

知能情報実験 iii（データマイニング班）
顔画像に基づく美男美女の識別と一般人との比較による
特徴抽出

215706D：KIM HYUNWOO, 235221E：山脇大輝,
235701B：松田遼平, 235732B：長瀬一生

提出日: 2025 年 7 月 18 日

目次

1	はじめに	1
1.1	実験の目的と達成目標（アプローチの全体像を含む）	1
1.2	意図していた実験計画との違い	1
2	実験方法	1
2.1	実験目的	1
2.2	データセット構築・前処理	1
2.3	モデル選定	5
2.4	パラメータ調整	5
3	実験	5
3.1	役割分担について	5
3.2	用意したデータセット画像	5
4	実験結果	6
4.1	custom model	6
4.2	resnet 18	7
4.3	resnet 34	7
4.4	resnet 50	8
4.5	EfficientNet	8
5	失敗事例分析について：Grad-Cam を用いたモデルの信頼性の検証	9
5.1	実験結果の考察	11
6	白背景画像を用いた失敗事例分析の改善	11
7	考察	11
8	反省・今後の課題	11
8.1	時間の都合上省いた項目	11
9	まとめ	12

概要

本グループでは、美男美女（有名人）と一般人を分類する機械学習システムを構築した。美男美女ランキングから対象者の顔写真をスクレイピングで収集し、一般人データとのバランスを図るため前処理を実施した。前処理として画像サイズ調整や顔位置の正規化を行い、分類器を構築した。実験では美男美女と一般人の顔画像の識別を試み、機械学習モデルから人間の美的感覚に近い判断をどの程度行えるかというデータを明らかにする。

1 はじめに

1.1 実験の目的と達成目標（アプローチの全体像を含む）

本グループでは「顔画像データを用いて、一般的に“美男美女”と称される有名人と一般人を識別・分類し、特徴差を明らかにすること」をテーマとして設定した。

人の美しさや魅力は主観的に表されることが多く、機械学習を用いて客観的に分析することを試みる。fairface のモデル構築の際に用いられた画像群と独自に美男美女の画像から構築したデータセットを用いて、画像認識のための深層学習モデルである ResNet18, ResNet34, ResNet50 と EfficientNet_b0 を用いて、美男美女（有名人）と称される顔写真と一般人の写真を識別・分類するモデルの構築を問題として設定した。最終的には、有名人と一般人の差から美しさや魅力度合いがどのように違うのかを定量化・可視化し、美的評価の要素を客観的に明らかにする。また、人の美しさは黄金比や対称性といった要素で評価されることが多いが、モデルが人間の美的感覚に近い判断をどの程度行えるかというデータを明らかにする。

1.2 意図していた実験計画との違い

fairface には、データセットと学習済みモデルの 2 種類が公開されているが、当初の調べ不足が原因でデータセットのみと勘違いして実験計画を立てていた。fairface では、すでに画像収集・前処理・ResNet34 による分析まで行い高精度を叩き出している。これらと同様のことを本実験で行っても車輪の再開発となり実験の意義は少ない。そこで fairface に追加学習を行うことにした。fairface に用意されているデータセットに加え、独自にデータセットを構築し、さらに有名人・一般人のラベル付をした画像を学習することで意義深い実験ができることを志した。

2 実験方法

2.1 実験目的

世界で美男・美女 (有名人) と呼ばれる人の顔写真と一般人の顔写真を分類するモデルを作り、美男・美女と一般人の分類を行う。

2.2 データセット構築・前処理

有名人と一般人を分類するモデルを構築するために、2 つのデータセットを準備し、前処理を行い、モデルの学習・評価を行う。

2.2.1 FairFace とは

本実験では FairFace というモデルを用いて、分類評価を行うと同時に、独自構築した分類モデルも作成した。ここで、FairFace というモデルについて概要を説明すると、従来の多くの顔画像データセットでは、特定の人種（白人）や性別にデータの偏りが見られる傾向にある。FairFace では、7つの人種グループ：白人、黒人、インド人、東アジア人、東南アジア人、中東人、ラテンアメリカ人でデータ数を均一して、人種差の少ないデータセット構築を行った。これにより、機械学習モデルの特定の人種・性別への偏り・バイアスの軽減を実現した顔画像データセット及びモデルである。構築されたデータセットについては、主に大規模なパブリックデータセット yahoo YFCC100M（Flickr 画像）から、意図的に人種バランスを考慮したサンプリング手法を用いて収集され、加えて Twitter やオンライン新聞からの画像も含まれている。YFCC100M 全体からランダムに顔画像をサンプリングし、各国の人口構成を推定し、データセットが特定（白人）の人種に偏らないように画像数を調節した。これにより他の人種を過小評価するデータセットのバイアスを軽減している。

2.2.2 定義

本実験では、以下のように一般人と有名人の定義を行う。

- 一般人：fairface のデータセットで作成されたデータセット。
- 有名人：2024 年度世界の美男美女ランキング上位 top50 を有名人として定義する。

2.2.3 データセット構築

- 一般人のデータセット：FairFace のデータセットを使用
- 美男・美女のデータセット：独自構築を行った。

(1) 美男・美女の基準を決定

[2] や [3] より 2024 年度美男美女ランキング top50 を美男美女として扱う

(2) データの収集

対象者の画像は、「2024 年度美男美女ランキング」top50 の人名から bing 検索エンジンを用いて web スクレイピングした。

* ソースコードを以下の GitHub にて公開する。

* https://github.com/e235221/info3dm_racial_classification

2.2.4 前処理

- リサイズ
 - ー 収集したデータの顔部分だけ切り取り、サイズを 300×300 とし、FairFace のデータセットと合わせて調整する。

- 正面・側面の判定
 - 顔の向きが学習に与える影響を排除するために、hopenet を用いて正面を向いている画像のみ抽出した。側面を向いている画像はデータセットから除外している。
 - HopeNet とは、1 枚の顔写真から Yaw(左右の向き), Pitch (上下の向き), Roll (傾き) を求め、その人がどの方向を向いているのかを推定する深層学習モデルである。感情認識などに用いられている。
- アップサンプリング・ダウンサンプリング
 - アップサンプリング：データ数の少ない美男美女データの各画像を左右反転させることでデータ数を 2 倍に増やした。
 - ダウンサンプリング：データ数の多い有名人データセットからランダムにデータを削除することで、有名人データセットと一般人データセットのデータ件数を調節した。
 - 画像数については、train で 31000 枚ずつ、test で 7800 枚ずつの画像を用意した。
- 美男美女・一般人データセットに対して、ラベル付け
 - web スクレイピングして収集した画像データに対して、fairface で用意されているラベル付けと同様に美男美女データセットに対してもラベル付けを行い、それに加えて一般人か有名人かを判別するためにそれぞれ 0 と 1 を付与した。
 - 新規作成した美男美女データセットについては、手作業で検索し、国籍・誕生日のラベルを付与した。
 - `file_name, age, gender, race, 0/1` の 5 種類を列名として csv を作成した。
`file_name` に指定された path で画像を読み込み、学習を行う。
- ファイルのリネーム

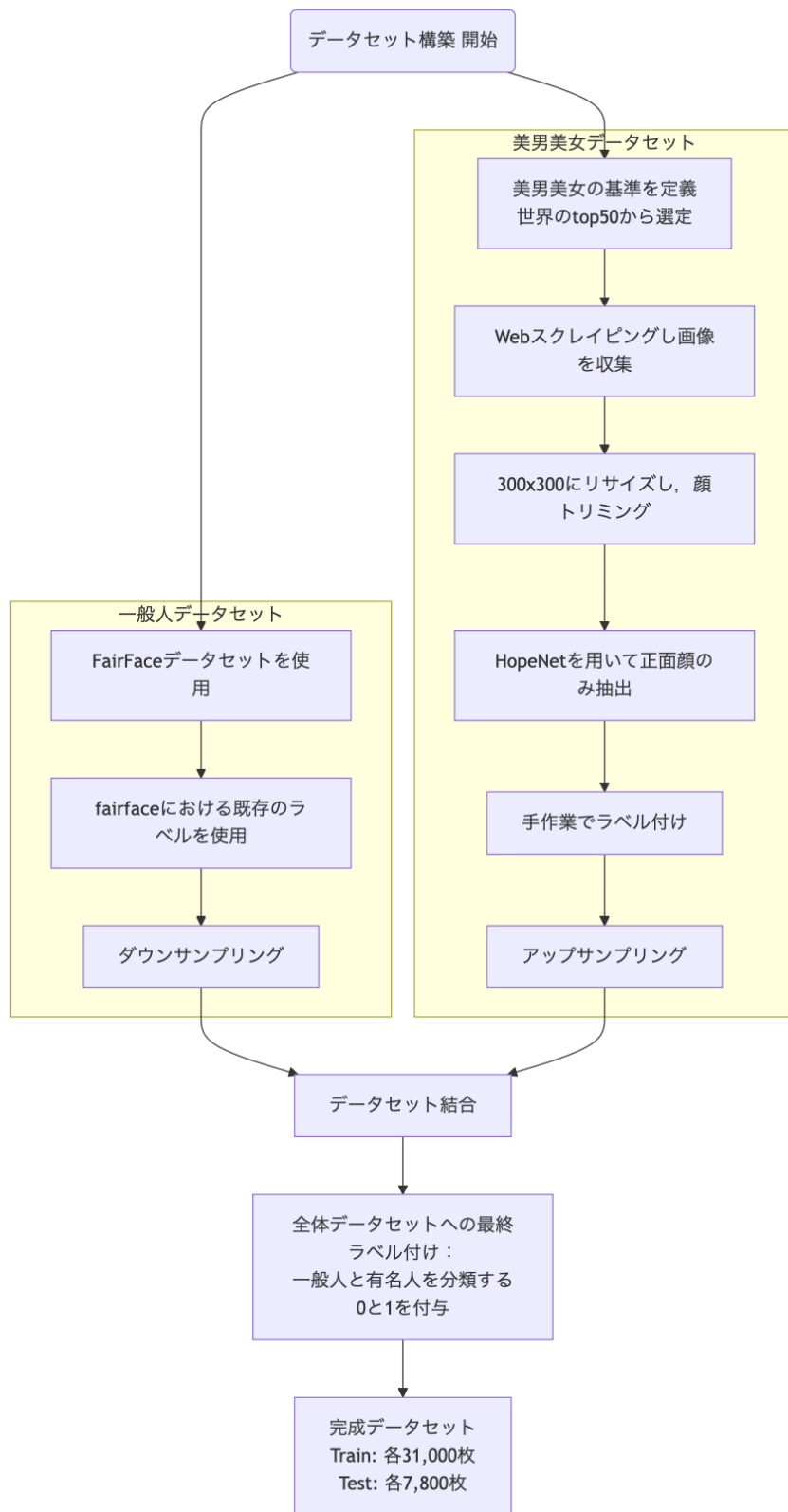


図 1 前処理のフローチャート

2.3 モデル選定

fairface が提供する学習済みモデルを本研究のタスクに合わせてカスタマイズした。これには ResNet34 が使用されていたため、そのまま ResNet34 を用いる。詳しくは [1] を参照されたい。

(後述するが、この fairface のカスタマイズモデルは精度が 100% になったため、別アプローチとして本実験では fairface を用いず新規に ResNet18, ResNet34, ResNet50 も用いた。)

2.4 パラメータ調整

デフォルト値のまま使用している。

3 実験

3.1 役割分担について

- 215706D : KIM HYUNWOO
 - 正面画像抽出・ラベル付・複数の ResNet と EfficientNet での実行・Grad Cam
- 235221E : 山脇大輝
 - FairFace の調査, カスタマイズ, Web スクレイピング・リサイズ・トリミング・ラベル付
- 235701B : 松田遼平
 - GitHub の操作補助
- 235732B : 長瀬一生
 - アップサンプリング・ダウンサンプリングコード作・背景のマスキング実施

3.2 用意したデータセット画像

図 2, 図 3 に用意したデータセットの画像の一部を示す。これは筆者がランダムに選択したものであり、レポートの可読性を向上させるために用意した。

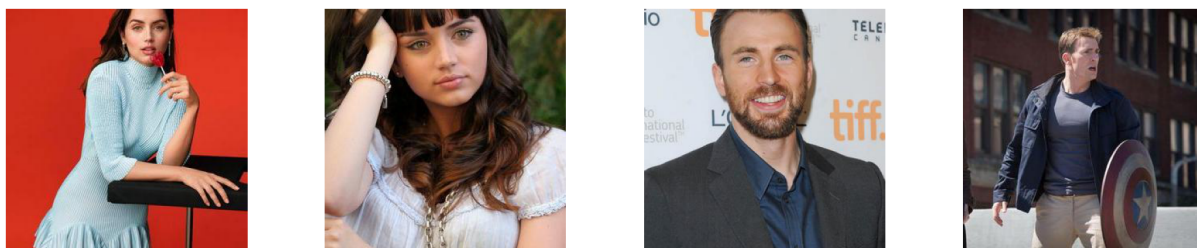


図 2 有名人データセットの画像例

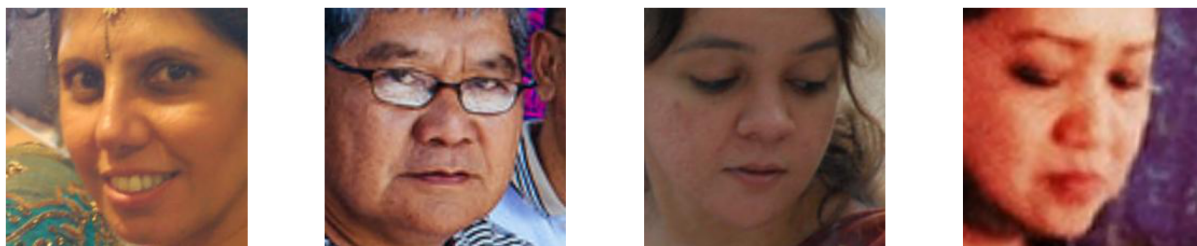


図 3 一般人データセットの画像例

4 実験結果

fairface に追加学習をすることによって有名人と一般人の差異を明らかにすることであった。しかし、fairface に追加学習したモデル (custom) は精度が 100% になった。過学習の影響も考えられるが、fairface はそもそも学習済みモデルであり、本実験はこれに追加学習を行うアプローチのため原因の切り分けが困難である。

そこで、アプローチとして ResNet18,34,50 と EfficientNet で新規に一般人と有名人の学習を行い、その結果を見る。

実行結果を表 1, 表 2, 表 3, 表??, 表 5 に示す。

4.1 custom model

表 1 Custom Model

Epoch	Train Accuracy	Validation Accuracy
1	99.48	100.00
2	100.00	99.99
3	100.00	99.99
4	99.97	100.00
5	100.00	100.00
6	100.00	100.00
7	100.00	100.00
8	100.00	100.00
9	100.00	100.00
10	100.00	100.00

考察：

4.2 resnet 18

表 2 ResNet-18

Epoch	Train Accuracy	Validation Accuracy
1	97.92	99.28
2	99.23	99.22
3	99.47	99.53
4	99.61	99.48
5	99.68	99.52
6	99.76	99.38
7	99.83	99.59
8	99.83	99.51
9	99.88	99.45
10	99.88	99.52

考察：

4.3 resnet 34

表 3 ResNet-34

Epoch	Train Accuracy	Validation Accuracy
1	97.74	99.08
2	99.27	99.34
3	99.40	99.27
4	99.55	99.50
5	99.59	99.40
6	99.65	99.27
7	99.72	99.31
8	99.80	99.49
9	99.79	99.56
10	99.84	99.47

考察：

4.4 resnet 50

表 4 ResNet-50

Epoch	Train Accuracy	Validation Accuracy
1	95.26	98.58
2	98.69	98.76
3	99.10	99.20
4	99.32	99.25
5	99.39	99.34
6	99.50	99.37
7	99.54	99.40
8	99.66	99.28
9	99.71	99.36
10	99.75	99.39

考察：

4.5 EfficientNet

表 5 EfficientNet

Epoch	Train Accuracy	Validation Accuracy
1	94.47	98.76
2	99.00	99.19
3	99.40	99.54
4	99.63	99.59
5	99.79	99.79
6	99.85	99.82
7	99.89	99.94
8	99.92	99.96
9	99.93	99.87
10	99.95	99.96

考察：

5 失敗事例分析について：Grad-Cam を用いたモデルの信頼性の検証

Grad-CAM (Gradient-weighted Class Activation Mapping) という CNN が画像のどの部分に注目して一般人か美男美女かの判断を下したのかをヒートマップとして可視化する技術がある。図 4, 図 5 に Grad-CAM を用いてモデルごとに有名人 (good) と一般人 (normal) の各データの平均したヒートマップを示す。

各画像は以下に示すモデルを適用した結果である。

- 1 行目
 - 追加学習したモデル：custom
 - EfficientNet_b0
 - ResNet18
- 2 行目
 - ResNet34
 - ResNet50

5.1 実験結果の考察

有名人の判断根拠について、custom は顔全体、特に口元がよく見られているが、それ以外の ResNet・EfficientNet については、全体の傾向として、目と口が判断根拠になっていることがわかる。

一方、一般人の判断根拠について、custom モデルは目元の上 (額と思われる) を見ており、それ以外のモデルは人物の右背景を見ていることがわかる。右背景を見ているのでは目的であった顔を見て分類する目的を果たせておらず、実験の目標は達成できていない。この原因については、図 2, 図 3 に示したように有名人の画像の全体的な傾向として、記者会見などで用いられるバックパネルやスタジオで撮影されていると思われる画像など、背景がシンプルなものが多い。一方で、一般人の画像は、Flickr 画像と Twitter やオンライン新聞の画像などがデータセットに使われており、人物の横に他の人の体の一部が写っていたりカメラを向いていなかったりするものが多い傾向にある。そういった構築したデータセットの画像の特徴から、一般人の分類において背景で判断されたと考えられる。そこで次の章では、画像を白背景にして再度を学習を行った。

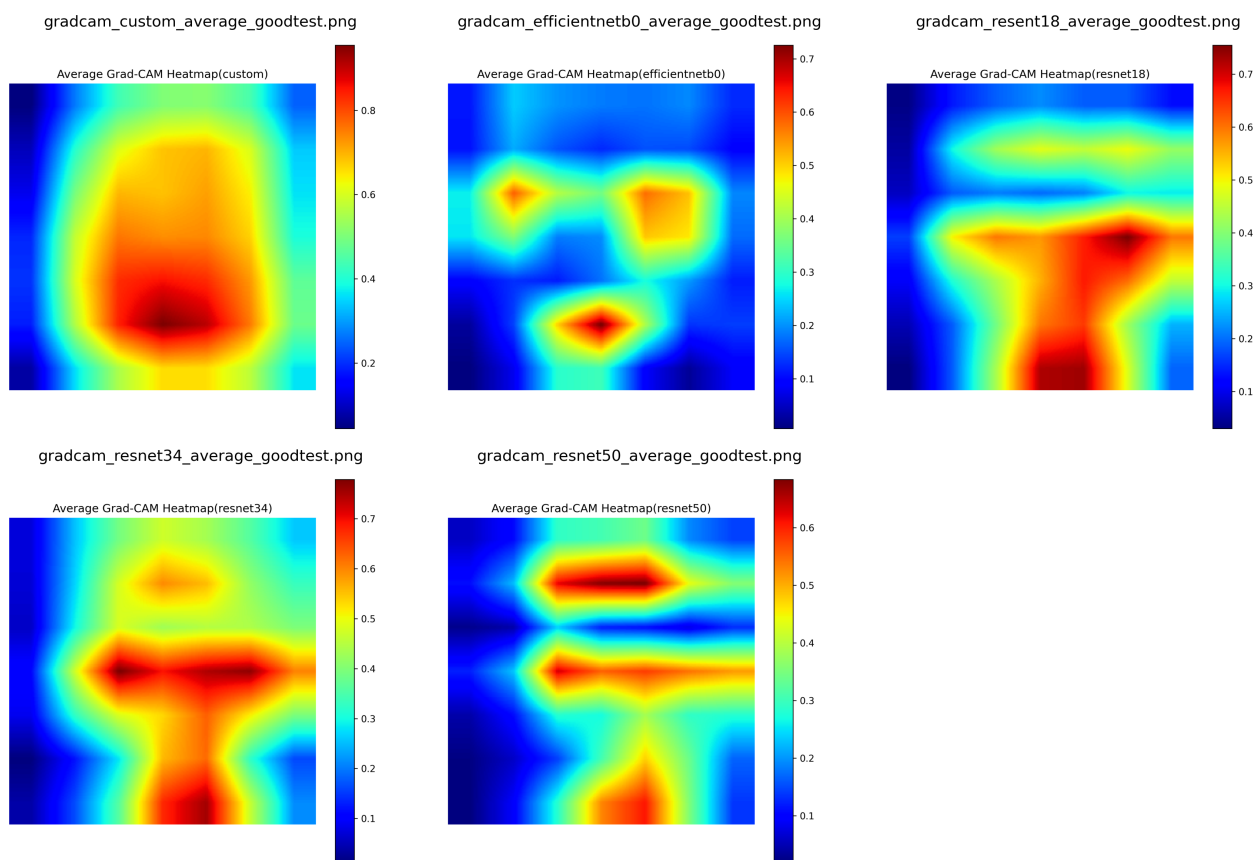


図4 Grad-CAM を用いた有名人の判断根拠の可視化

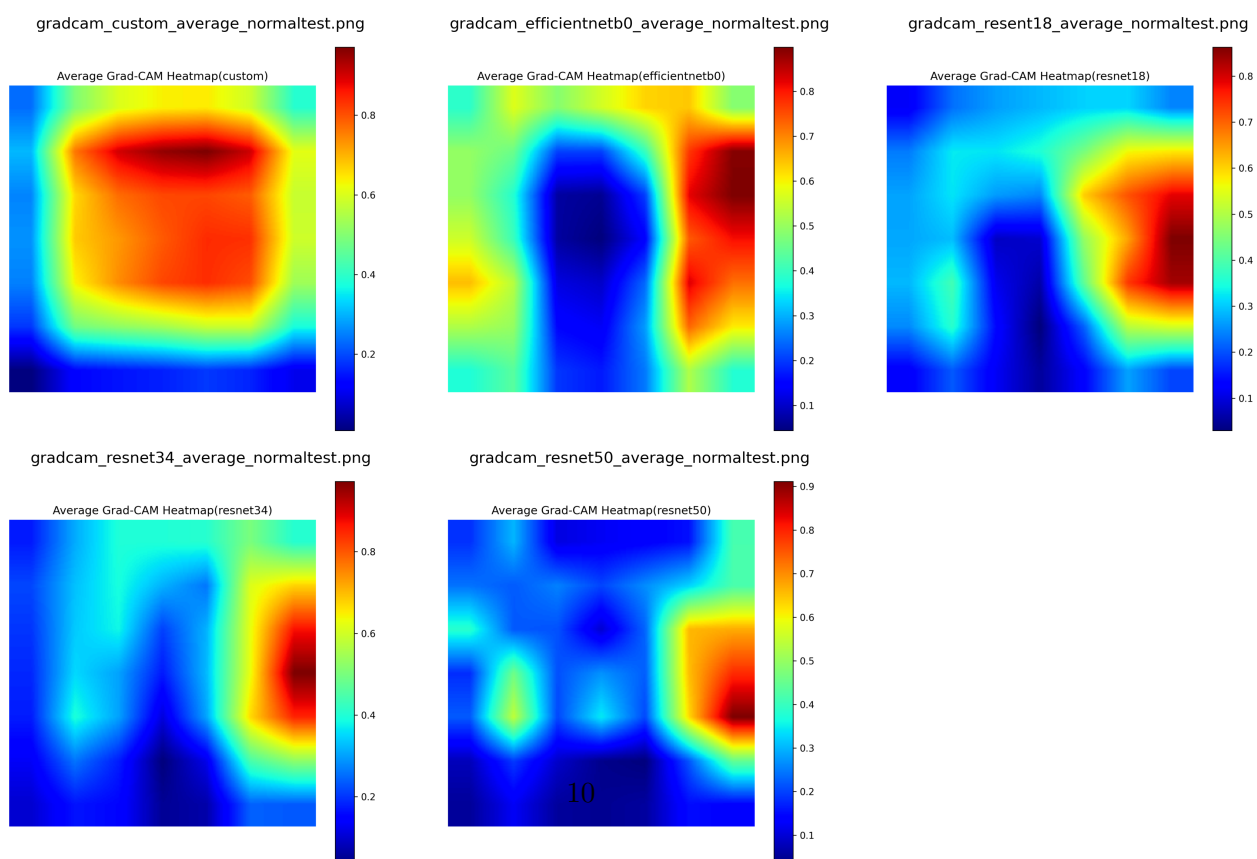


図5 Grad-CAM を用いた一般人の判断根拠の可視化

6 白背景画像を用いた失敗事例分析の改善

7 考察

8 反省・今後の課題

github を使おうとしたが、使える人数が少なく github でのバージョン管理が実質できていなかった。さらに amane で実行する場合、さらにソースコードのバージョンが不透明になってしまった。役割分担・実験計画が適切にできておらず、実験の後半で急いで学習を行った。計画よりも前処理に時間がかかり、その後も実行しその結果を考察する時間をファイルの欠損の対応・コンテナプラットフォーム Singularity 使用のための準備といった対応に終わってしまった。

8.1 時間の都合上省いた項目

8.1.1 撮影日時による人物データの抽出

画像を web スクレイピングで収集したものについて、ネットに存在する有名人の画像は撮影された日時がバラバラである。若い時の写真もあれば現在の写真もあり、年齢によって発生する特徴量の変化については考慮できていない。検索キーワードをもとにスクレイピングをする際に、人名に年齢を含むアプローチも行ったが、得られる画像が少ない・検索してもそもそも年齢で絞れないという問題が発生した。本実験で調べる限りは実験の限界と判断し、他の前処理を丁寧に行う方を優先した。

8.1.2 構築したデータセットの人名の偏り改善

本実験では、2024 年美男美女ランキング top50 をもとに、50 名ずつ抽出し、アップサンプリング（左右反転）を施した。これは一般人のデータセットと比較して、偏りが生じる。他のアップサンプリングの方法として回転・輝度変更といった手法があるが、これらでさらにデータを増やすよりは、美男美女の人名数を 50 名ずつよりさらに増やす方がバイアスを削減できると考える。本実験では時間の都合・第一回の実験時の計画がずれていたため省略した。

9 まとめ

fairface で構築されているデータセットは、本当にイケメンでない人が含まれているのかという疑問から、分担して画像を確認していったが、いわゆる芸能人並のイケメンの数は少なかった。

参考文献

- [1] Karkkainen, Kimmo and Joo, Jungseock. FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1548–1558, 2021.
<https://arxiv.org/abs/1908.04913>
- [2] Most Handsome Man In The World 2024, shiningawards.com,
<https://shiningawards.com/most-handsome-man-in-the-world-2024/>
- [3] Most Beautiful Faces 2024, gigazine.net,
https://gigazine.net/gsc_news/en/20241229-most-beautiful-faces-2024/
- [4] Nataniel Ruiz Eunji Chong and James M. Rehg (Georgia Institute of Technology). Fine-Grained Head Pose Estimation Without Keypoints In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 2074–2083, 2018.
<https://arxiv.org/abs/1710.00925>
- [5] Head Pose Estimation, <https://github.com/natanielruiz/deep-head-pose>