TuberBot

By: Christina Matthews, Casey O’Brien, Wanlong He, David Cabrera

What is TuberBot?

- **Goal**: The goal of this project is to create an application that quickly creates an avatar of the user without it being overly expensive and time-consuming.

- Many Vtubers want exact look-alikes for avatars but don't want to pay for this. This app makes it easier for them.

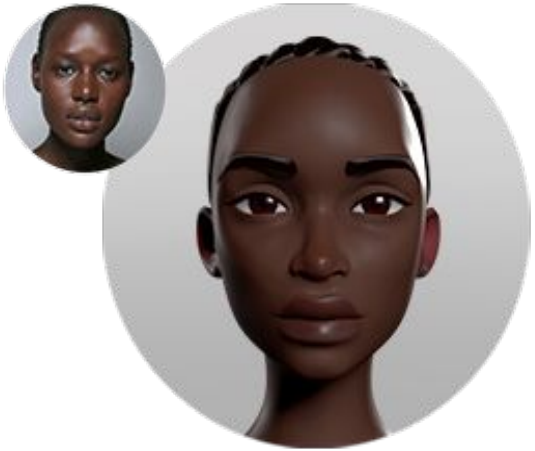
What is the goal and the Target Users for the Application?

- **Target Users**: Gamers, streamers, and people in their 20's.
- A woman in her early 20's who streams for a living and is interested in video games, anime, and Japanese culture.

Describe Similar AI Apps and Their Pros and Cons

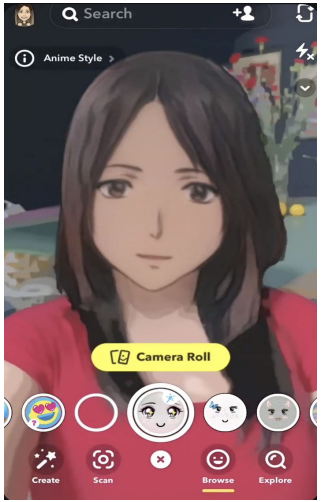
Pinscreen

- **Pros**: Free, face-tracking, quick rendering
- **Cons**: Doesn't create a personalized avatar, it simply imposes a photo of the user's face, avatars aren't liked by users



ToonMe

- **Pros**: Quickly takes images and turns them into cartoon images.
- **Cons**: Only works for still images and not facial recognition.



Snapchat Filter

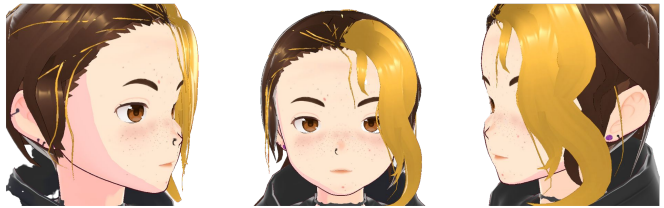
- **Pros**: Create customized anime faces based on the person's facial features and the filter makes sure it looks similar to your actual image.
- **Cons**: Highly dependent on camera angles and environmental lighting.

How Is It Different Than Its Competitors ?

Our app quickly creates an original face based off the user, whereas other apps imposes on the user's face and places it on the user. Our app helps users save time and money.

VRoid Studio Creation

- As a proof of concept, and also to validate our claims about the effort it takes to create a Vtuber, we built a Vtuber model ourselves.
- Using a free software, VRoid Studio, a Vtuber version of Casey was created.
  - This took 2 attempts and around 4 hours to ensure it was as accurate as possible using the anime style model.



The Prototype

The Research Behind TuberBot

1.A novel approach for face detection using hybrid skin color model.

2.“Robot Artist Performs Cartoon Style Facial Portrait Painting

3.Encoding in style: A styleGAN encoder for image-to-image translation.

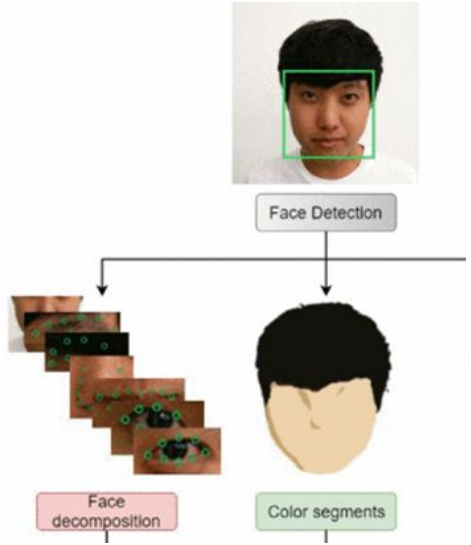
Paper#1: Face Detection Challenges

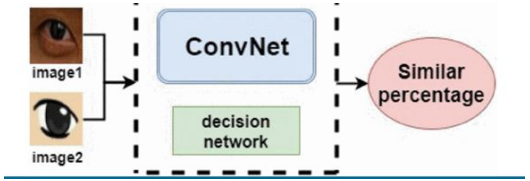
1. Pose / Camera - Face Angle
2. External Facial Features (Glasses, Ears, Beards)
3. Facial Expressions (Sad, Happy, Angry)
4. Environment Lighting / Camera Conditions

Yadav, S., Nain, N. A novel approach for face detection using hybrid skin color model. *J Reliable Intell Environ* 2, 145–158 (2016). https://doi.org/10.1007/s40860-016-0024-8

Paper #2: Face Decomposition Detection

1. Face Decomposition
    a. Left/Right Eye
    b. Eyebrows
    c. Face Shape
    d. Nose
    e. Mouth
2. Identify Each Individual Facial Feature Type
    a. Right Eye Color (Green / Brown/ Blue)
    b. Right Eye Shape (Almond / Monolid / Hooded)

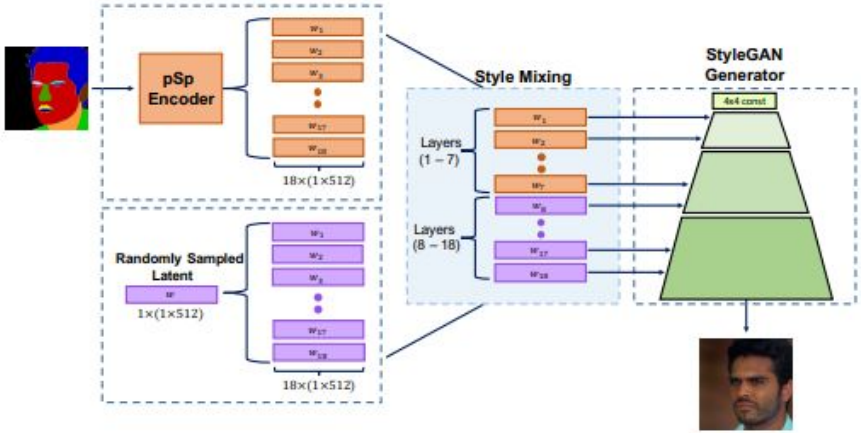


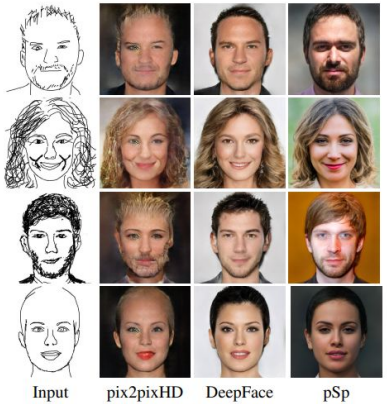


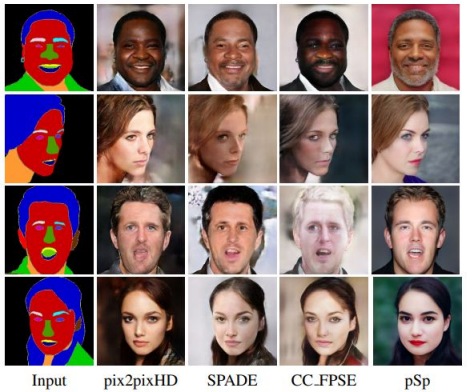
Luo, Ren C, and Yu Jung Liu. “Robot Artist Performs Cartoon Style Facial Portrait Painting.” 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018. 7683–7688. Web.

Paper#3: StyleGAN Image Translation Framework

1. Face from Sketch
2. Face from Segmentation Map



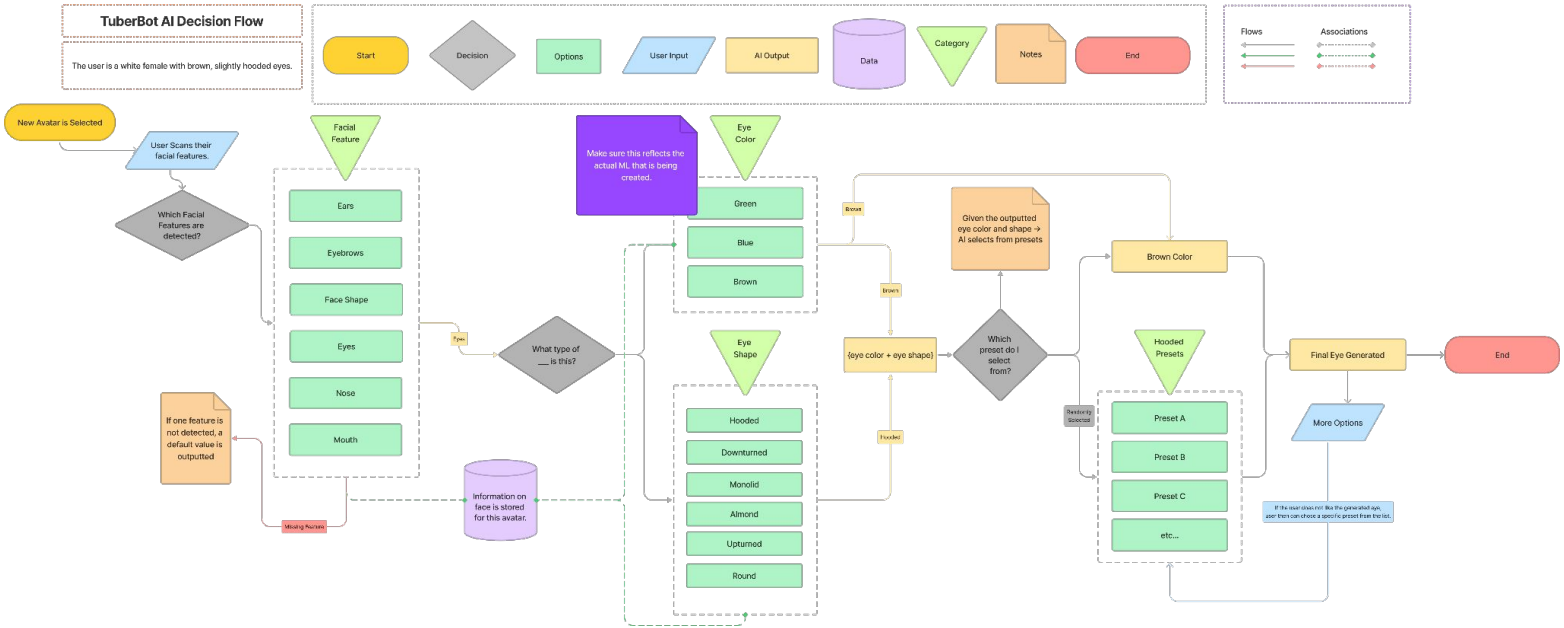




Richardson, E., Alaluf, Y., Patashnik, O., Nitzan, Y., Azar, Y., Shapiro, S., & Cohen-Or, D. (2021, April 21). Encoding in style: A stylegan encoder for image-to-image translation. arXiv.org. Retrieved December 11, 2022, from https://arxiv.org/abs/2008.00951

The Technical Background Behind TuberBot (Figjam) I

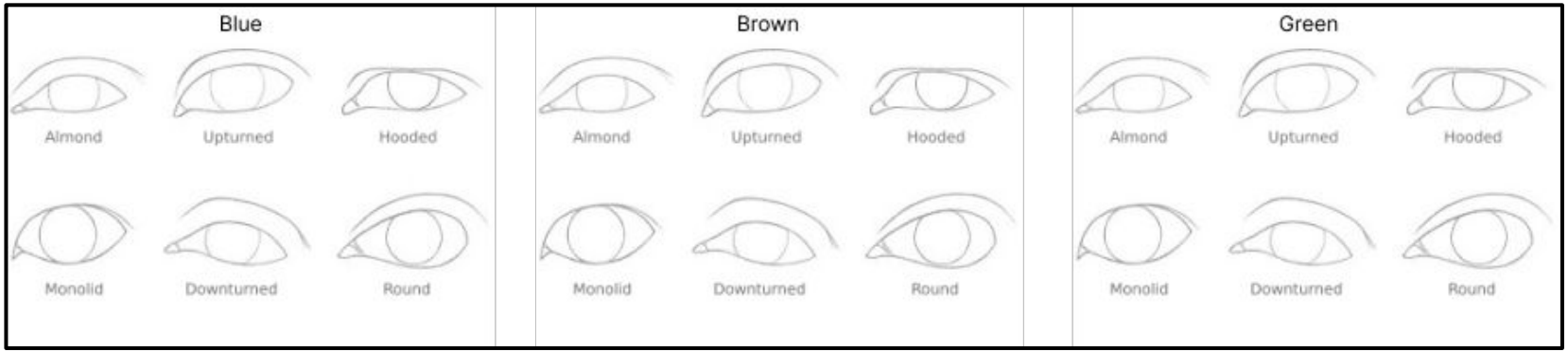
The AI Logic Flow



The Technical Background Behind TuberBot (ML) II

1.Data Collection - Method 1

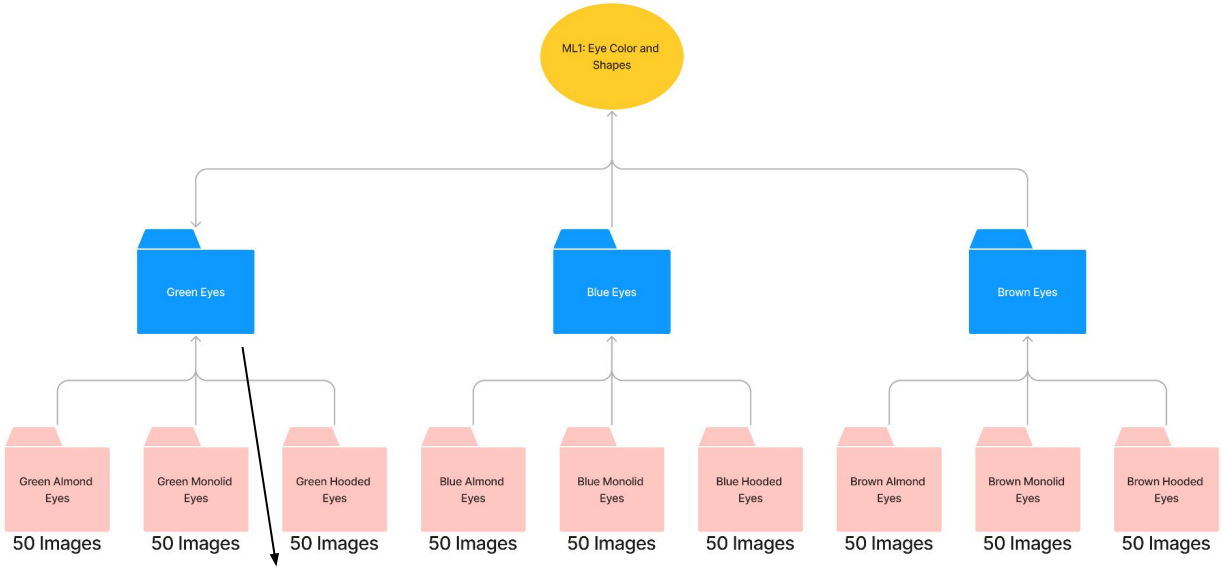
3 eye color file along with 6 eye shapes each, each shape 50 images.

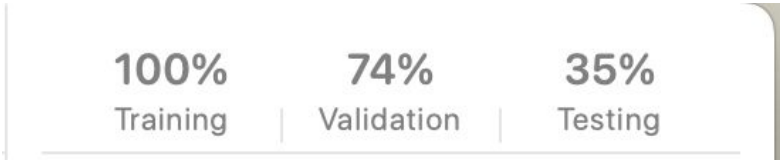


Challenges:

- A lot data needs to be collected (900 images) with no database available
- We’re not good at differentiating eye shapes

1.Data Collection - Method 2

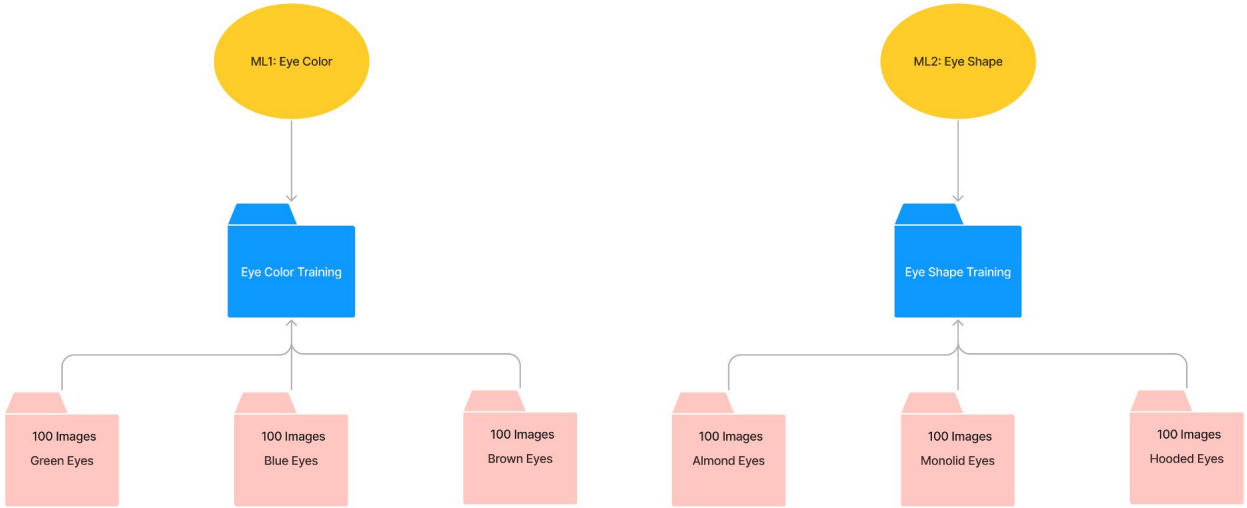




Challenges:

- Monolid eyes is most common among brown eyes.
- It's hard to find monolid eyes with green & blue color
- Create ML doesn't allow more than 2 levels of files being trained together.

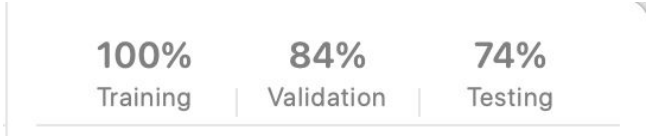
1.Data Collection - Method 3

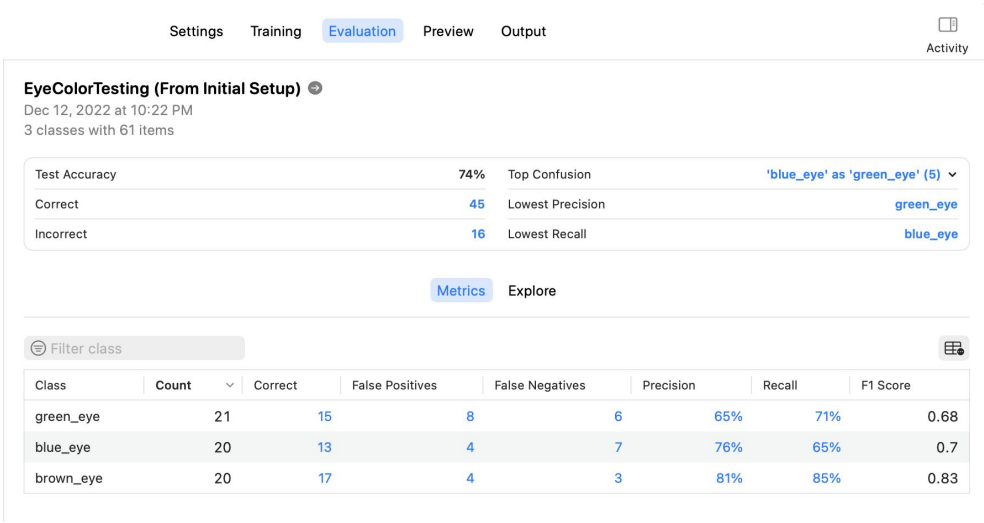


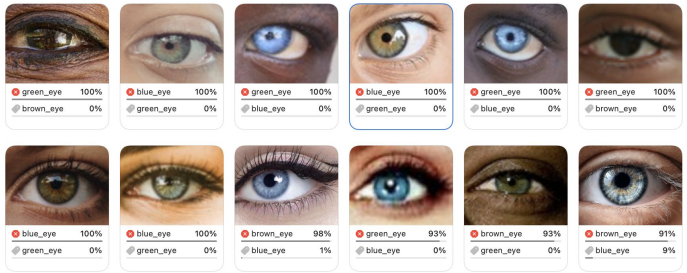
Advantages:

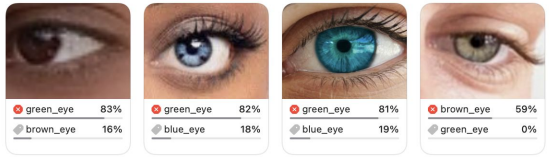
- Reflect on our AI decision flow
- Simple rearrange of data we collected with last method.

2.Eye Color Model ML Result - Part 1









2.Eye Color Model ML Result - Part 2

Analysis:

- Training, Validation and Testing have decent accuracy
- Low precision on green eye, low recall on blue eyes

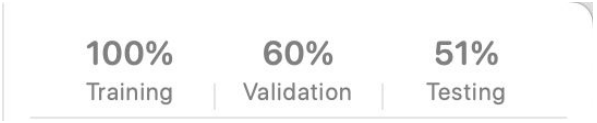
Challenges:

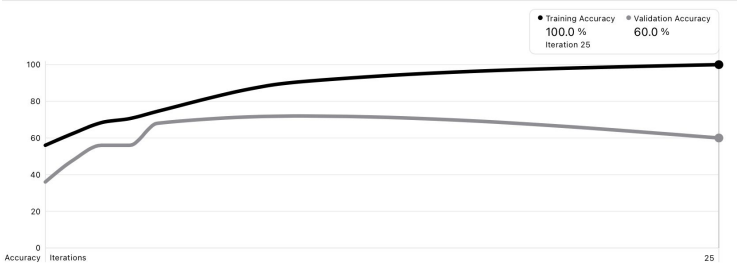
- The model struggle to identify eyes with darker skin - not enough skin tone diversity in green / blue eye image training
- Blue and Green eyes are hard to differentiate - similar color
- Green and Brown Eyes are hard to differentiate - included hazel as green.

Improvement:

- More diversified image training data
- Include more eye color (hazel, gray)

3.Eye Shape ML Result - Part 1





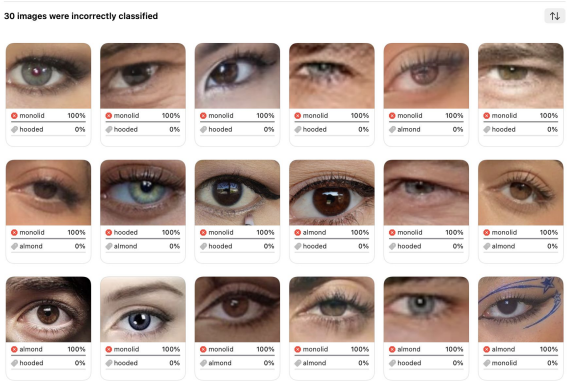
EyeShapeTesting (From Initial Setup)
Dec 13, 2022 at 6:35 PM
3 classes with 61 items

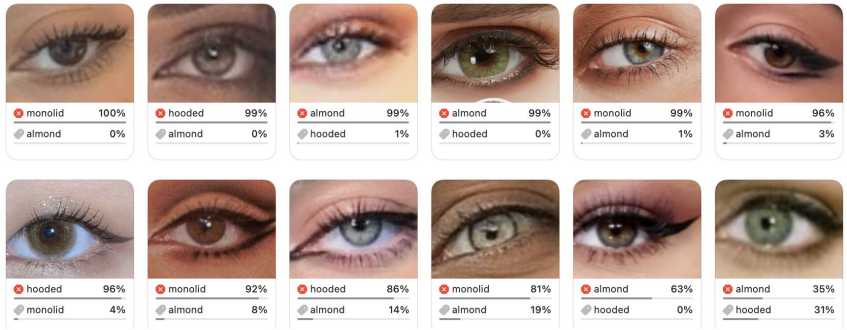
Test Accuracy	51%	Top Confusion	'almond' as 'monolid' (10)
Correct	31	Lowest Precision	monolid
Incorrect	30	Lowest Recall	hooded

Metrics | Explore

Filter class

Class	Count	Correct	False Positives	False Negatives	Precision	Recall	F1 Score
hooded	21	6	4	15	60%	29%	0.39
almond	20	7	8	13	47%	35%	0.4
monolid	20	18	18	2	50%	90%	0.64





3.Eye Shape ML Result - Part 2

Analysis:

- Validation and Testing have worse accuracy
- Low precision on monolid eyes, low recall on hooded eyes

Challenges:

- Differentiate eye shape is hard - we can't guarantee the training & testing images are being sorted correctly (human error)
- Image condition varies (quality, lighting, angle, facial expression etc..) and it matters for the training - the difference between shapes are not as obvious as eye colors.

Improvement:

- More accurate eye shape image collection
- Standardized training & testing image conditions

Conclusion / Summary

- It's a practical idea
  - Possible to build out and design
  - Doesn't need complicated AI
  - Fits the current online trend and live streaming market.
- Development Challenges
  - Require larger quantity of training database that are correctly collected
  - Require lots of human effort to create 3D presets that fits all facial features
- What should we do in the future?
  - Dig deeper into facial expressions' effect on facial features
  - Collab with fashion group to incorporate their AI on outfits.

ML Testing

Thanks for Listening!