | 2.5点 |         |       |
| 全く好きではない  | 1点   |         |       |

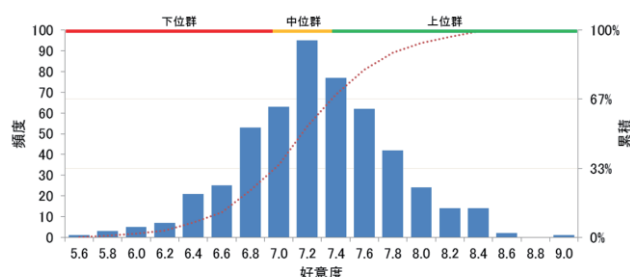


図1 好意度のヒストグラム



図2 パッケージ画像例

また、解析には調査に供した 350mL ビール缶パッケージの正面画像を用いた (図 2)。調査対象となったパッケージ画像は 509 種類であり、上位 33%(167 枚)を上位群、中位 34%(175 枚)を中位群、下位 33%(167 枚)を下位群とした。

連絡先: 篠原裕之, キリン株式会社基盤技術研究所, 神奈川県横浜市金沢区福浦 1-13-5, 080-2370-9301, Hiroyuki.Shinohara@kirin.co.jp

尚、調査は非明示であるものの、当該製品が既存製品のリニューアルである場合には消費者がその製品を既知である事が好意度に影響する事が予想された。また、当該製品が季節限定品である場合には販売する季節が好意度に影響することが予想された。そこで、当該製品が①新規製品であるか/②既存製品のリニューアルであるか/③春季限定製品であるか/④夏季限定製品であるか/⑤秋季限定商品であるか/⑥冬季限定商品であるかを追加の製品属性として用いた。

3. 手法

本研究の目的に従い、以下の検討を行った。本項ではそれぞれについて詳細を記す。

3.1 上位群下位群 2 クラス分類及び要因の可視化

限られた調査結果でモデルを構築する必要がある事から、VGG16 [Simonyan 14]をファインチューニングすることにより好意度調査上位群及び下位群について 2 クラス分類を行った。その際に分類に寄与した特徴量を可視化する事により、パッケージデザインの中で好意度に対して好影響を及ぼしている要素と悪影響を及ぼしている要素を可視化した。

具体的には、VGG16 の畳み込み層を用いて特徴量抽出を用い、得られた特徴量について Global Average Pooling [Lin 13]を行った後に新規に構築した全結合層に入力した。季節等の製品属性は 6 次元の配列として別途構築した全結合層に入力した。2 つの全結合層の出力を加算し、得られた特徴量を用いて 2 クラス分類を行った。モデル概要については図 2 に示す。また、その他解析条件詳細については表 2 に示す。

各クラス 25 枚をホールドアウト画像とし、残りの画像について 6 ラウンドのクロスバリデーションによりモデルのトレーニングを行った。学習回数は最大 250 エポックとし、その中でバリデーション画像の正答率が最も高かったエポックのモデルをそのラウンドの結果とした。

トレーニングの結果得られた 6 つのモデルを平均アンサンブルし最終モデルとした。最終モデルを用いたホールドアウト画像の分類の正答率からモデルを評価した。

寄与した特徴量の可視化については Grad-CAM [Selvaraju 17]を用いた。得られた 6 モデルそれぞれに Grad-CAM を適用し、得られた画像の平均画像を特徴量の可視化結果とした。また、得られた可視化結果の妥当性については麒麟株式会社内でビール缶パッケージをデザインしている担当者の有する知見との整合性を確認する事により評価した。

3.2 好意度の予測

3.2.1 回帰モデルの性能目標値設定

パッケージデザインの平均好意度と平均的な調査パネルの好意度との相関係数を回帰モデルの性能目標値とする事とした。しかし、好意度調査における個々のパネルの回答結果を入手出来なかった為、回答者による回答のばらつきを定量化する為の調査を行った。

今回の研究に用いたパッケージ画像の内 10 枚をビール事業に直接従事した事が無い麒麟株式会社従業員 23 名にランダムな順序で提示し、好意度について 7 段階で回答を得た。回答結果を最高 7 点、最低 1 点となるように好意度を点数化した。また、全回答者の点数の平均値をパッケージ画像の平均好意度とした。10 枚の画像に対する個々の従業員の好意度と平均好意度の相関係数を求め、得られた 23 の相関係数の平均値を回帰モデルの性能目標値とした。

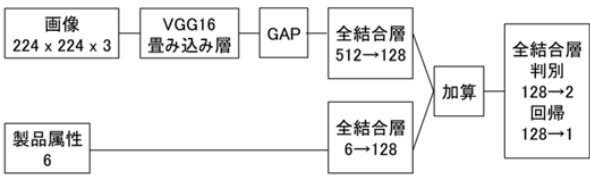


図3 モデル構造

| 表2 2クラス分類における解析条件詳細 |              |                      |
|---------------------|--------------|----------------------|
| 使用データ               | 入力① 画像       | 224x224x3            |
|                     | 入力② 製品属性     | 6                    |
|                     | 画像数          | 334枚 (167枚x2群)       |
|                     | ホールドアウト画像数   | 50枚 (25枚x2群)         |
| Data Augmentation   | 回数           | 9回                   |
|                     | 最大回転角度       | 15度                  |
|                     | 最大垂直移動距離     | 15%                  |
|                     | 最大水平移動距離     | 15%                  |
| 学習条件                | Optimizer    | Adam                 |
|                     | 損失関数         | binary cross-entropy |
|                     | エポック数        | 250エポック              |
|                     | 初期学習率        | 1.00E-06             |
|                     | 初期学習率保持エポック数 | 5エポック                |
|                     | 最終学習率        | 1.00E-09             |
|                     | 学習率変化率       | 0.97倍/エポック           |

| 表3 回帰における解析条件詳細   |              |                                          |
|-------------------|--------------|------------------------------------------|
| 使用データ             | 入力① 画像       | 224x224x3                                |
|                   | 入力② 製品属性     | 6                                        |
|                   | 画像数          | 509枚                                     |
|                   | ホールドアウト画像数   | 60枚 (20枚x3群)                             |
| Data Augmentation | 回数           | 9回                                       |
|                   | 最大回転角度       | 15度                                      |
|                   | 最大垂直移動距離     | 15%                                      |
|                   | 最大水平移動距離     | 15%                                      |
| 学習条件①             | Optimizer    | Adam                                     |
|                   | 損失関数         | 平均平方二乗誤差率<br>ただし、絶対誤差率が20%を下回った場合は0と見なす。 |
|                   | エポック数        | 50エポック                                   |
|                   | 初期学習率        | 1.00E-05                                 |
|                   | 初期学習率保持エポック数 | 5エポック                                    |
|                   | 最終学習率        | 1.00E-06                                 |
|                   | 学習率変化率       | 0.95倍/エポック                               |
| 学習条件②             | Optimizer    | Adam                                     |
|                   | 損失関数         | 実測値と予測値の相関係数を<br>1から減じた値                 |
|                   | エポック数        | 250エポック                                  |
|                   | 初期学習率        | 1.00E-06                                 |
|                   | 初期学習率保持エポック数 | 3エポック                                    |
|                   | 最終学習率        | 1.00E-08                                 |
|                   | 学習率変化率       | 0.98倍/エポック                               |

3.2.2 回帰モデルの学習条件及び評価方法

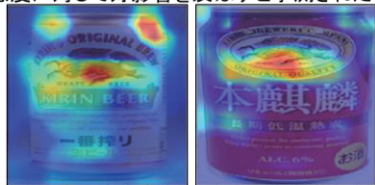
3.1 で検討したモデル構造の最終出力を変更し、回帰モデルとした。また、損失関数については平均平方二乗誤差率及び真値と予測値のピアソンの積率相関係数(以下、相関係数とする)を 1 から減じた値を用いた。その他解析条件詳細については表 3 に示す。

上位群、中位群、下位群各 20 枚をホールドアウト画像とし、残りの画像について 6 ラウンドのクロスバリデーションによりモデルのトレーニングを行った。はじめに損失関数を平均平方二乗誤差率とし、最大 50 エポックのトレーニングを行った。但し、平均平方二乗誤差率算出の際に、絶対誤差率が 20%を下回った画像の平方二乗誤差率は 0 と見なした。次に、得られたモデルについて、損失関数を真値と予測値の相関係数を 1 から減じた値に変更し、最大 250 エポックのトレーニングを行った。尚、前半の学習についてはバリデーション画像の平均平方二乗誤差率が最も低かったエポックのモデルを後半の学習に供した。後半の学習についてはバリデーション画像の真値と予測値の相関係数を 1 から減じた値が最も小さかったエポックのモデルをラウ

表4 2クラス分類結果

| ホールドアウト<br>画像分類の正答率 |       |
|---------------------|-------|
| 試行 1                | 82%   |
| 試行 2                | 82%   |
| 試行 3                | 82%   |
| 試行 4                | 76%   |
| 試行 5                | 76%   |
| 試行 6                | 68%   |
| 平均                  | 77.7% |
| 相対標準偏差              | 7.2%  |

好意度に対して好影響を及ぼすと予測された領域



好意度に対して悪影響を及ぼすと予測された領域



図3 ホールドアウト画像におけるGrad-CAMを用いた要因可視化結果例

ンドの結果とした。得られた 6 つのモデルを平均アンサンブルし最終モデルとした。

また、トレーニング時に用いる真値については、人間の評価値に直線性があるか不明であった為、①好意度 ②好意度を対数変換した値 ③好意度を 10 で除し、その値で 10 を累乗した値 の 3 条件の換算方法を比較する為、それぞれ 3 回トレーニングを行った。

最終モデルを用いてホールドアウト画像の好意度の予測を行い、②及び③についてはそれぞれの変換方法の逆関数を適用してから好意度との相関係数を算出しモデルを評価した。

## 4. 結果

### 4.1 上位群下位群 2 クラス分類及び要因の可視化

検討したモデルについて 6 回トレーニングを行った結果、ホールドアウト画像の分類の平均正答率は 77.7%、相対標準偏差は 7.2%であった(表 4)。また、その際の分類に寄与する領域の可視化結果(図 3)から、「麒麟のマークは好意度に好影響を及ぼす」「製品ロゴの英語表記は好意度に悪影響を及ぼす」といった仮説が得られた。これらはデザイン担当者の有する知見とも合致したことから、適切なモデルが構築できたと判断した。

### 4.2 好意度の予測

#### 4.2.1 回帰モデルの性能目標値設定

10 枚のパッケージ画像について個々のパネルの好意度と平均好意度との相関係数を求めた結果(表 5)、その平均値は 0.611 であった。この数値を回帰モデルの性能目標値とする事とした。

表5 各被験者の好意度と平均好意度の相関係数

| 相関係数   |        | 相関係数   |       |
|--------|--------|--------|-------|
| 被験者 1  | -0.140 | 被験者 13 | 0.723 |
| 被験者 2  | 0.034  | 被験者 14 | 0.746 |
| 被験者 3  | 0.177  | 被験者 15 | 0.759 |
| 被験者 4  | 0.235  | 被験者 16 | 0.763 |
| 被験者 5  | 0.436  | 被験者 17 | 0.782 |
| 被験者 6  | 0.474  | 被験者 18 | 0.782 |
| 被験者 7  | 0.536  | 被験者 19 | 0.802 |
| 被験者 8  | 0.576  | 被験者 20 | 0.879 |
| 被験者 9  | 0.652  | 被験者 21 | 0.887 |
| 被験者 10 | 0.673  | 被験者 22 | 0.925 |
| 被験者 11 | 0.705  | 被験者 23 | 0.933 |
| 被験者 12 | 0.706  | 平均     | 0.611 |

表6 好意度の変換方法の違いによる相関係数の差

| 真値     | 好意度   | 好意度を<br>対数変換した値 | 好意度を10で除し<br>その値で10を累乗した値 |
|--------|-------|-----------------|---------------------------|
| 試行 1   | 0.467 | 0.501           | 0.667                     |
| 試行 2   | 0.701 | 0.520           | 0.591                     |
| 試行 3   | 0.528 | 0.591           | 0.647                     |
| 平均     | 0.565 | 0.537           | 0.635                     |
| 相対標準偏差 | 21%   | 8.8%            | 6.2%                      |

表7 好意度予測結果

| 相関係数   |       | 平均絶対誤差率 |
|--------|-------|---------|
| 試行 1   | 0.667 | 24%     |
| 試行 2   | 0.591 | 35%     |
| 試行 3   | 0.647 | 16%     |
| 試行 4   | 0.542 | 23%     |
| 試行 5   | 0.629 | 39%     |
| 試行 6   | 0.641 | 17%     |
| 平均     | 0.620 | 26%     |
| 相対標準偏差 | 7%    | 36%     |

#### 4.2.2 好意度予測モデルの学習結果

トレーニング時に用いる真値を①好意度 ②好意度を対数変換した値 ③好意度を 10 で除し、その値で 10 を累乗した値とした条件について各 3 回のトレーニングを行いその結果を比較した(表 6)。その結果、好意度を 10 で除し、その値で 10 を累乗した値を用いた場合に最も良いモデルが得られた。そこで、当該条件のみさらに 3 回のトレーニングを行った結果、相関係数の平均値が 0.620、相対標準偏差が 7.3%であった(表 7)。性能目標値が相関係数 0.611 であったことから、適切なモデルが構築できたものと判断した。

尚、これらのモデルについてホールドアウト画像の好意度とその予測値の平均絶対誤差率を算出した結果、平均 26%、相対標準偏差 36%であった。

## 5. まとめ

本研究により、限られた数のパッケージデザインの好意度調査結果から作成したモデルを用いて、未調査のパッケージデザインの中から好意度に対して好影響を及ぼすと予測される要素と悪影響を及ぼすと予測される要素を可視化し、それをデザイナーにフィードバックする事と、複数のパッケージデザイン候補の中から実際の調査の対象にするパッケージデザインを絞り込むプレスクリーニングテストを行う事が可能になったと考える。一方で、好意度そのものを高精度で予測するには予測結果の誤差が大きく、更なる改善が必要であると考えられた。

## 参考文献

[石井 10] 石井裕明, 恩蔵直人:価値視点のパッケージ・デザイン戦略, マーケティングジャーナル 30(2), 31 - 43(2010)

- [Talebi 17] Talebi, H. Milanfar, P.: NIMA: Neural Image Assessment, *arXiv preprint arXiv*, 1709.05424, 2017
- [Kong 16] Kong, S. Shen, X. Lin, Z. Mech, R. Fowlkes, C.: Photo Aesthetics Ranking Network with Attributes and Content Adaptation, *arXiv preprint arXiv*, 1606.01621, 2016
- [Simonyan 14] Simonyan, K. Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv preprint arXiv*, 1409.1556, 2014
- [Lin 13] Lin, M. Chen, Q. Yan, S.: Network In Network, *arXiv preprint arXiv*, 1312.4400, 2013
- [Selvaraju 17] Selvaraju, R. Cogswell, M. Das, A. Vedantam, R. Parikh, D. Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization, *2017 IEEE International Conference on Computer Vision (ICCV)*, 618-626, 2017