

Age Recognition

Mohammed Hassan Ali

20201149

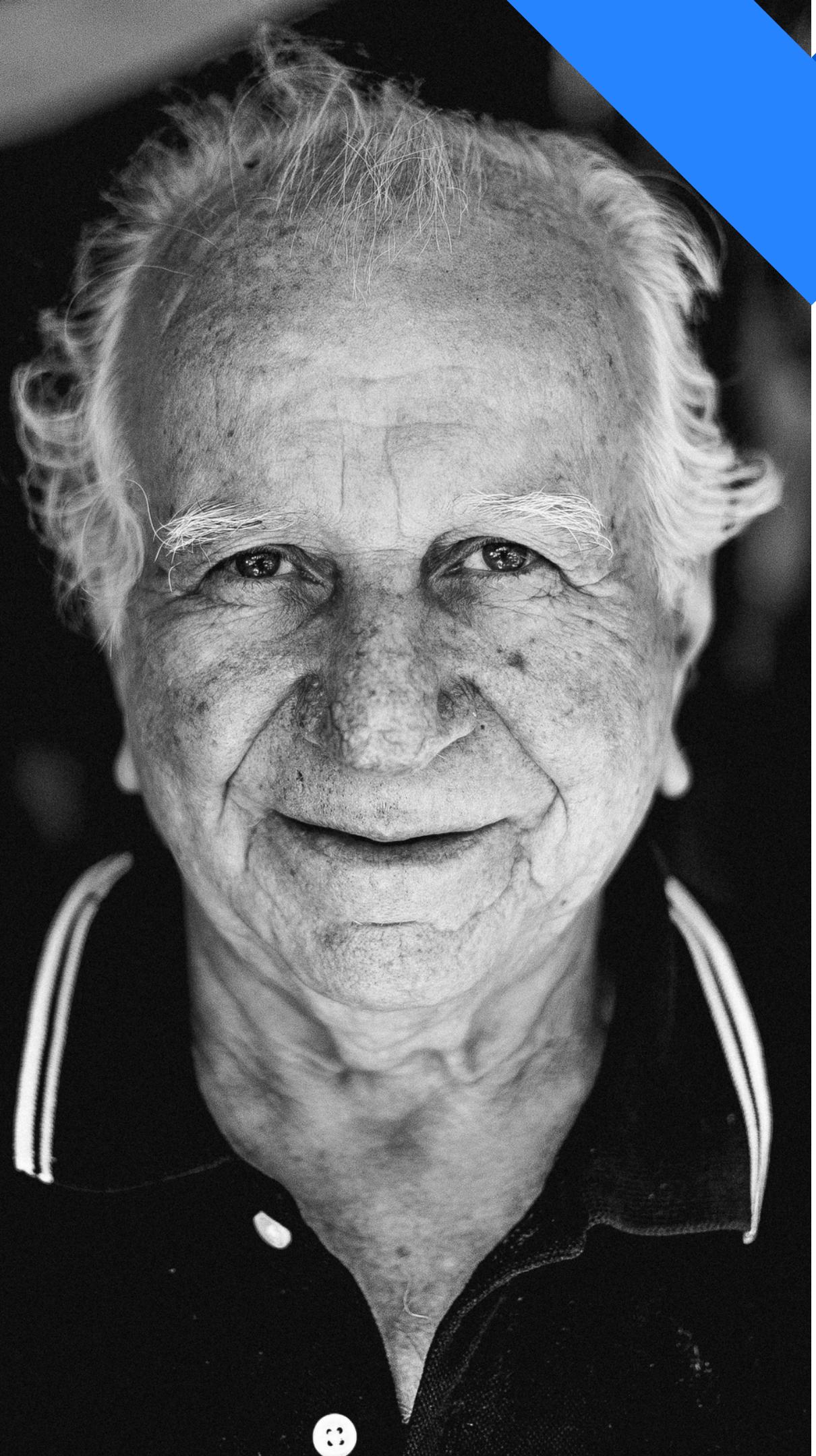
Table of Content

- Problem Definition
- About Dataset
- 4 Ranges & The Conversion
- Main Challenge
- Target Solution
- Pre-processing Techniques
- Feature Extraction/Selection
- Classifiers
- Ensemble Classifiries

Problem Definition

The problem definition of an age recognition dataset project involves accurately predicting the age of individuals based on their facial features. The dataset typically includes images of people along with their corresponding age labels. The goal is to develop a machine learning model that can accurately predict the age of new individuals based on their facial features and classify them into four categories, which could have various applications in fields such as security, marketing, and healthcare.





About Dataset

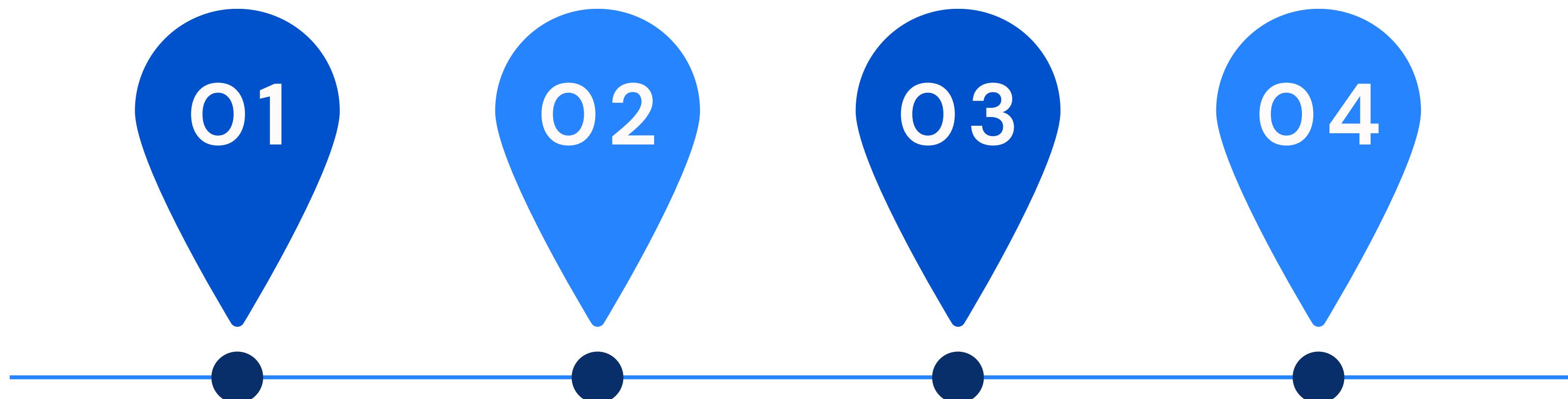
The Age Recognition Dataset is a collection of data that consists of **4** categories, each of which describes an age interval.

The categories in the dataset are defined as follows:

1. Category 1: Ages **6-20**
2. Category 2: Ages **25-30**
3. Category 3: Ages **42-48**
4. Category 4: Ages **60-98**

The dataset includes images of human faces and their corresponding age interval category. It is balanced and can be used for training and testing age recognition algorithms. The dataset is useful for developing new age recognition techniques.

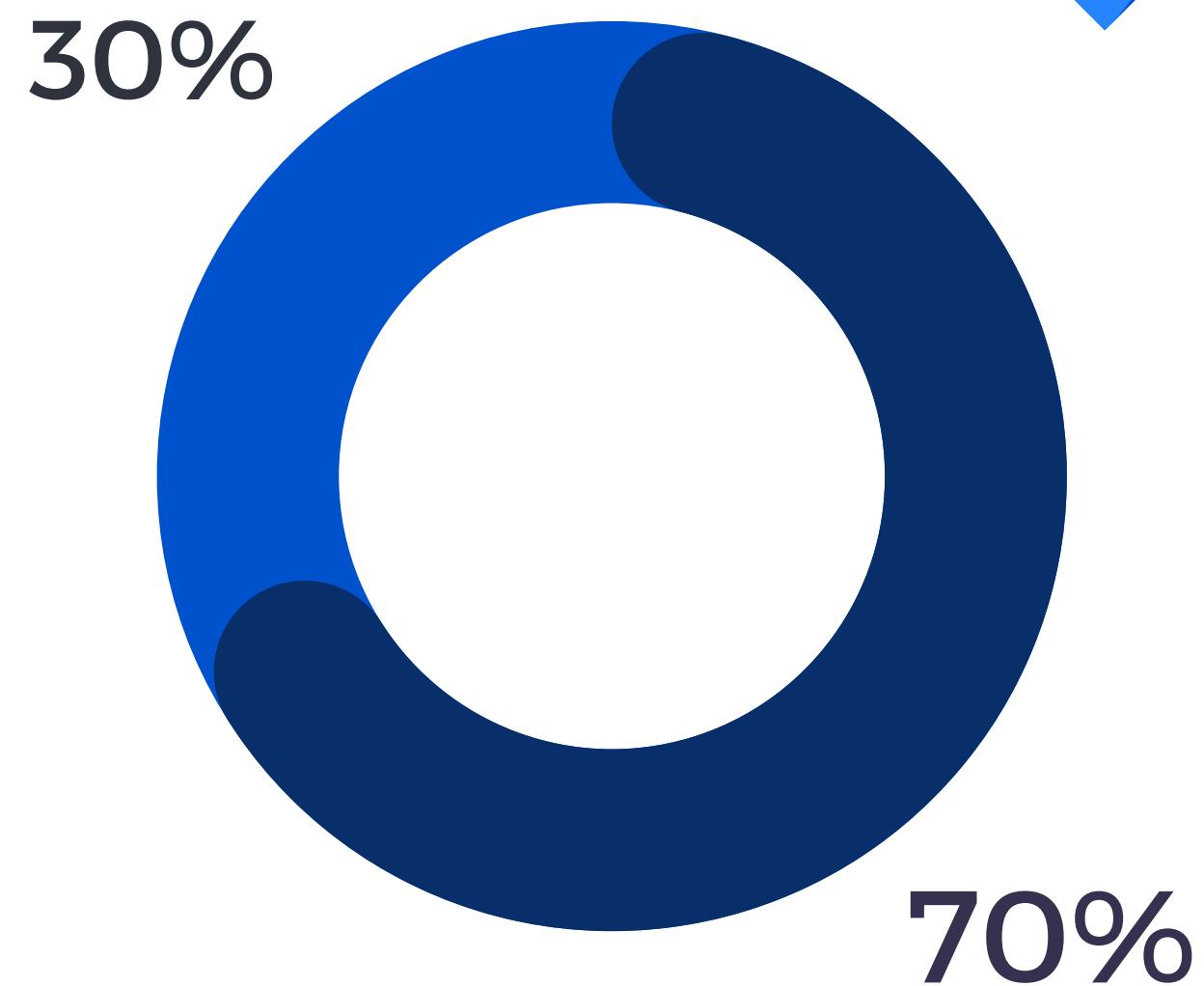
The 4 Ranges:



Total **9097** files

Conversion of Data

Category	Train (70%)	Test (30%)
6-20	1760	754
25-30	1747	748
42-48	1007	432
60-98	1854	795



The data was multi-selectional data (categories), then I converted it into a binary dataset by dividing it manually into 70% for the training dataset and 30% for the testing dataset, each dataset contains 4 files of the 4 categories.

The Main Challenge

The main challenge that the Age Recognition Dataset project tackles is developing accurate age recognition algorithms. Accurately determining a person's age based on their facial features is a difficult task due to variations in facial features and changes in appearance over time.

To address this challenge, the project provides a balanced dataset of images of human faces and their corresponding age intervals. This dataset can be used to train and test age recognition algorithms and to develop new techniques for accurately determining a person's age based on their facial features.

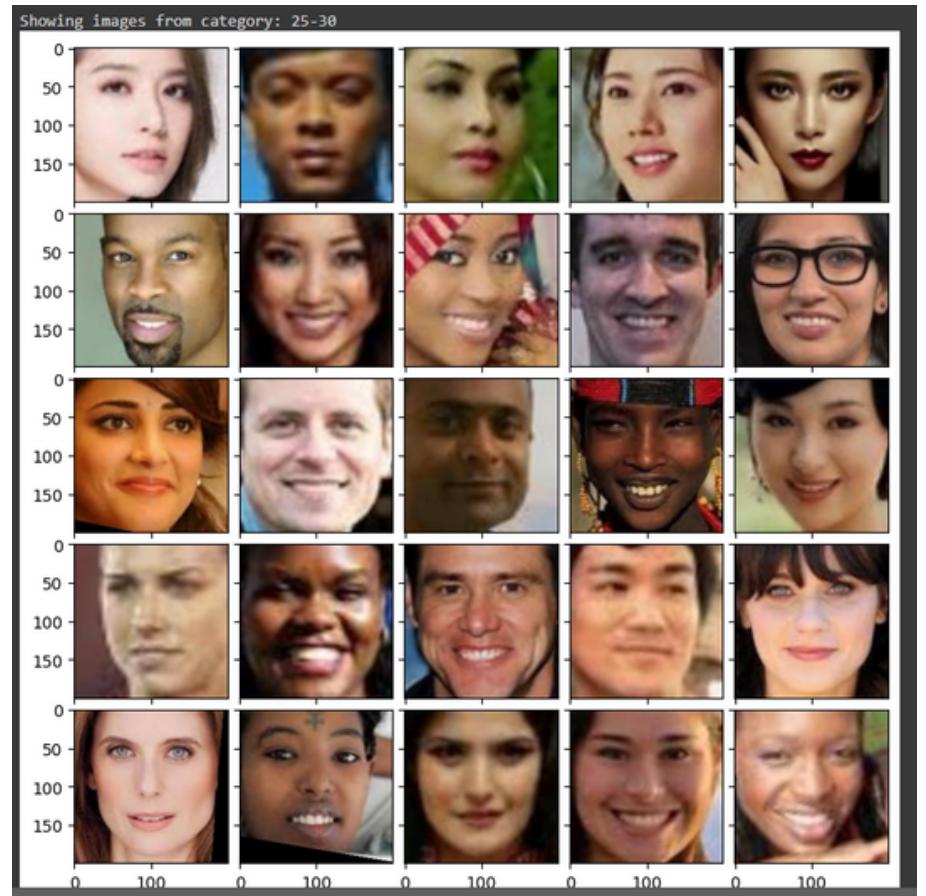
The Target Solution

The target solution is to improve the accuracy of age recognition algorithms, which can have applications in various fields such as face recognition, age-based marketing, and age-based content filtering. By providing a dataset that can be used to develop and test these algorithms, the Age Recognition Dataset project aims to advance the state-of-the-art in age recognition technology.

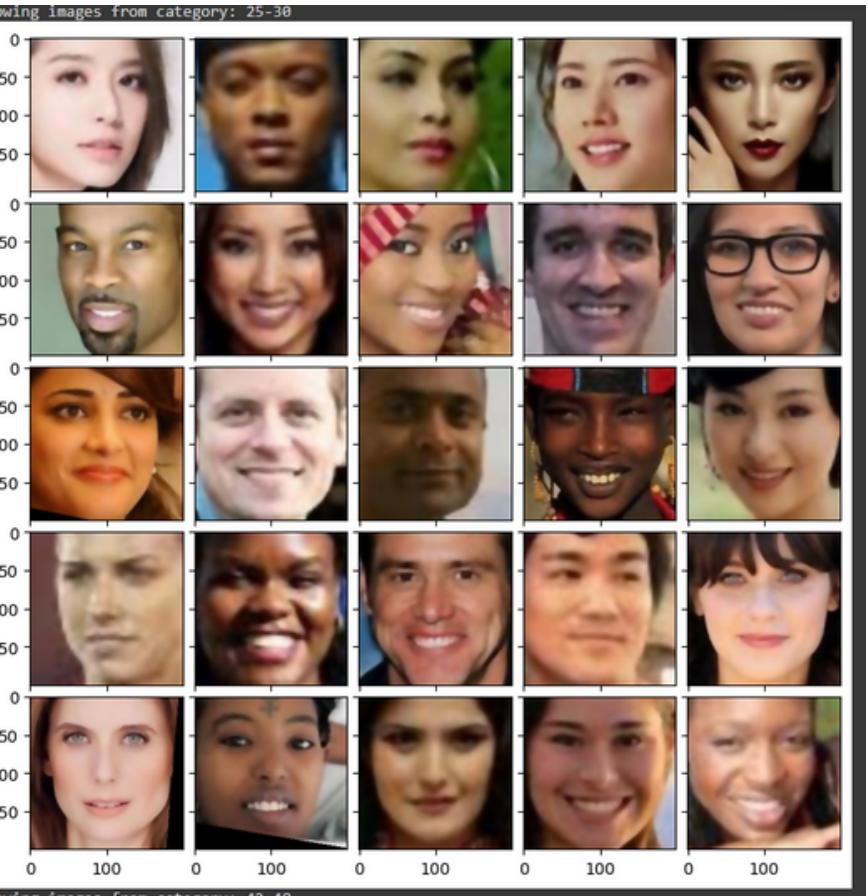


Pre-processing Techniques

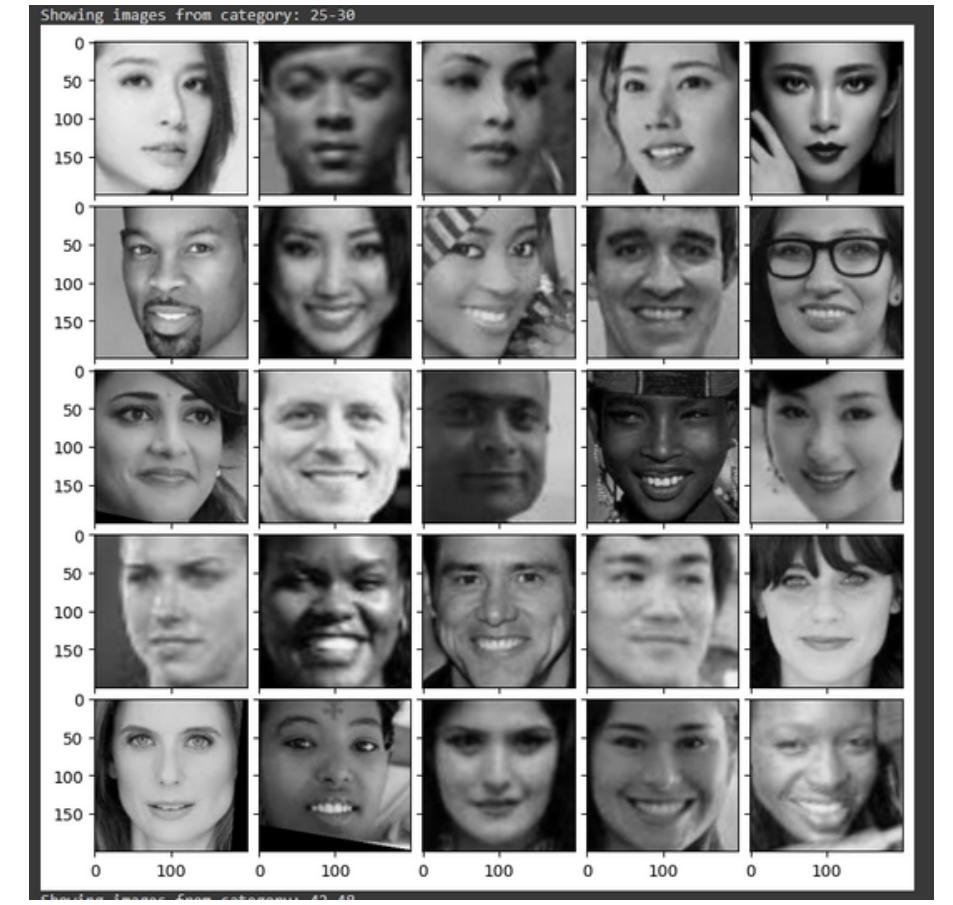
- Used **resize**, to make all data the same size.
- **Bilateral filtering** is a type of edge-preserving smoothing filter that reduces noise while preserving edges in an image. It works by taking a weighted average of the pixels in a neighborhood, where the weights depend on both the spatial distance between the pixels and their intensity similarity.
- **Brightness enhancement** to optimize the brightness
- Applied **gray scale** to reduce the color information in the images and convert them into black and white.
- Applied **normalization** to standardize the data and make it easier to process.
- Used **Median Filter** to remove noise from the images and enhance the quality of the data



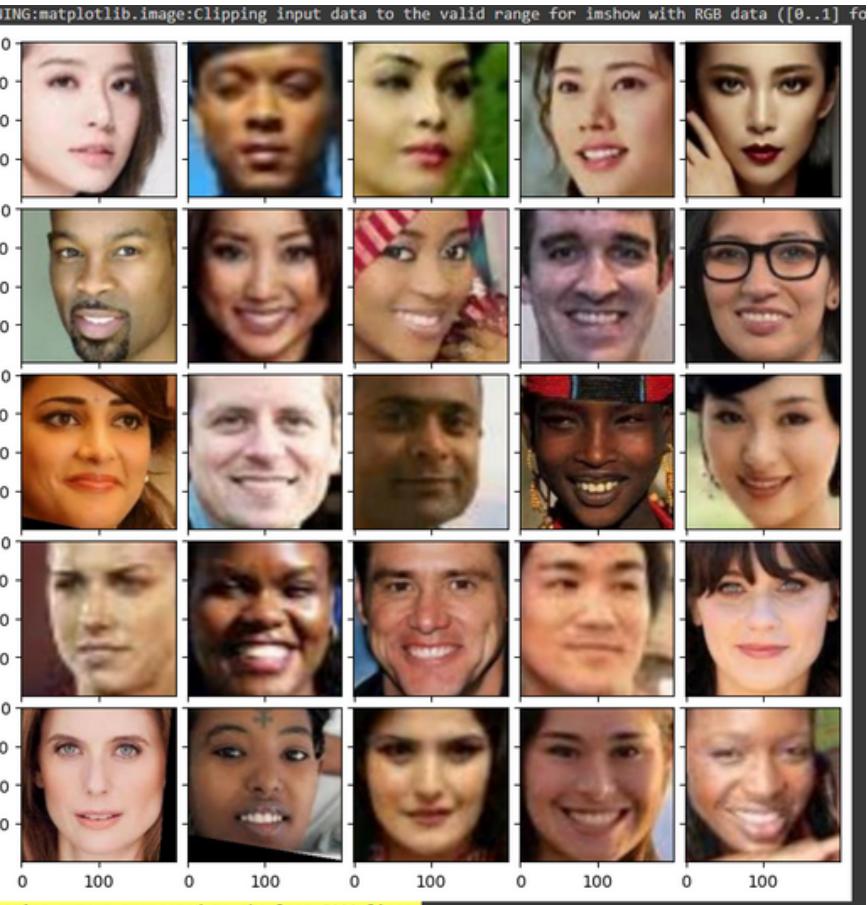
Resize



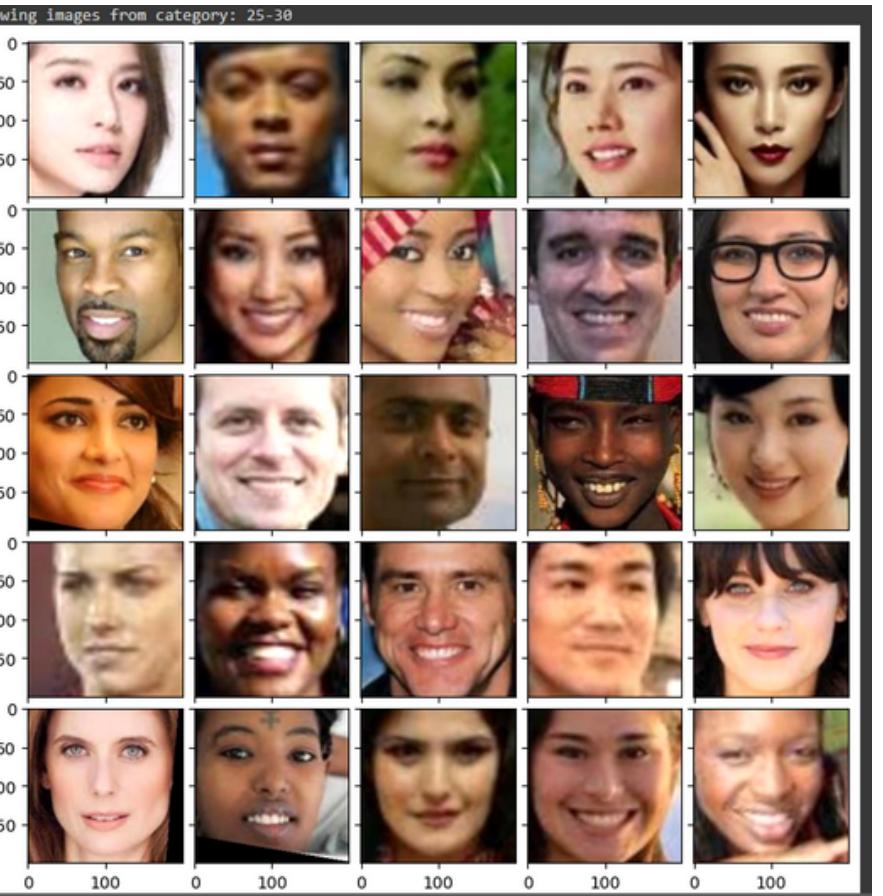
Bilateral filtering



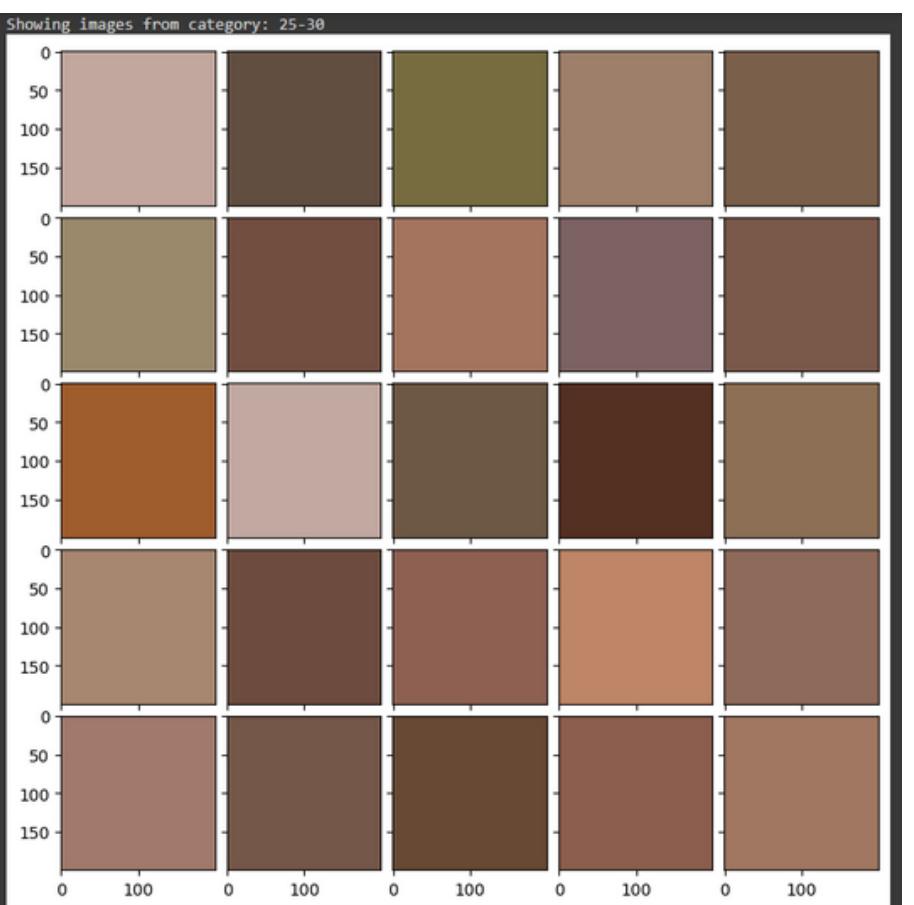
Gray Scale



Normalization



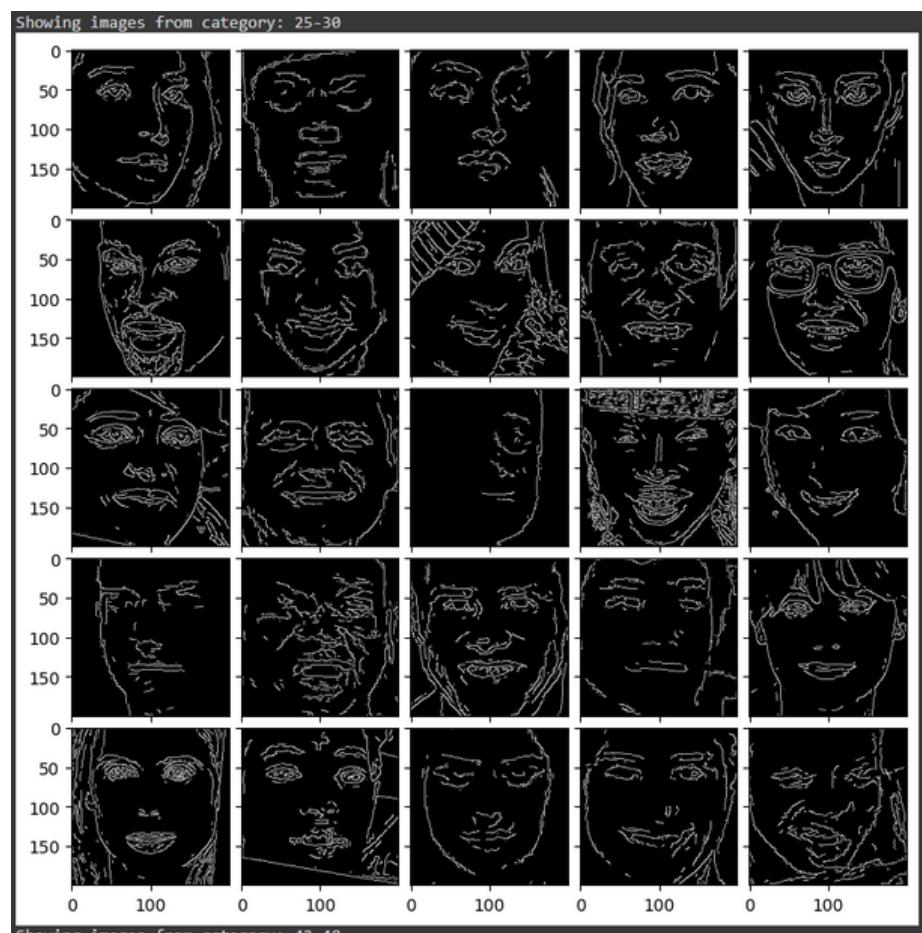
Brightness enhancement



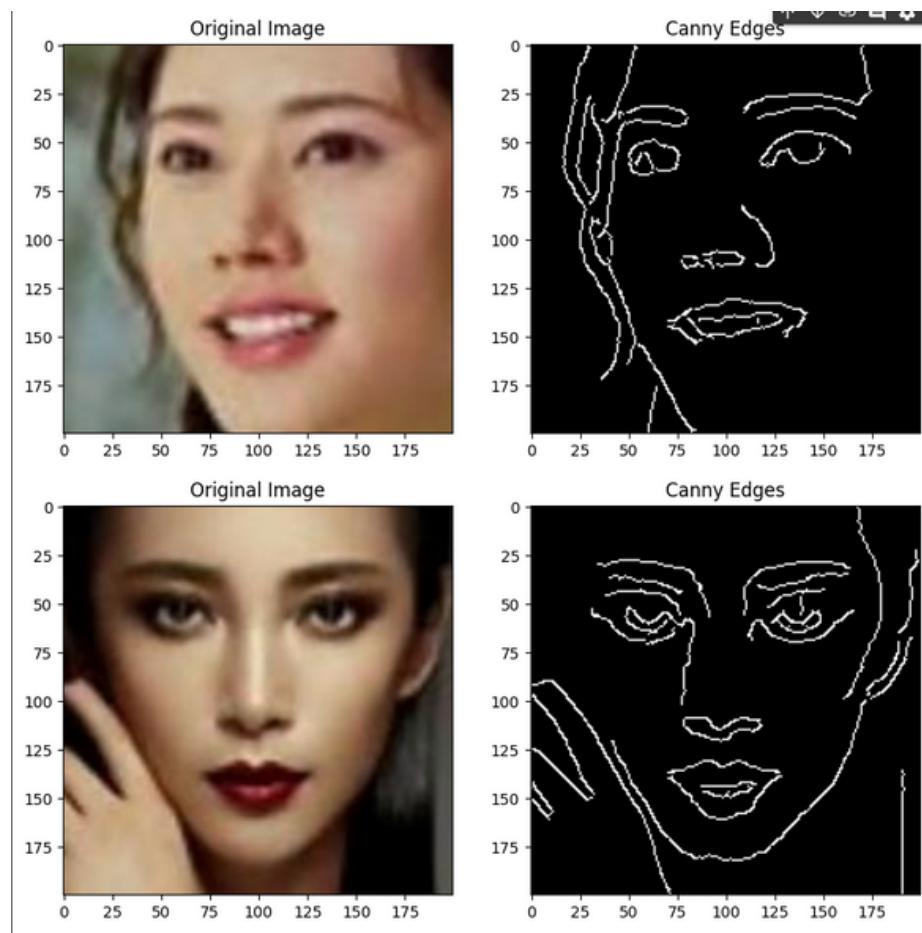
Median Filter

Feature Extraction/Selection

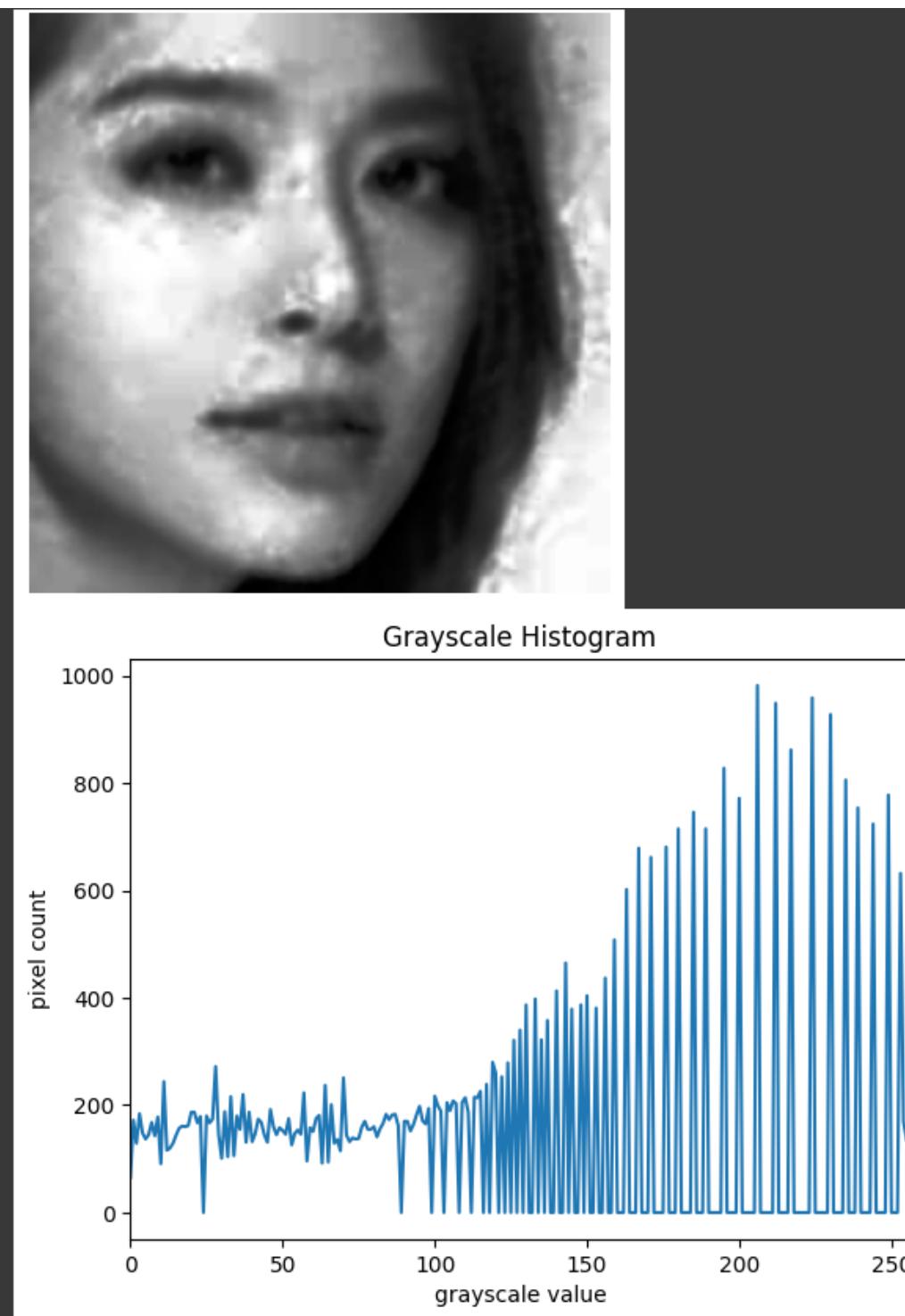
- Face Detection
- Gray Scale histogram
- Canny edge detection
- HOG -> Histogram of Oriented Gradients
- LBP -> Local Binary Patterns
- PCA -> Principle Component Function



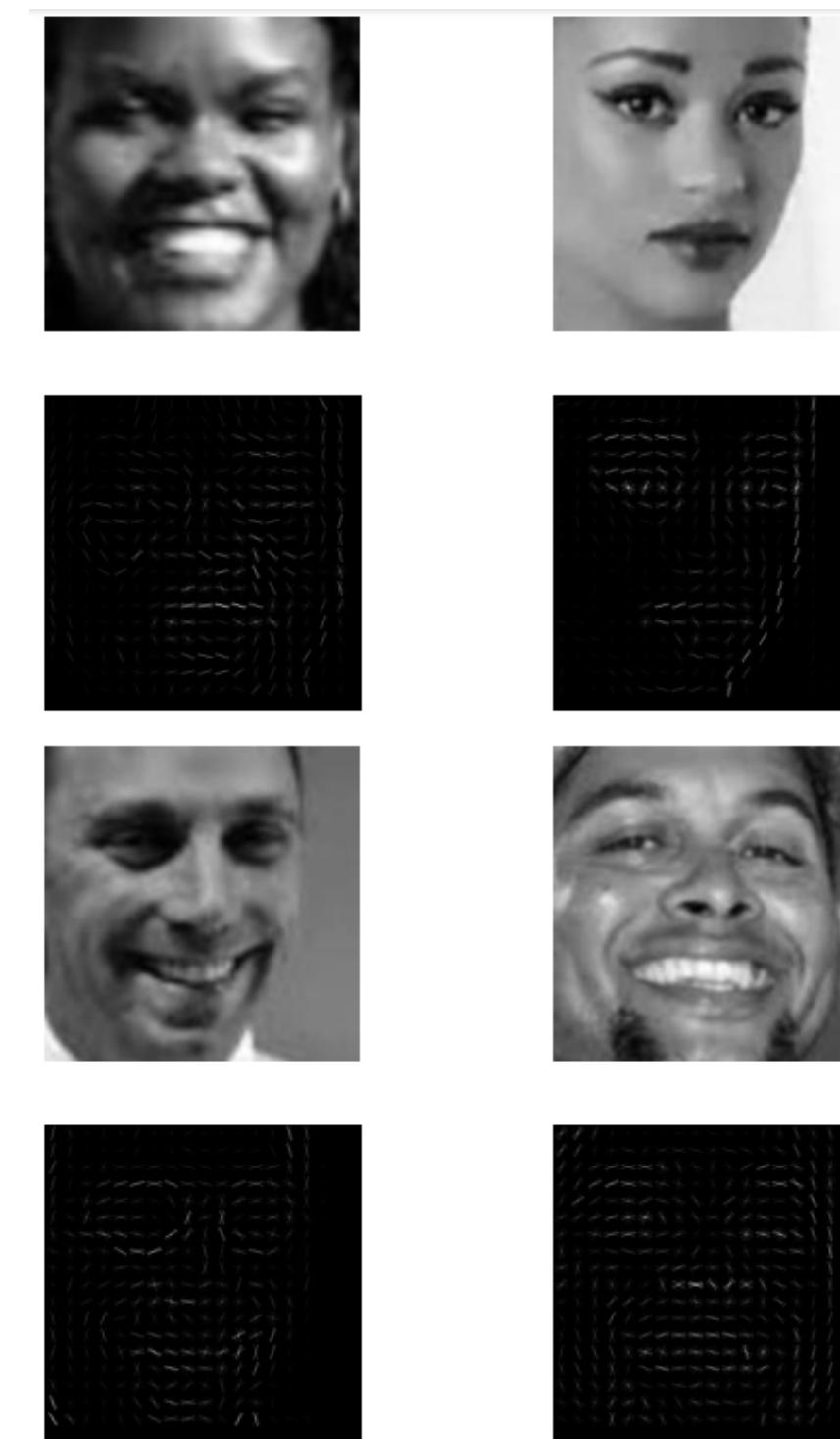
Canny edge detection 1



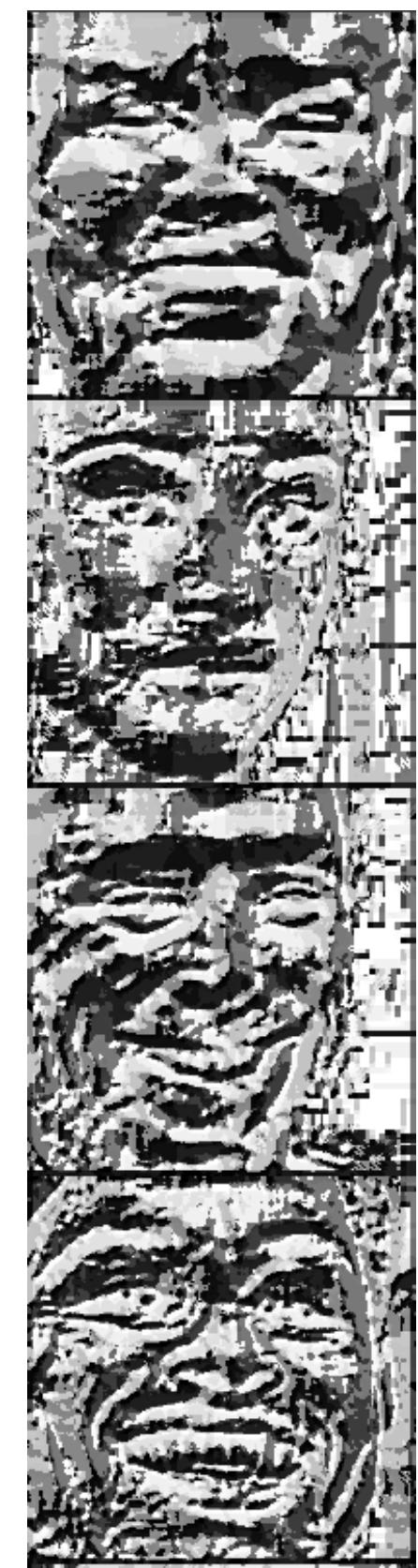
Canny edge detection 2



Gray Scale
Histogram

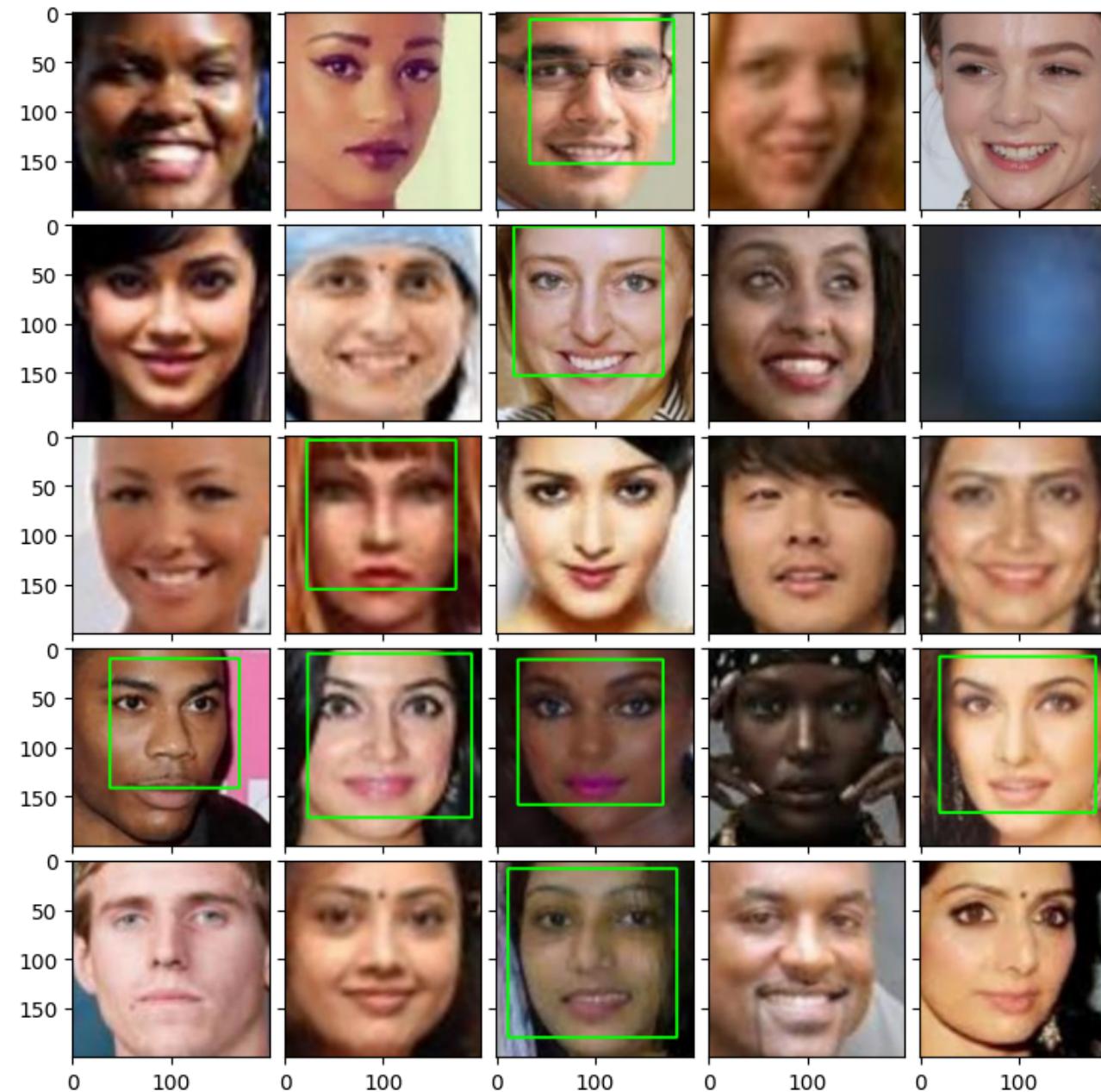


Histogram of Oriented
Gradients (HOG)



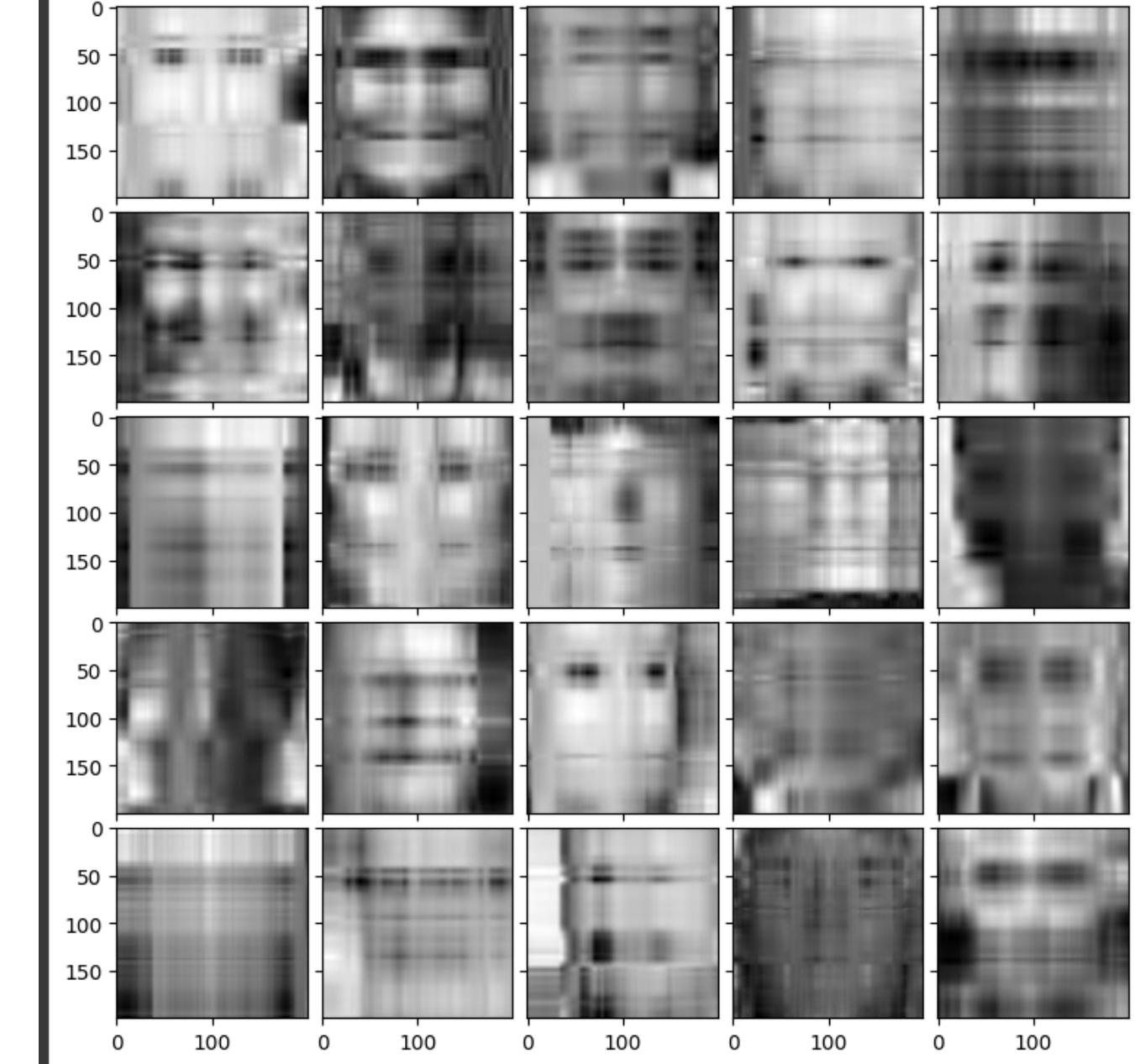
Local Binary
Patterns (LBP)

Showing images from category: 25-30



Face detection

Showing images from category: 60-98



PCA

Classifiers

- **Classifier 1:** SVM -> Support Vector Machine
- **Classifier 2:** RFT -> Random Forest Tree

Support Vector Machine with poly kernel

Accuracy: 0.6585557299843015
True Label: 25-30
Predicted Label: 60-98



True Label: 60-98
Predicted Label: 60-98



True Label: 60-98
Predicted Label: 60-98



True Label: 42-48
Predicted Label: 42-48



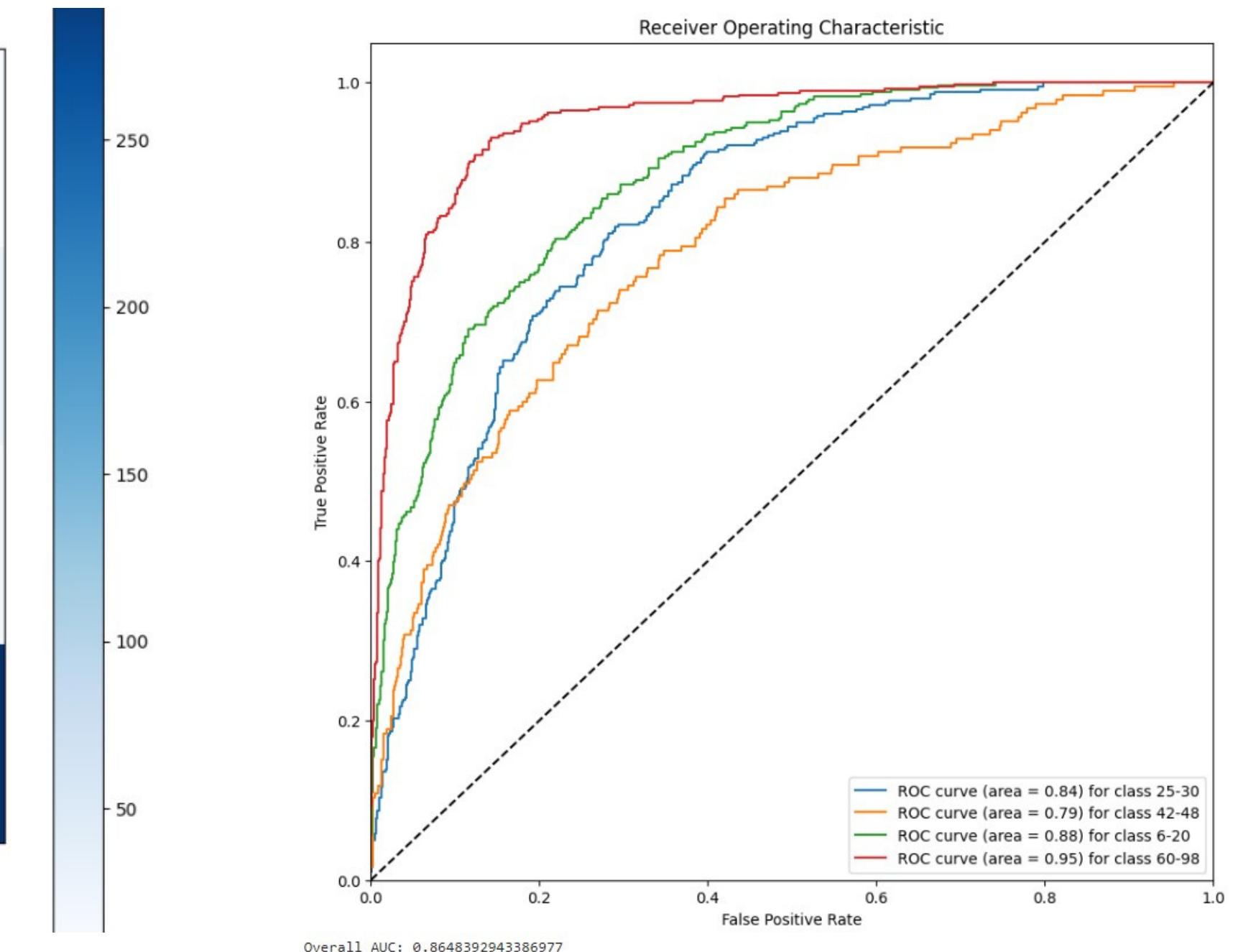
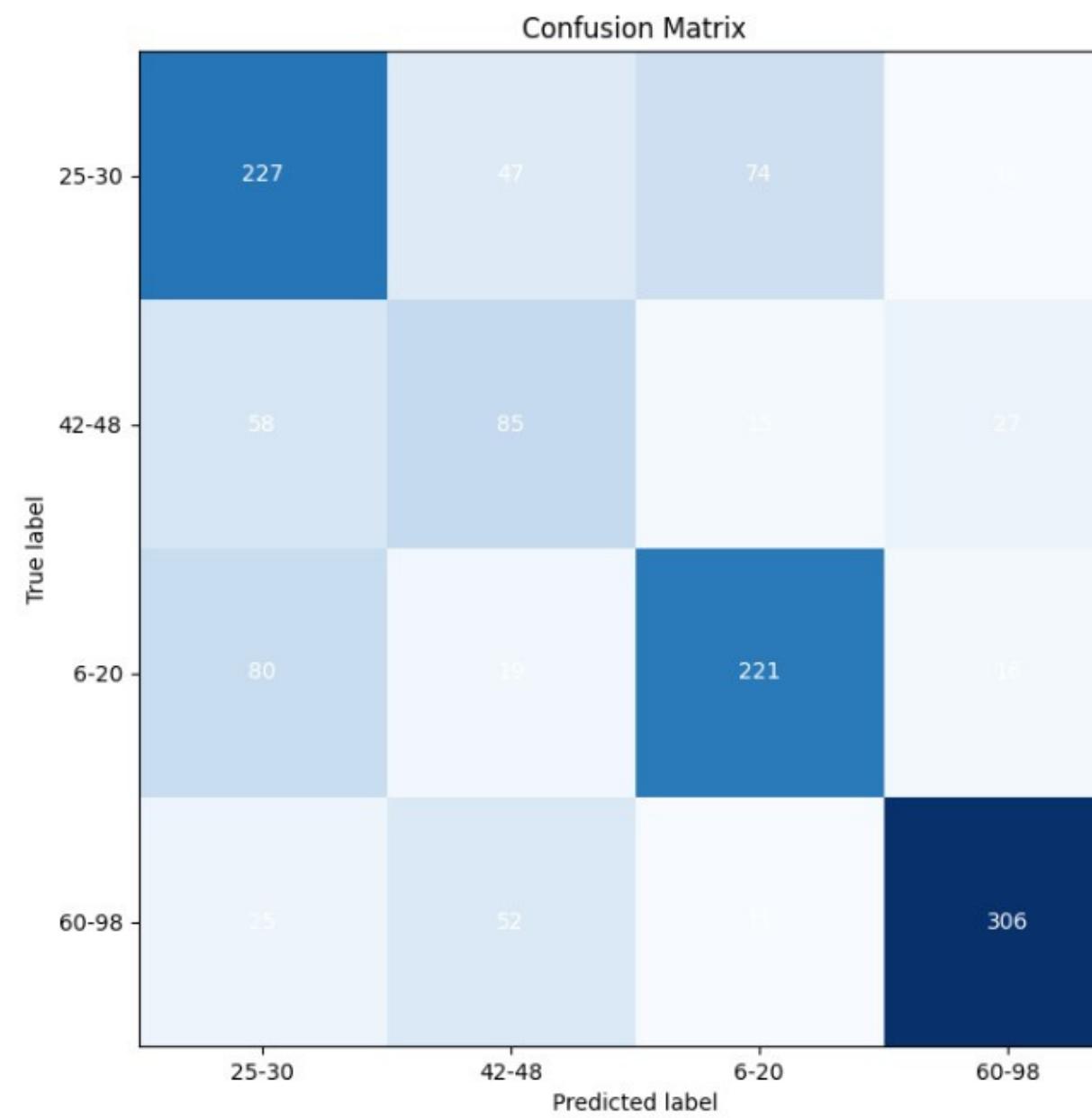
True Label: 6-20
Predicted Label: 6-20



True Label: 6-20
Predicted Label: 6-20



True Label: 25-30
Predicted Label: 25-30



Support Vector Machine with RBF kernel

Accuracy: 0.6836734693877551

True Label: 60-98

Predicted Label: 60-98



True Label: 6-20

Predicted Label: 6-20



True Label: 25-30

Predicted Label: 25-30



True Label: 60-98

Predicted Label: 60-98



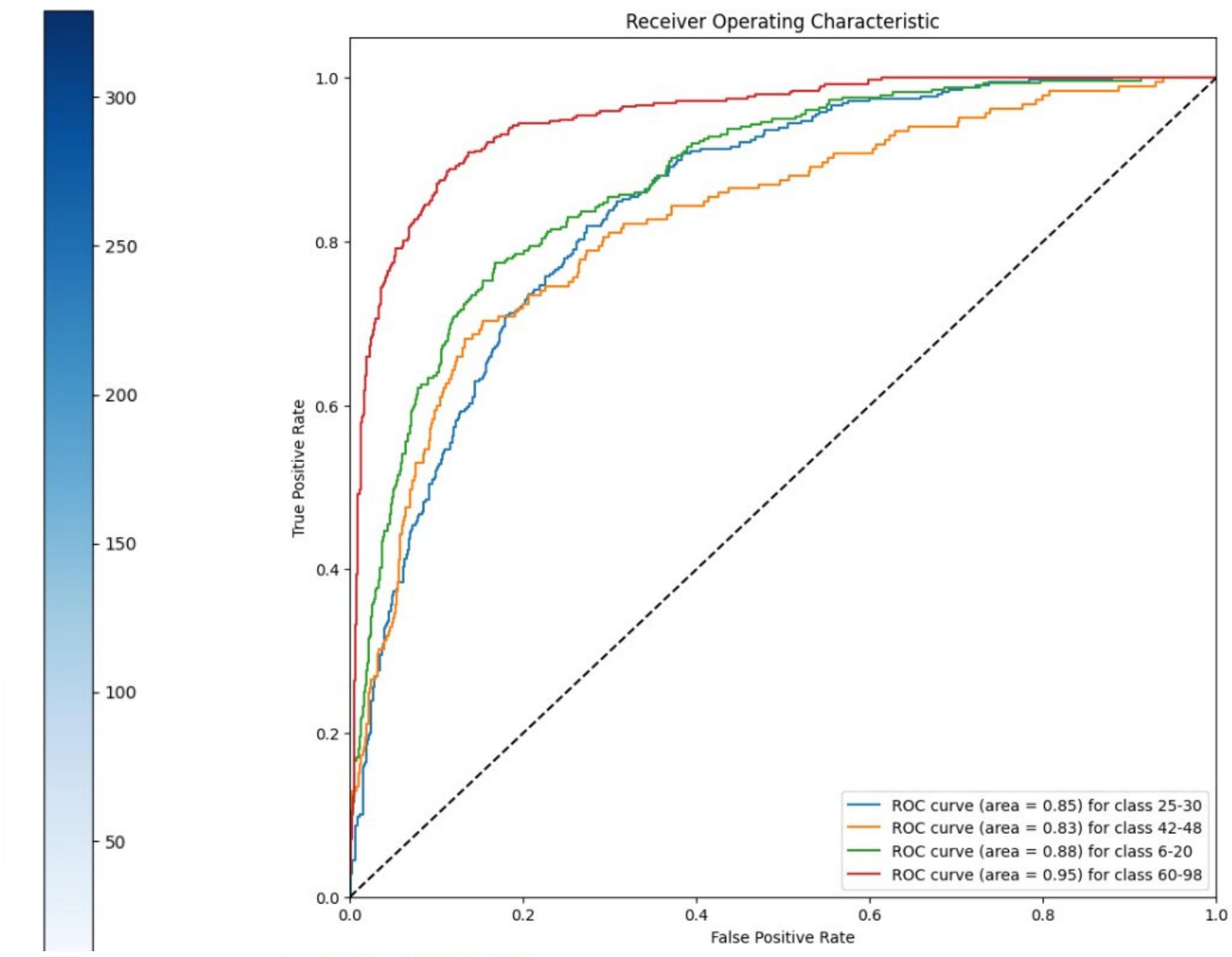
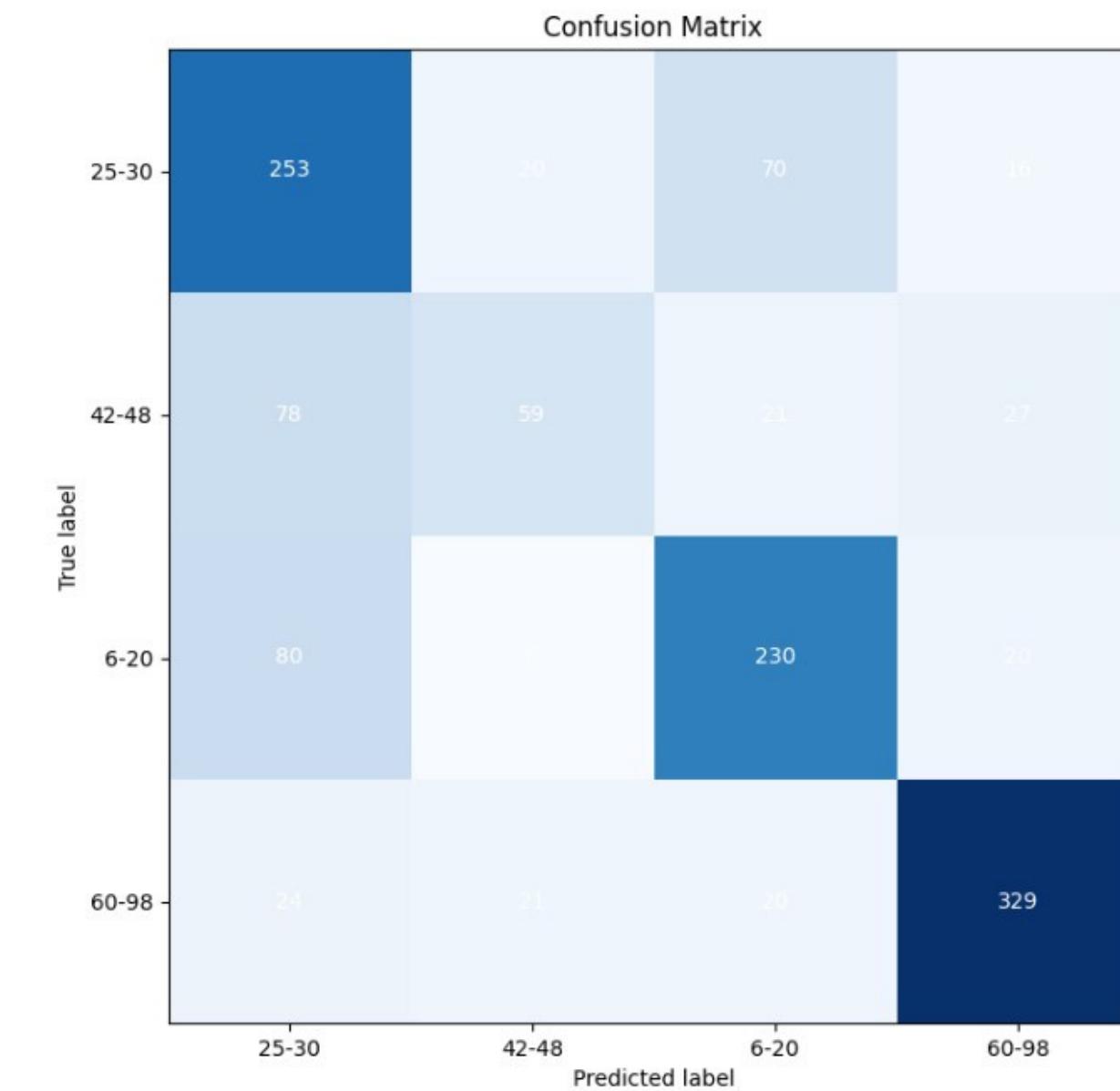
True Label: 60-98

Predicted Label: 60-98



True Label: 6-20

Predicted Label: 6-20



Random Forest Tree

Accuracy: 0.6475667189952904

True Label: 25-30

Predicted Label: 25-30



True Label: 6-20

Predicted Label: 25-30



True Label: 60-98

Predicted Label: 60-98



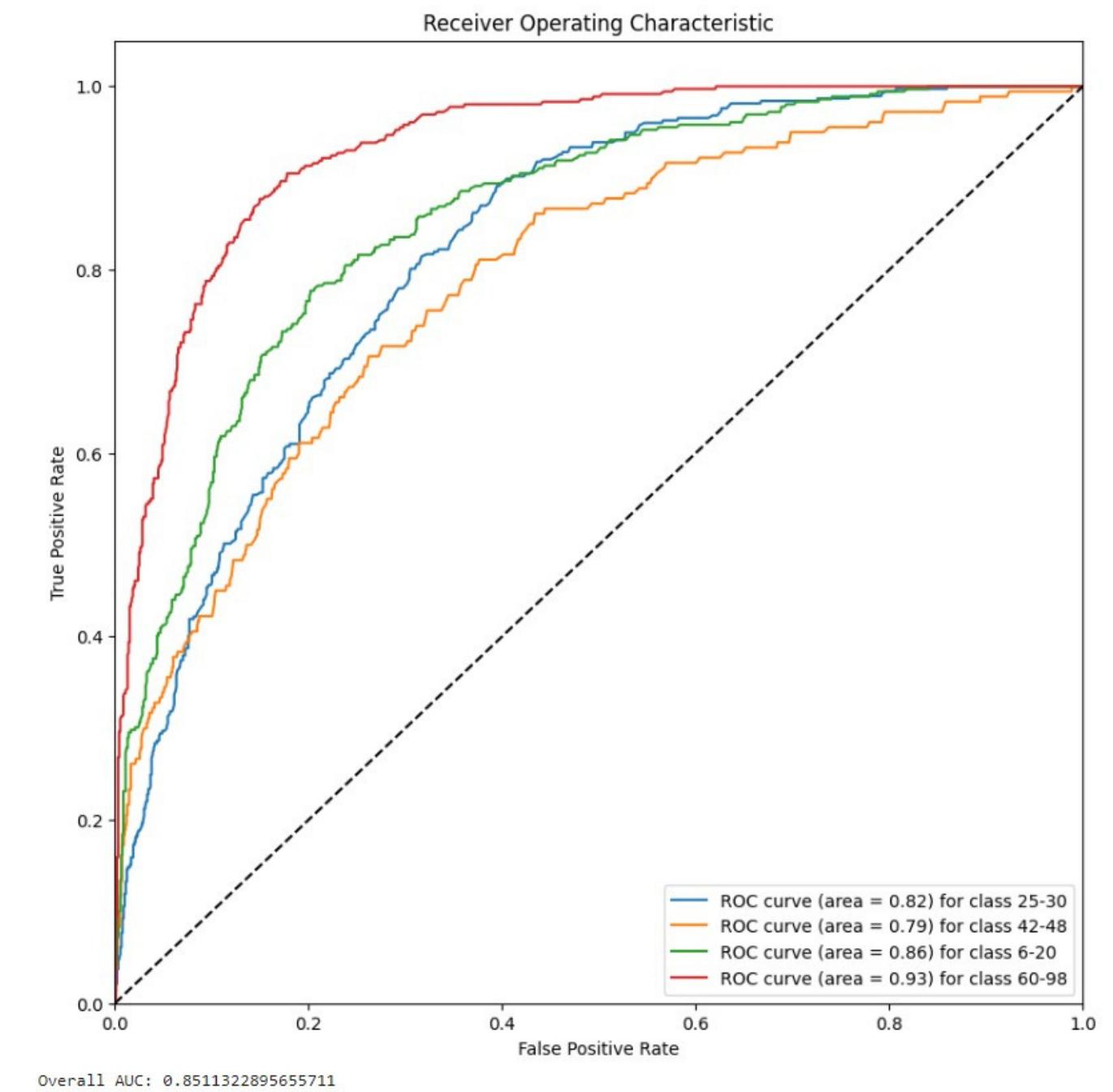
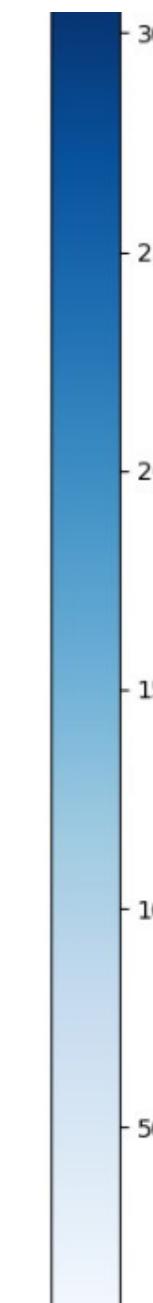
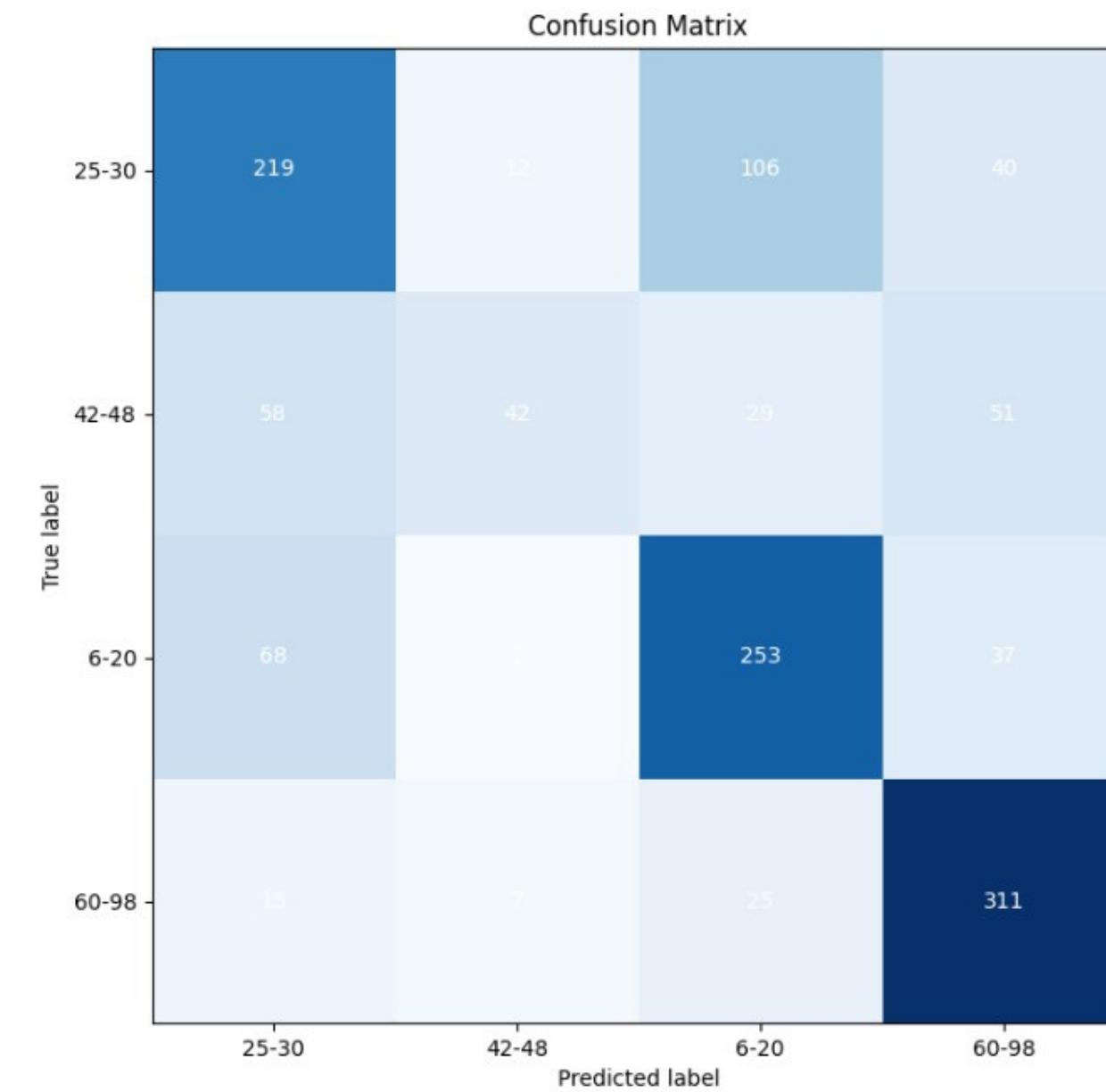
True Label: 60-98
Predicted Label: 60-98



True Label: 6-20
Predicted Label: 6-20



True Label: 42-48
Predicted Label: 60-98



Random forest classifier with bilateral filter

Accuracy: 0.6475667189952904
True Label: 25-30
Predicted Label: 25-30

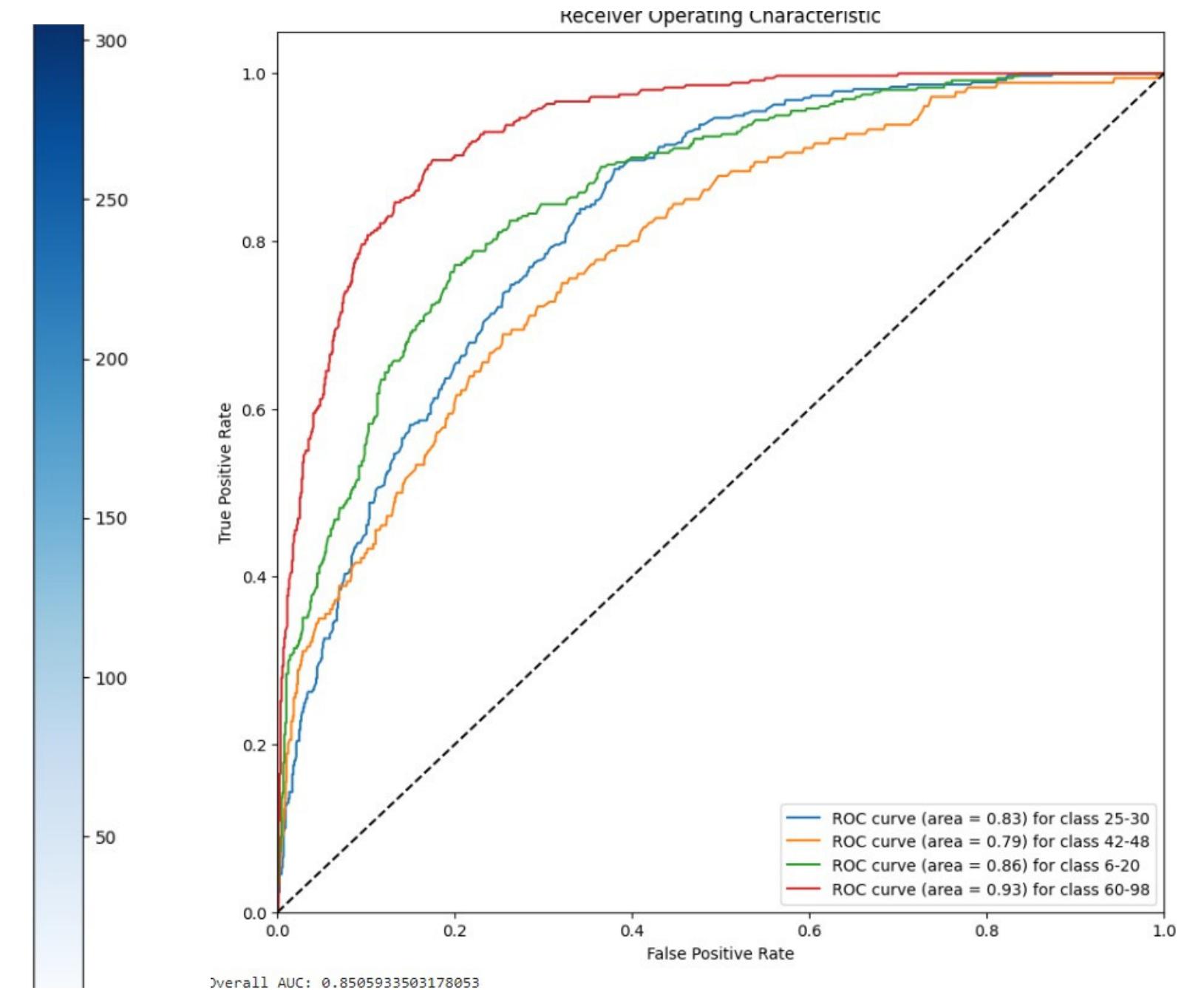
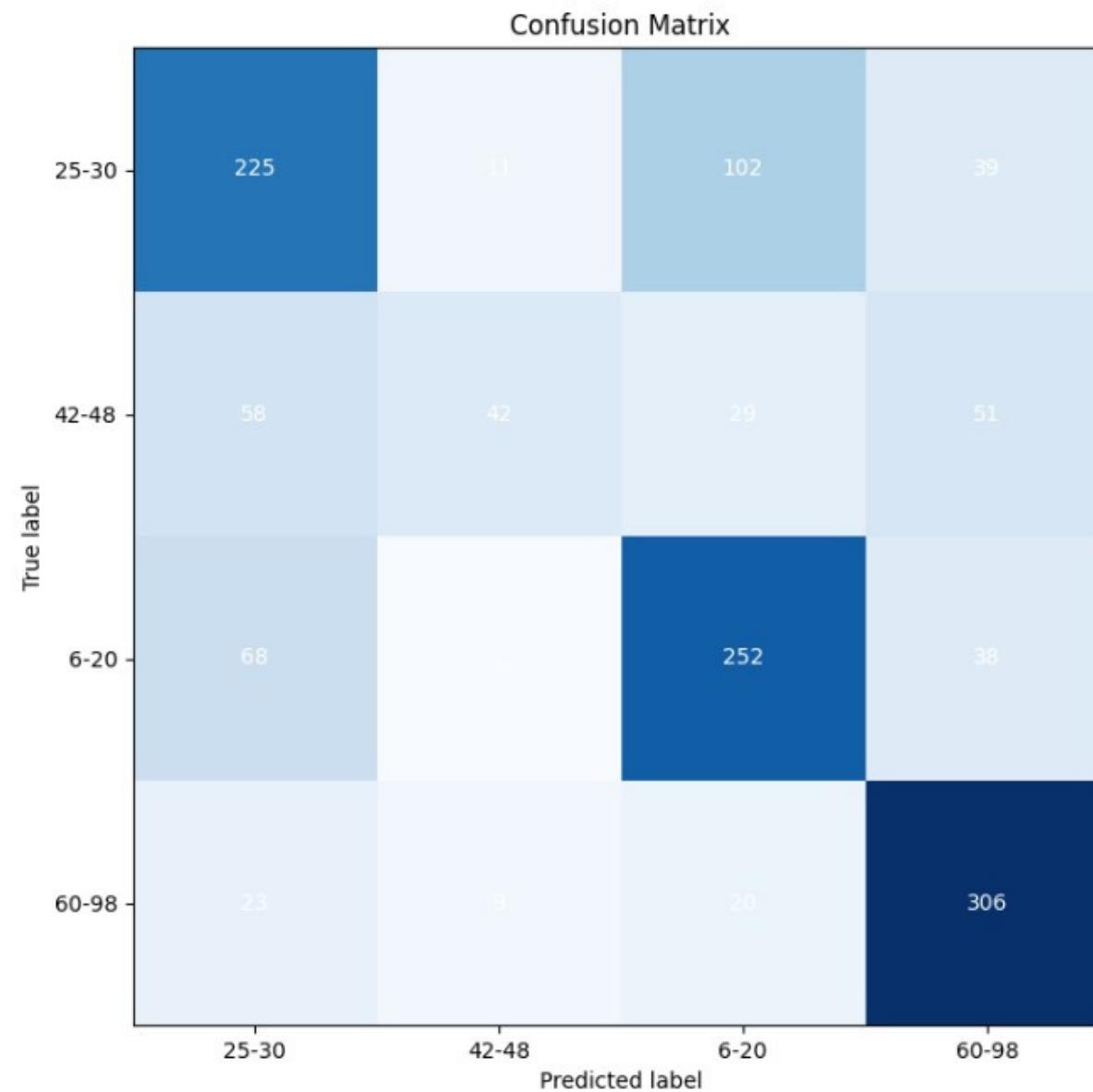

True Label: 6-20
Predicted Label: 6-20


True Label: 60-98
Predicted Label: 60-98


True Label: 42-48
Predicted Label: 25-30


True Label: 60-98
Predicted Label: 60-98


True Label: 6-20
Predicted Label: 25-30

Ensemble Classifications

- Stacking Classifier
- Bagging Classifier
- Boosting Classifier

Stacking classifier with LogisticRegression, KNeighborsClassifier,DecisionTreeClassifier

Accuracy: 0.4521193092621664

True Label: 42-48

Predicted Label: 25-30



True Label: 25-30

Predicted Label: 25-30



True Label: 6-20

Predicted Label: 6-20



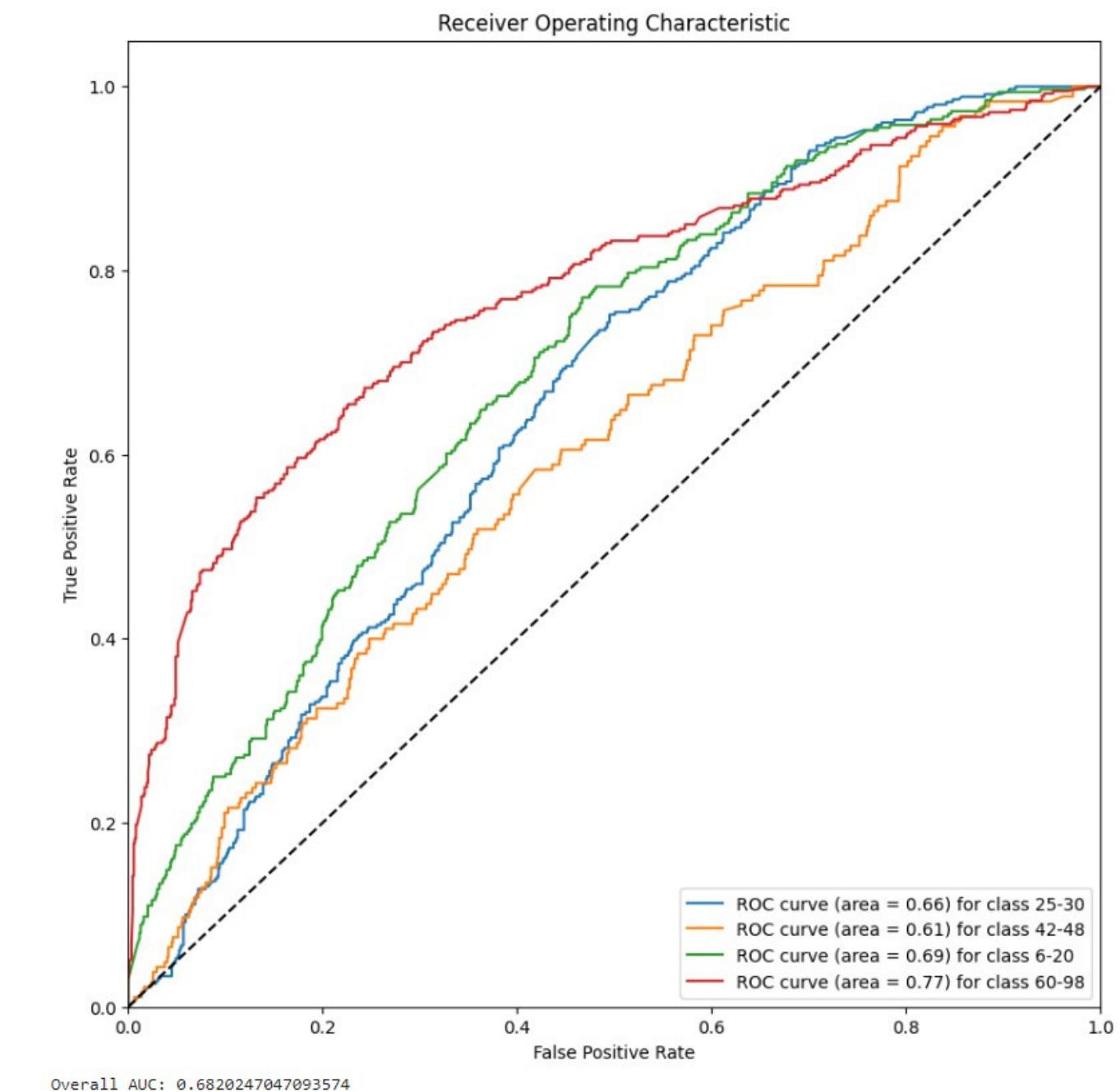
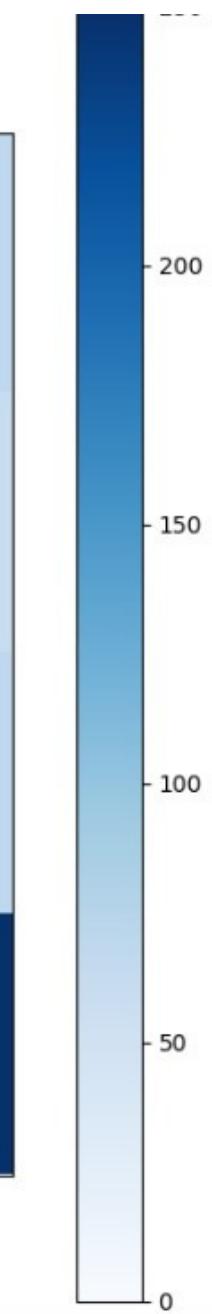
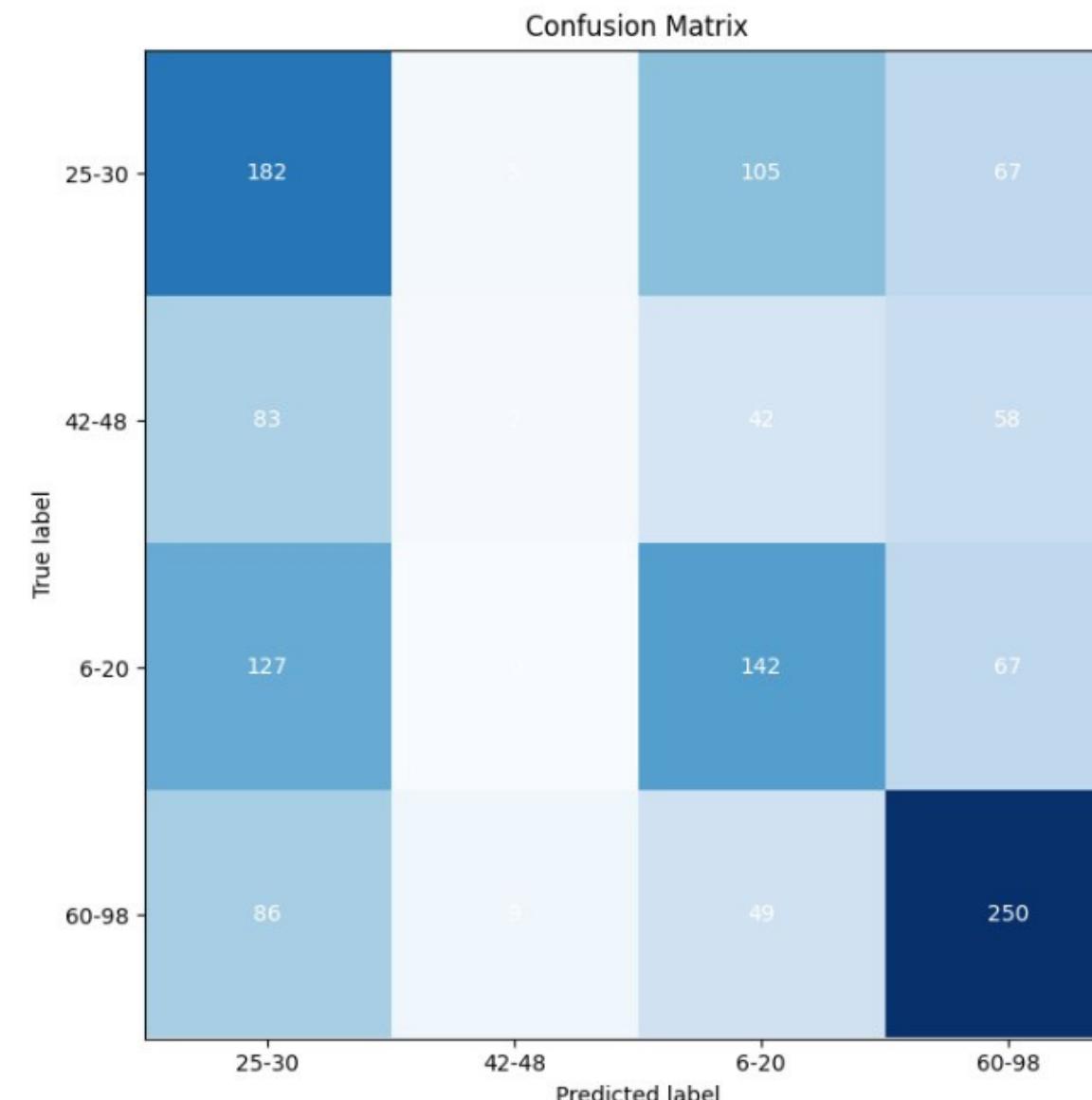
True Label: 42-48

Predicted Label: 25-30



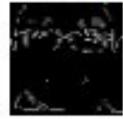
True Label: 6-20

Predicted Label: 25-30



Bagging with LogisticRegression, KNeighborsClassifier,DecisionTreeClassifier

Accuracy: 0.45682888540031397
True Label: 6-20



True Label: 25-30



True Label: 6-20



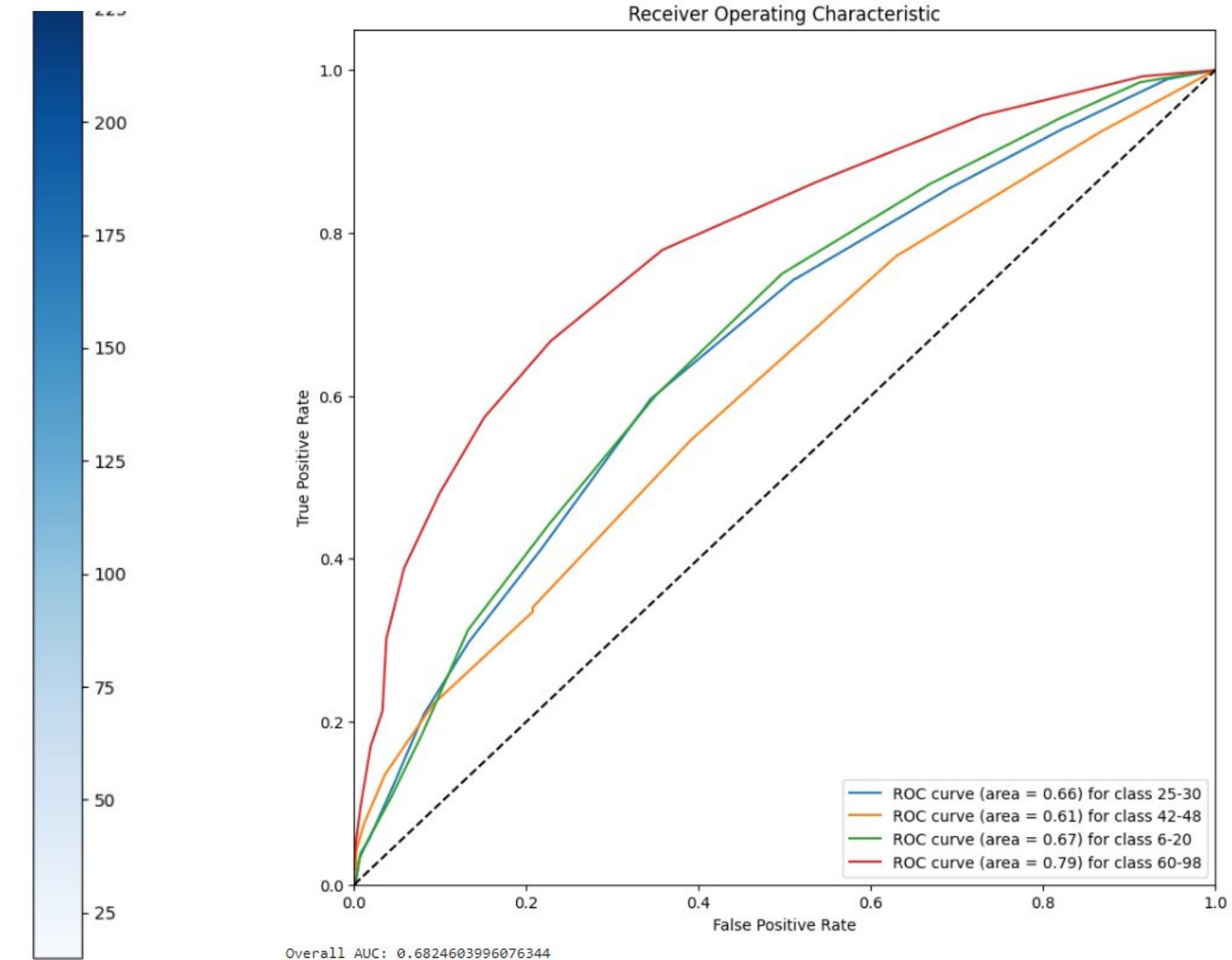
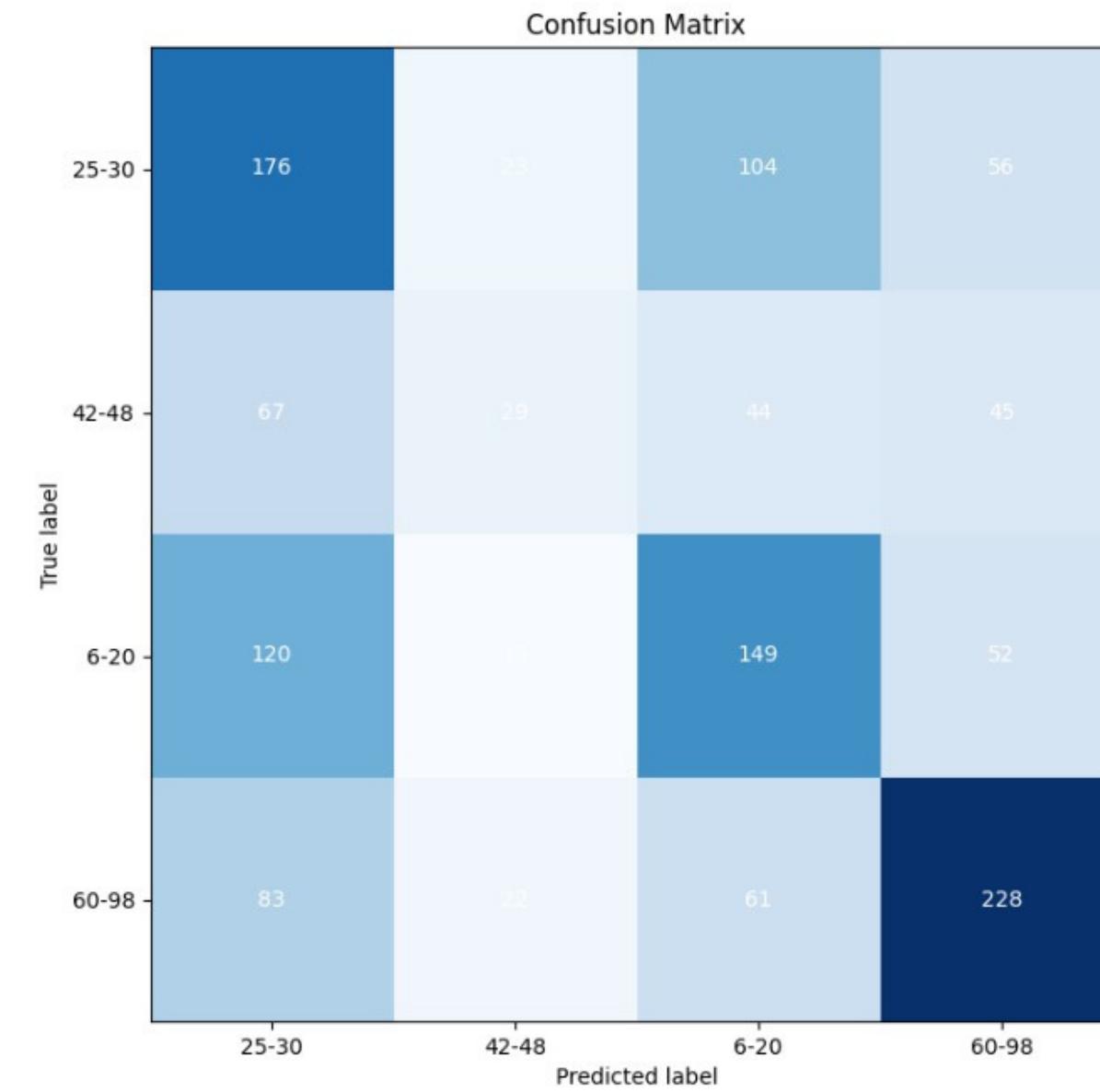
True Label: 6-20



True Label: 42-48



True Label: 60-98



Boosting classifier with LogisticRegression, KNeighborsClassifier,DecisionTreeClassifier

→ Accuracy: 0.41130298273155413

True Label: 25-30



True Label: 60-98



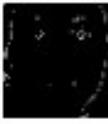
True Label: 6-20



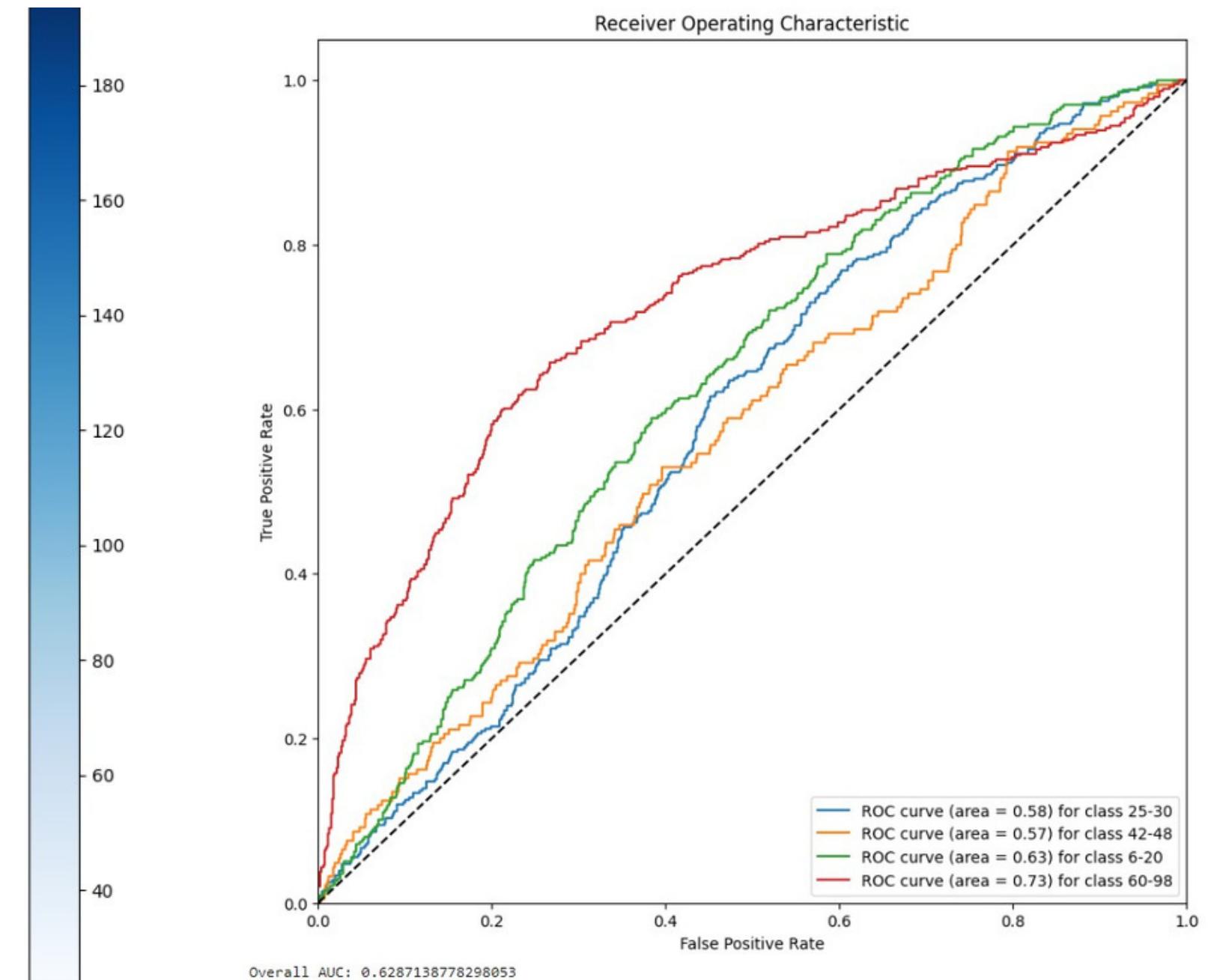
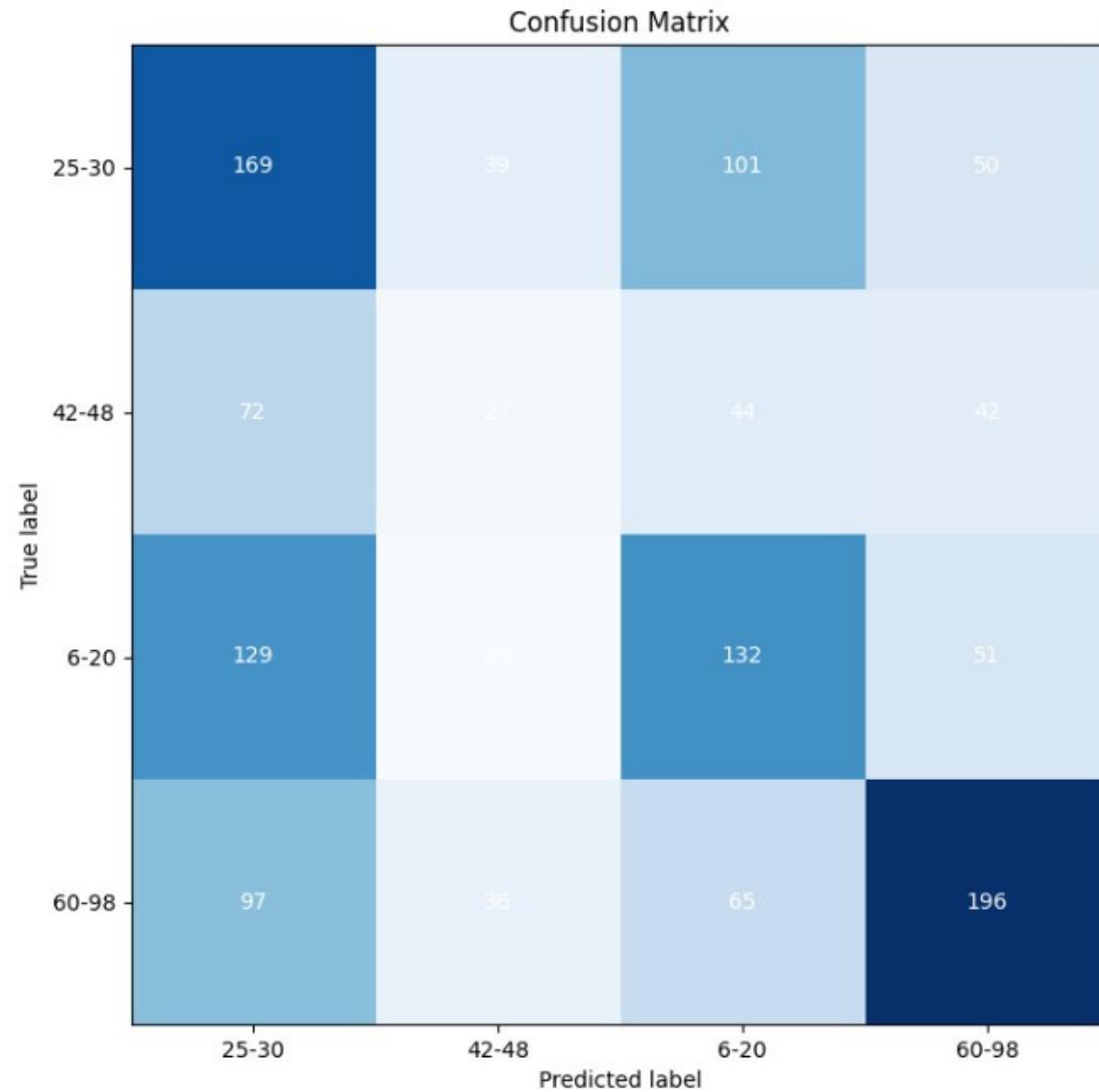
True Label: 6-20



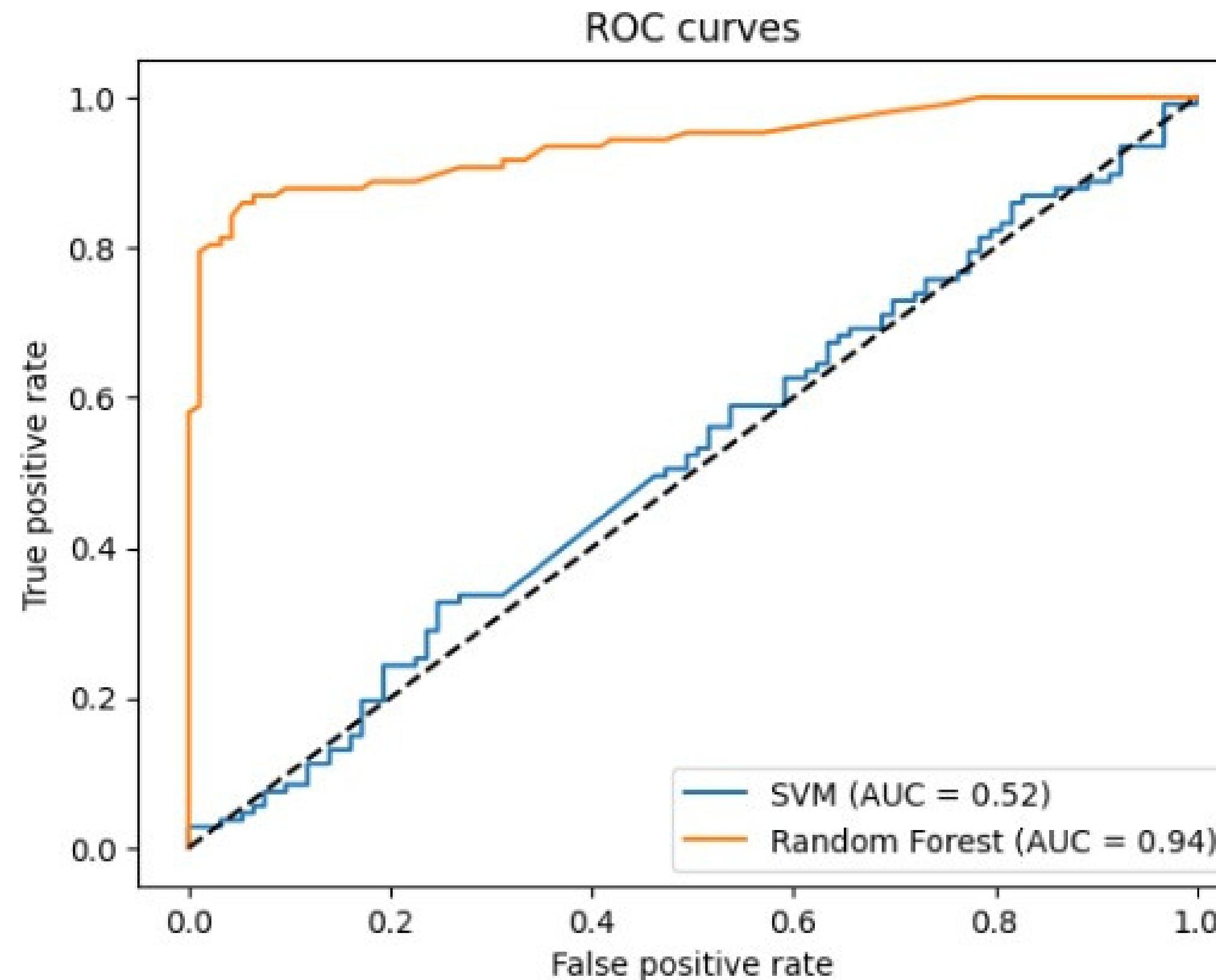
True Label: 25-30



True Label: 6-20



ROC & AUC



The background features a large, stylized 'X' shape composed of overlapping blue and white circles and rectangles. The design is minimalist and modern, with soft shadows creating depth.

Thank You