

# 畳み込みニューラルネットワークを用いた 顔魅力要因の可視化と考察

## Visualization and Discussion of Facial Attractiveness Factors Using Convolutional Neural Network

佐野 貴紀<sup>†</sup>

Takanori Sano

<sup>†</sup>株式会社構造計画研究所

Kozo Keikaku Engineering Corporation

takanori-sano@kke.co.jp

### 概要

近年、顔の特徴を用いて魅力度を予測する研究が盛んに行われており、特に、畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用いた研究では高い精度での予測が達成されることが報告されている。また、構築した CNN モデルの隠れ層を可視化することで、予測に重要な特徴の確認が可能となる。本研究では、CNN を用いた隠れ層の可視化による顔魅力要因の調査を行った。その結果、本モデルから抽出された特徴は、心理学研究における知見とのいくつかの共通点が確認された。

キーワード：深層学習、畳み込みニューラルネットワーク、隠れ層、心理学

### 1. はじめに

「Beauty is in the eye of the beholder」(美に対する感じ方は人それぞれ) と長年の間考えられてきたが、数々の研究の中で顔の魅力については性別や文化間においても高い一貫性があることが明らかになっている[1]。顔の平均性[2]、シンメトリー性[3]、性ホルモンマーカ- [3] (いわゆる男らしさ、女らしさ) 等が人間の顔の魅力の知覚に影響を与える重要な要因であることが報告されているが、魅力知覚に影響する顔の特徴やその心理メカニズムの詳細は未だ明らかではない部分が多い。

近年では、これらの顔の特徴を用いて魅力度を予測する研究が機械学習の分野で盛んに行われており [4][5]、顔の魅力予測が化粧推薦や画像検索などの応用 [6] に有用であることが示されている。その中でも特に、畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用いた研究では高い精度での予測が達成されることが報告されている [7]。さらに、構築した CNN モデルの隠れ層を可視化することで、魅力予測に重要となる特徴について考察を行っている研究もある。例えば Xu et al(2015)[8]の研究では、CNN モデルを用いて顔魅力予測モデルを構築し隠れ層の可視化を行った結果、顔の輪郭や目と口が顔の魅力の予測

に重要な特徴として抽出され、これは人間の視覚認識の直感と一致することを提唱している。このように、CNN 等の予測モデルを用いることにより、顔魅力における定量的に観測可能な普遍性の高い特徴の抽出が可能となる。しかし、隠れ層の可視化による特徴の抽出は、先行研究では予測精度を高める手法の検討を目的にするものが多く、人の魅力知覚に影響する顔の特徴や魅力知覚メカニズムの解明を目的とした研究はあまり行われていない。

### 2. 目的

本研究では、CNN を用いて魅力予測モデルの構築及び隠れ層の可視化を行う。可視化結果を確認し、モデルが重要と判断している特徴とこれまでの心理学研究における知見との共通点や差異を調査することで、人の魅力知覚に影響する顔の特徴や魅力知覚メカニズムの理解を拡張することを目的とする。

### 3. 検証の手続き

検証は、ベンチマークデータベースの SCUT-FBP5500[9]を用いて行う。SCUT-FBP5500には、魅力得点が 60 人の評価者によってラベル付けされた男性画像 2750 枚、女性画像 2750 枚の顔画像が収録されている。

最初の検証では、構築する魅力予測モデルの精度検証を行う。Xu et al(2015)[8]の研究では女性顔データセットのみを用いていたが、本研究では、性差についても今後考察を行うことを念頭に、男女画像を別々に分割し、それぞれで検証を行う。検証では、2750 枚の画像の中から 2200 枚の画像をランダムに選択して学習データセットを作成し、残りをテストデータセットとする。データ分割による偏りを減らすために、すべての検証は 5 分割の交差検証を行う。また、先行研究

[10][11]にならない、予測性能を評価するための精度指標は、画像のラベル値と予測結果間のピアソン相関を用いる。さらに、本研究で構築する CNN モデルと既往手法による予測精度を比較するため、機械学習のベンチマークモデルとして Linear Regression(LR), Gaussian Process Regression(GPR), Support Vector Regression(SVR) を用い、それぞれの手法についても同条件で予測精度の算出を行う。

次の検証では、構築したモデルを用いて隠れ層の可視化を行う。可視化結果を確認し、モデルが重要と判断した特徴の結果と、心理学研究において人の魅力知覚に重要と提唱されている内容との共通点や差異を探り、魅力知覚における顔の特徴に関する考察を行う。

#### 4. 構築モデル

本稿で用いたモデルのネットワーク図を図 1 に示す。図中の各層の上数字はその層のサイズを表し、下の記

載はその層における処理を表す。各プーリング層と全結合層の直後に Dropout の処理を行っている。活性化関数は ReLU を使用し、ミニバッチサイズは 64 で固定し、1000epoch で学習を行った。

#### 5. 精度検証結果

CNN モデルを構築し、精度検証を行った結果、男性顔データセットを用いた検証ではテストデータに対して平均 0.771 の相関が達成され、女性顔データセットを用いた検証ではテストデータに対して平均 0.789 の相関が達成された。機械学習のベンチマークモデルとなる LR, GPR, SVR の予測精度は男性顔データセットの場合はそれぞれ 0.225, 0.356, 0.704 女性顔データセットの場合はそれぞれ 0.199, 0.354, 0.703 であり、本モデルは既往の機械学習手法と比較しても高い精度であった。5 分割の交差検証の結果を表 1 に示す。

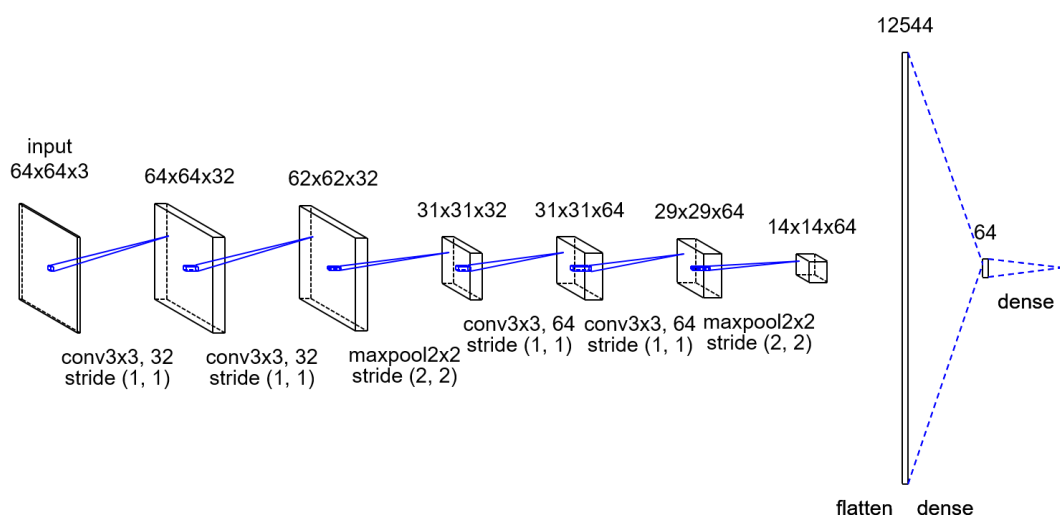


図 1 モデルネットワーク図

表 1 精度比較結果

model	sex	1	2	3	4	5	average
本モデル (CNN)	male	0.753	0.798	0.750	0.750	0.806	0.771
	female	0.792	0.787	0.794	0.807	0.790	0.794
SVR	male	0.743	0.730	0.696	0.676	0.672	0.704
	female	0.702	0.715	0.678	0.713	0.707	0.703
GPR	male	0.386	0.340	0.388	0.361	0.304	0.356
	female	0.342	0.314	0.361	0.369	0.386	0.354
LR	male	0.234	0.167	0.279	0.286	0.160	0.225
	female	0.206	0.155	0.249	0.185	0.197	0.199

## 6. 隠れ層の可視化結果

次に、全データで学習を行った本モデルを用いて、モデルが何の特徴を根拠に魅力値を予測していたのかを確認するため、CNNの予測判断根拠を視覚的に把握するのに優れた Gradient-weighted Class Activation Mapping (Grad-CAM)[12]を用いて出力層に最も近い隠れ層の可視化を行った。ここでは、男性および女性それぞれの魅力得点のラベル値が高い上位10枚の画像について確認を行った(図2)。その結果、男女ともに、目や眉の周りが活性化している傾向があるものがいくつか見られた。これは目と眉の魅力度は顔全体の魅力と有意に相関する[13]という過去の知見と一致する。また、魅力的な顔の中には、活性がほとんど見られないもの

もいくつか確認された。これは、魅力的な顔は平均的[2]であることが知られているため、魅力顔は特徴が少ないこととの関連が考えられる。また、魅力得点が高い顔画像の中には、顔全体が活性化している傾向があるものも見られた。これは、顔の色やテクスチャが魅力評価に影響を及ぼすこと[14]が関連している可能性が考察される。

これらのように、隠れ層の可視化によりモデルから抽出された特徴は、心理学研究における知見とのいくつかの共通点を確認された。本稿は特に魅力得点の高い画像に注目して分析を行ったが、今後、魅力の強度による差異や、各画像におけるモデルの予測精度等も考慮することで、より詳細の分析が可能になると考える。



図2 Grad-Camを用いた可視化結果

## 7. まとめ

本モデルを用いることで、男性顔データセットではピアソン相関0.771、女性顔データセットではピアソン相関0.794の精度での顔魅力の予測が達成され、機械学習のベンチマークモデルと比較しても高い精度であることが示された。さらに、隠れ層の可視化によりモデルから抽出された特徴は、心理学研究における知見とのいくつかの共通点が確認された。引き続き、他の隠れ層の可視化結果も踏まえ、モデルの処理過程についても考慮し考察を深めることにより、本研究で検討しているアプローチが人の魅力知覚に影響する顔の特徴の理解に加え、魅力知覚の心理メカニズムの解明の一助に繋がると考える。

## 文献

- [1] Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M., & Smoot, M., (2000) "Maxims or myths of beauty?: A meta-analytic and theoretical review", *Psychological Bulletin*, 126, 390–423.
- [2] Perrett, D. I., May, K. A. & Yoshikawa, S.,(1994) "Facial shape and judgments of female attractiveness: Preferences for non-average characteristics", *Nature*, 386, 239–242.
- [3] Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E., (2010) "Men's facial masculinity predicts changes in their female partners' sexual interests across the ovulatory cycle, whereas men's intelligence does not", *Evolution and Human Behavior*, 31, 412–424.
- [4] Mao, H., Jin, L., & Du, M., (2009) "Automatic classification of Chinese female facial beauty using Support Vector Machine", in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 4842–4846.
- [5] Zahang, D., Zhao, Q., & Chen, F., (2011) "Quantitative analysis of human facial beauty using geometric features", *Pattern Recognition*, 44, 4, 940-950.
- [6] Liang, L., Jin, L., & Li, X., (2014) "Facial skin beautification using adaptive region-aware mask", *IEEE Trans. on Cybernetics*, 44, 12, 2600–2612.
- [7] Xu, J., Jin, L., Liang, L., Feng, Z., Xie, D., and Mao, H., (2017) "Facial attractiveness prediction using psychologically inspired convolutional neuralnetwork (pi-cnn)", in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1657–1661.
- [8] Xu, J., Jin, L., Liang, L., Feng, Z., Xie, D., (2015) "A new humanlike facial attractiveness predictor with cascaded fine-tuning deep learning model", 2015
- [9] Liang, L., Lin, L. Jin, L., Xie, D., and Li, M., (2018) "SCUT-FBP5500: A diverse benchmark dataset for multi-paradigm facial beauty prediction," in *ICPR*, pp. 1598–1603.
- [10] Chiang, W., Lin, H., Huang, C., Lo, L., and Wan, S., (2014) "The cluster assessment of facial attractiveness using fuzzy neural network classifier based on 3D Moir features," *Pattern Recognition*, 47, 3, 1249-1260.
- [11] Yan, H., (2014) "Cost-sensitive ordinal regression for fully automatic facial beauty assessment", *Neurocomputing*, 129, 334-342.
- [12] Selvaraju, R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D., (2016) Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *arXiv preprint arXiv:1610.02391*, 1–24.
- [13] 加藤隆, 阿磨大介, 森岡久美子, & 赤松茂. (1998) "顔の魅力判断におけるパーツの魅力の影響", *信学技報 (HIP 97-53)*, 17-22.
- [14] Fink, B., Grammer, K., & Thornhill, R., (2001) "Human (Homo sapiens) facial attractiveness in relation to skin texture and color", *Journal of Comparative Psychology*, 115(1), 92–99.