ML-Based Age Prediction
and Personal Appearance Insights

Team 6
2021150013 Gaeun Ko
2020140300 Seunghyun Choi

Problem Definition

NEWS RELEASE 28-FEB-2023

People spend 1/6th of their lifetime on enhancing their appearance

And not only to find the love of their life!

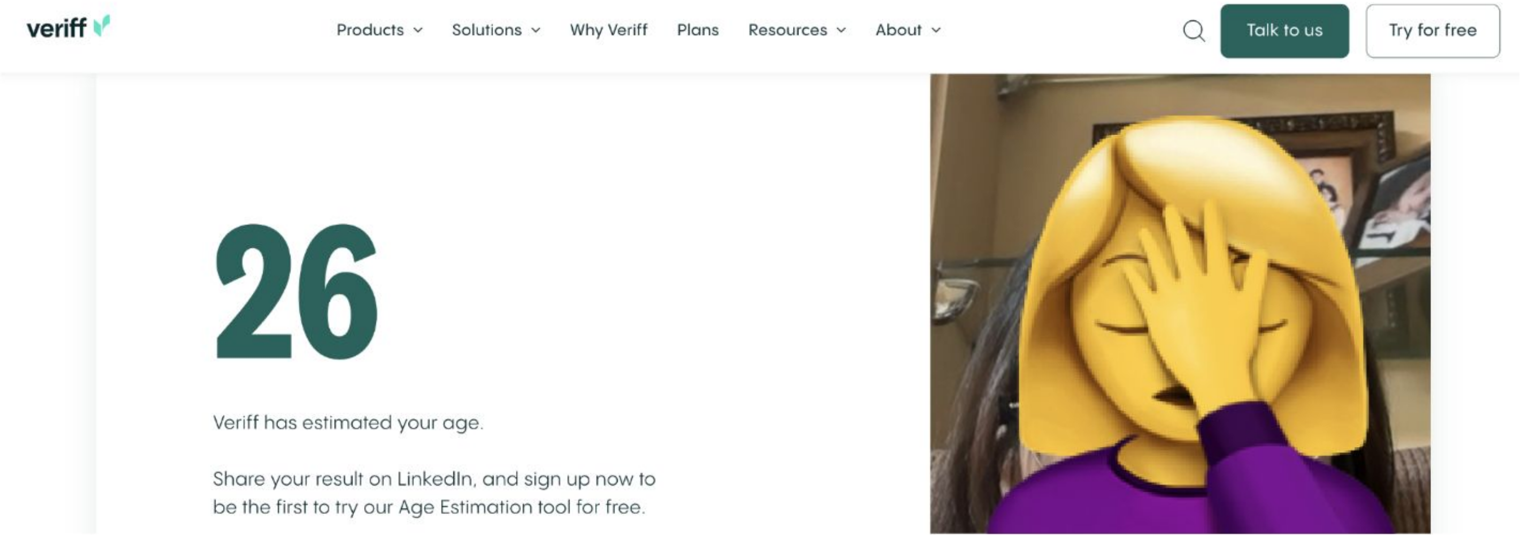
Peer-Reviewed Publication

NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS

People spend 1/6th of their lifetime trying to enhance their appearance.
Yet, they often lack objective feedback about how they look — or why.
There's a growing need for explainable, personalized insights to guide appearance-related decisions.

Problem Definition

Today, age estimation models can predict a person's age from facial images.
But most tools don't explain **what facial features** influence perceived age.
Users don't just want to know how old they look they want to know **why and what they can do about it**.



* This is the test result of one of our team members using the existing model

Dataset : All-Age-Faces(AAF)

Initial dataset : UTKFace

https://susanqq.github.io/UTKFace/

- All human race mixed
- Not cropped images (different face location, size)
- Poor model quality (Random Forest)
- Divided data set by race and gender

→ Asian showed best quality in both genders

Male - White | Samples: 2480 | MAE: 14.591, R^2: 0.4144
Male - Black | Samples: 205 | MAE: 20.43, R^2: -0.0294
Male - Asian | Samples: 796 | MAE: 6.389, R^2: 0.6696
Male - Indian | Samples: 548 | MAE: 10.136, R^2: 0.5079
Male - Others | Samples: 462 | MAE: 7.597, R^2: 0.6682
Female - White | Samples: 2836 | MAE: 14.806, R^2: 0.451
Female - Black | Samples: 203 | MAE: 12.588, R^2: 0.3988
Female - Asian | Samples: 890 | MAE: 7.898, R^2: 0.5742
Female - Indian | Samples: 943 | MAE: 8.11, R^2: 0.4937
Female - Others | Samples: 657 | MAE: 6.269, R^2: 0.3159

'Aging might show different patterns in different human races'

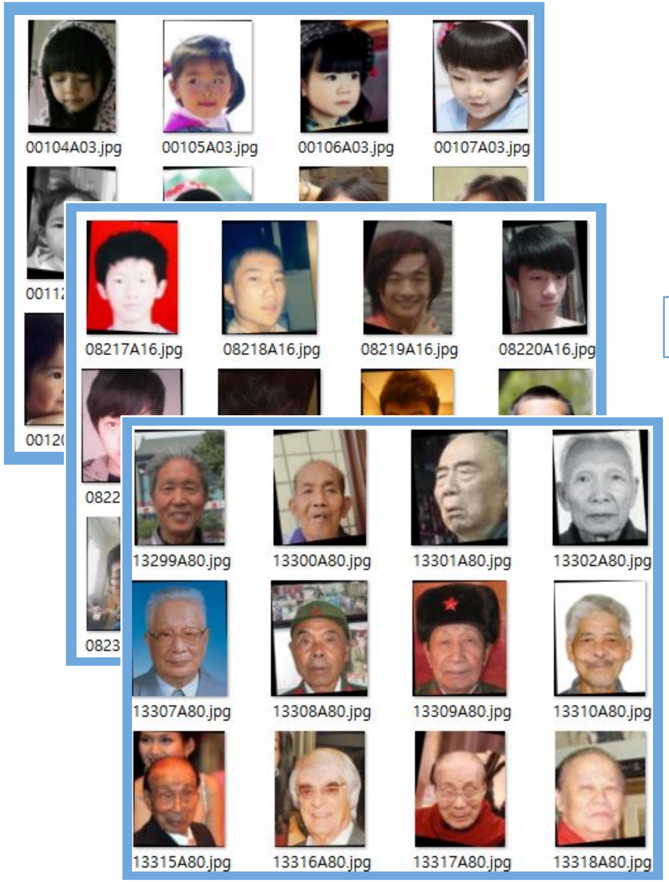
Dataset : All-Age-Faces(AAF)

New dataset : AAF

https://github.com/JingchunCheng/All-Age-Faces-Dataset

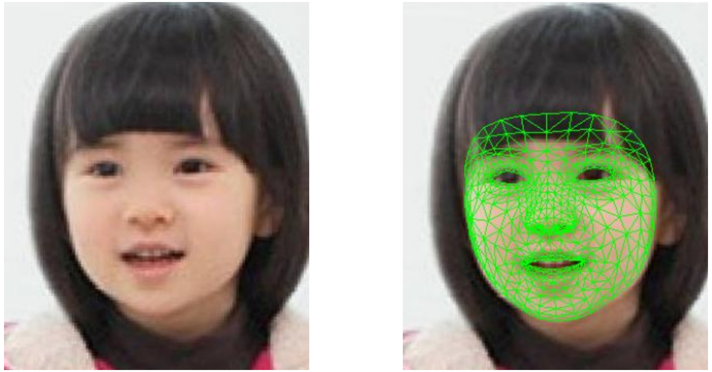
- Only containing Asian faces (13322 faces)
- Aligned images (File consists only faces)

XGBoost Model - Training



1. **Feature Extraction**

- Mediapipe-based structural and geometric features
- OpenCV-based texture and brightness-related features



Checking feature capturing quality

XGBoost Model - Training

2. Model Training (XGBoost)

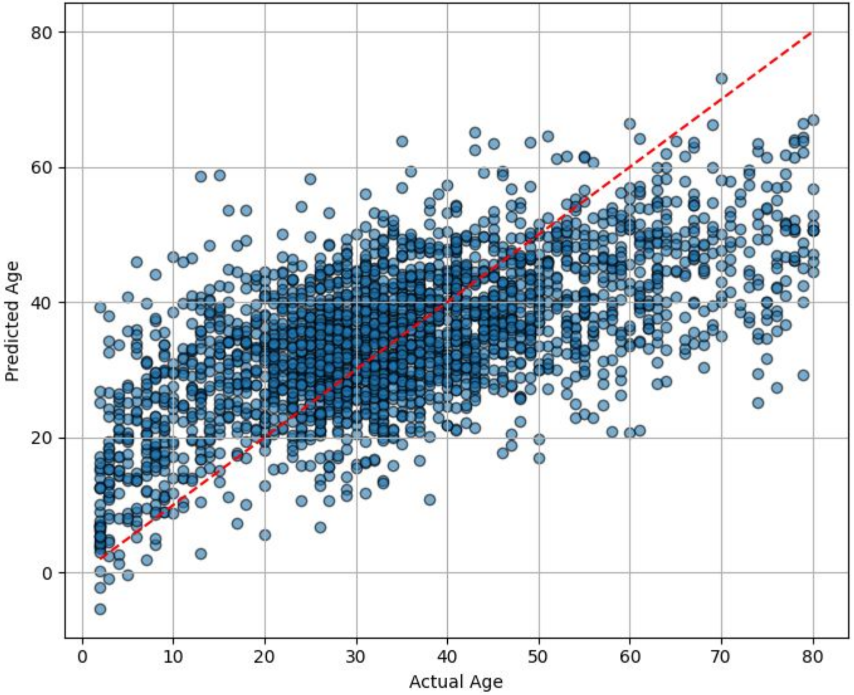
- Random Forest
- **XGBoost Regressor**
- XGBoost Classifier
- Ensemble model

Found best model from various combinations

- 20 Features
- 41 Features
- 100+ Features

41 Features → Age predicting model training

XGBoost Model - Results



Model Quality (All age)

- MAE : 10.367
- RMSE : 13.297
- R^2 : 0.397

✅Generally follows upward trend

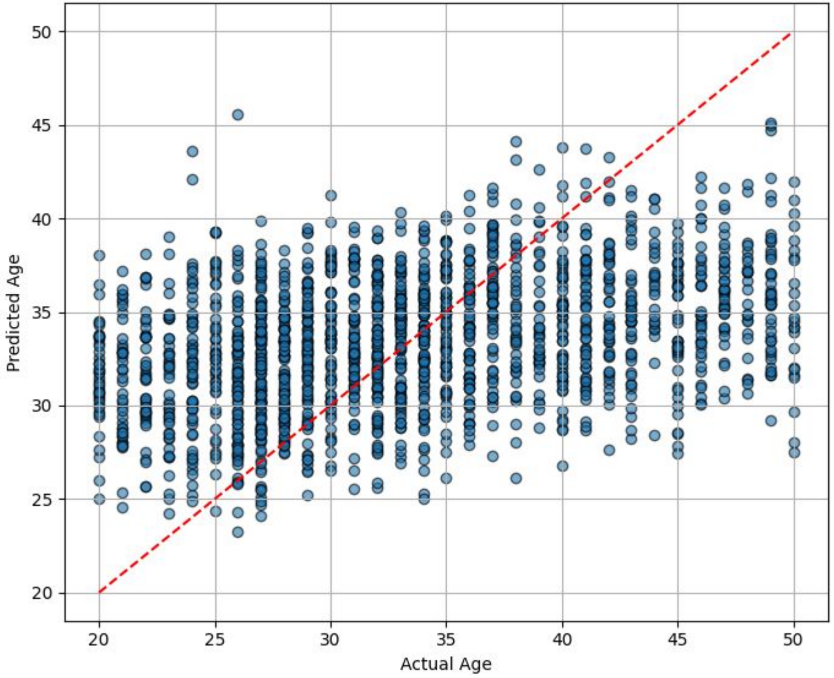
But..

🔺 Overprediction on young

🔻 Underprediction on old

→ **Predicted age clustering**

XGBoost Model - Results



Model Quality (Age 20 to 50)

- MAE : 5.946
- RMSE : 7.317
- R^2 : 0.136

✅ Lower MAE, better accuracy

❗ Still low R^2 → poor variance explanation

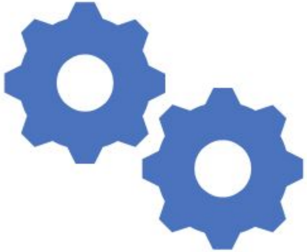
🎯 Predictions converge around mid-30s

XGBoost Model - Prediction

Age Prediction & Explanation

Input actual age

```
Your Actual Age : 23
```

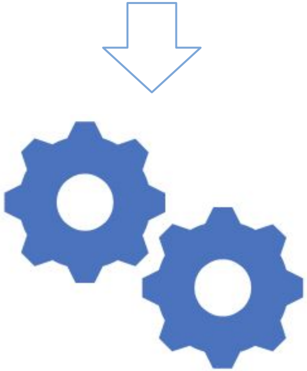


Age prediction

```
Actual Age: 23
Predicted Age: 41.26
 -->  older-looking (Difference: 18.26)
```

XGBoost Model - Prediction

Age Prediction & Explanation



SHAP based explanation

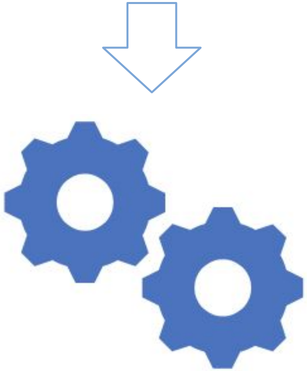
*SHAP : SHapley Additive exPlanations

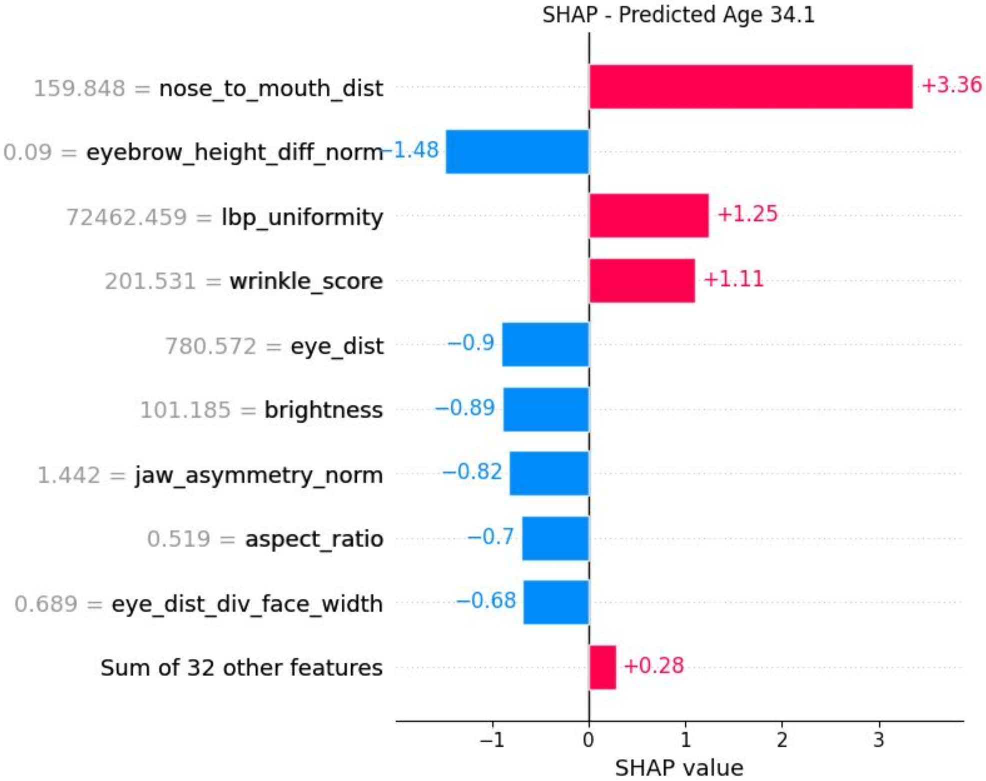
```
💡Main feature (Top 5):
                                       feature         value      shap
nose_to_mouth_dist            nose_to_mouth_dist    159.848251  3.358203
eyebrow_height_diff_norm  eyebrow_height_diff_norm    0.090414 -1.482256
lbp_uniformity                    lbp_uniformity  72462.458722  1.247929
wrinkle_score                      wrinkle_score    201.530972  1.106811
eye_dist                                eye_dist    780.571563 -0.902351
```

```
🔍 Top contributing facial features and their effects:
- More distance between nose and lips makes the person look older. (SHAP: 3.358)
- Less eyebrow height imbalance makes the person look younger. (SHAP: -1.482)
- More smoothness of skin pattern makes the person look older. (SHAP: 1.248)
- More visible wrinkles makes the person look older. (SHAP: 1.107)
- Less space between the eyes makes the person look younger. (SHAP: -0.902)
```

XGBoost Model - Prediction

Age Prediction & Explanation





XGBoost Model - Limitation

❗ **Limitation of XGBoost**

- Narrow prediction range & low R^2 → age clustering
- Overestimates youth / Underestimates elderly
- **Depends on handcrafted features (nose, mouth, brightness...)**

  ➔ Aging is too complicated to be captured by pre-defined features, regardless of its number (tried 100+ features)
- Misses subtle skin-level aging cues

➡ **Motivation to switch**: Deep learning model (CNN) that learns from raw image data.

Architecture

CNN (Resnet 18) → Age Prediction Model → Grad CAM / SHAP

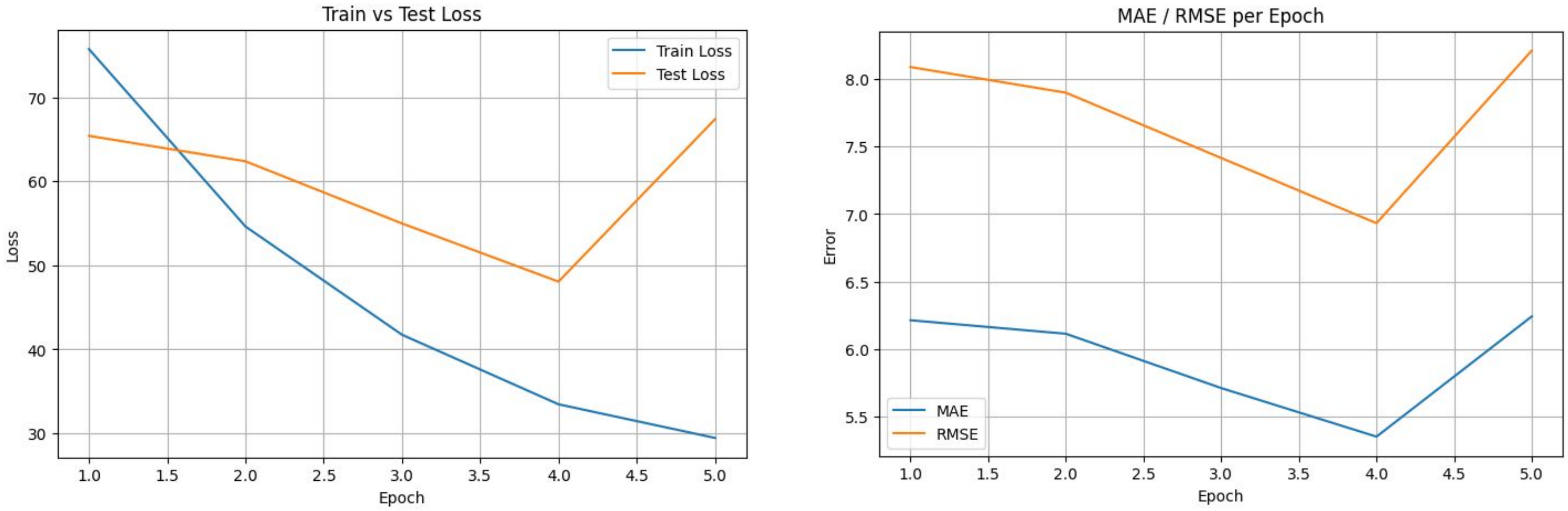
batchsize = 8
epoch = 4

MAE = 5.36

eye, mouth, forehead

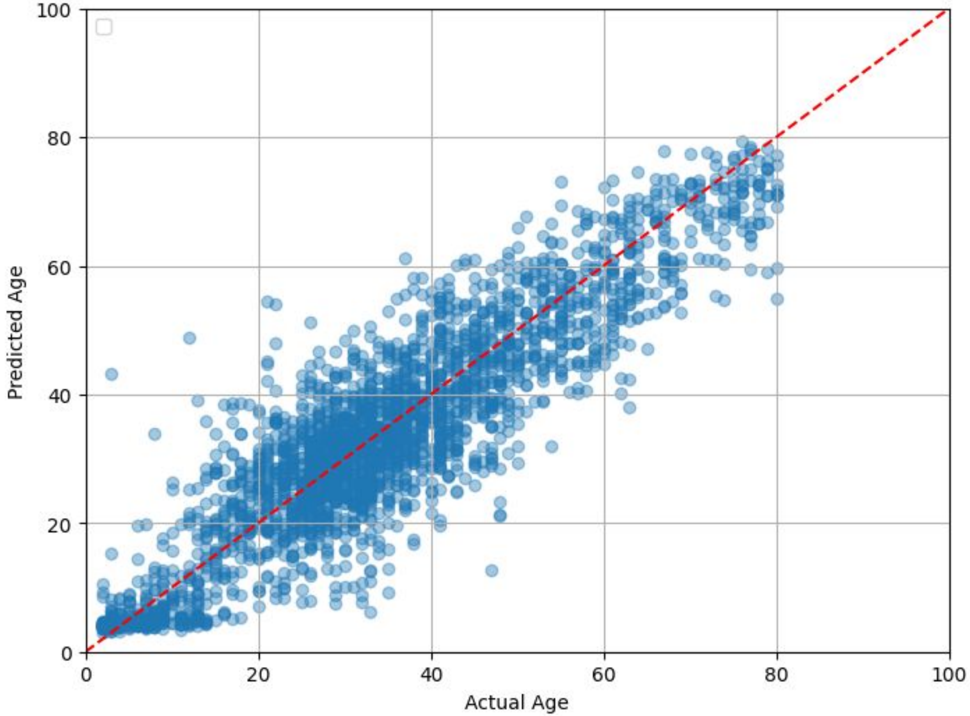
output:
image based age prediction
Classify as looking young or old
feature explanation

CNN (Resnet 18)



ResNet-18 outperformed ResNet-50, likely due to better generalization on a small, imbalanced dataset.
Age group imbalance was addressed using **inverse-frequency sampling weights.**
The model achieved optimal performance at **epoch 4**, with overfitting observed beyond that point.

Experimental Results



[0-10] MAE = 2.87 (n=223)
[11-20] MAE = 6.55 (n=263)
[21-30] MAE = 5.74 (n=666)
[31-40] MAE = 5.99 (n=652)
[41-50] MAE = 7.09 (n=401)
[51-60] MAE = 6.26 (n=209)
[61-70] MAE = 6.53 (n=151)
[71-100] MAE = 6.24 (n=100)

Model Quality

- MAE : 5.36
- RMSE : 6.93
- R^2 : 0.83

🔺 Outperformed XGBoost overall

🔻 Lower accuracy in older age groups
→ data scarcity and higher variance in perceived age

Experimental Results

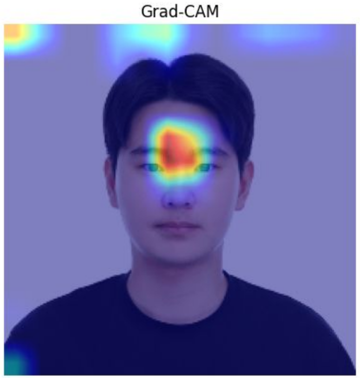
Age prediction

```
Predicted Age: 25.75 years
Actual Age: 23 years
The person looks appropriate for their age.
```

Grad CAM & SHAP based explanation

```
SHAP-based Feature Contributions:
- The eye area contributes to looking older.
- The forehead area contributes to looking younger.
- The mouth area contributes to looking younger.

Grad-CAM shows highest attention on the eye region (score: 0.099).
SHAP and Grad-CAM both highlight: eye, forehead
```



Conclusion

Conclusion

- Initially applied **XGBoost**, but it performed poorly due to limitations in capturing visual features.
- Switched to **CNN**, which showed better age prediction performance.
- Both models had lower accuracy for older age groups due to **data imbalance.**
- **CNN's black-box nature** made feature interpretation difficult.
- SHAP and Grad-CAM revealed important facial regions (e.g., eyes, forehead, mouth), but failed to explain **what specific cues** within those regions influenced predictions.

Future Improvements

- Use a **larger and more balanced dataset**, especially for older age groups, to improve model generalization.
- Explore **hybrid models** combining CNNs with interpretable methods for both accuracy and explainability.

Thank You

github (code and weight)

https://github.com/luckygeko/COSE471_XGBoost-team6.git

https://github.com/luckygeko/COSE471_CNN-team6.git