

PostSelectML

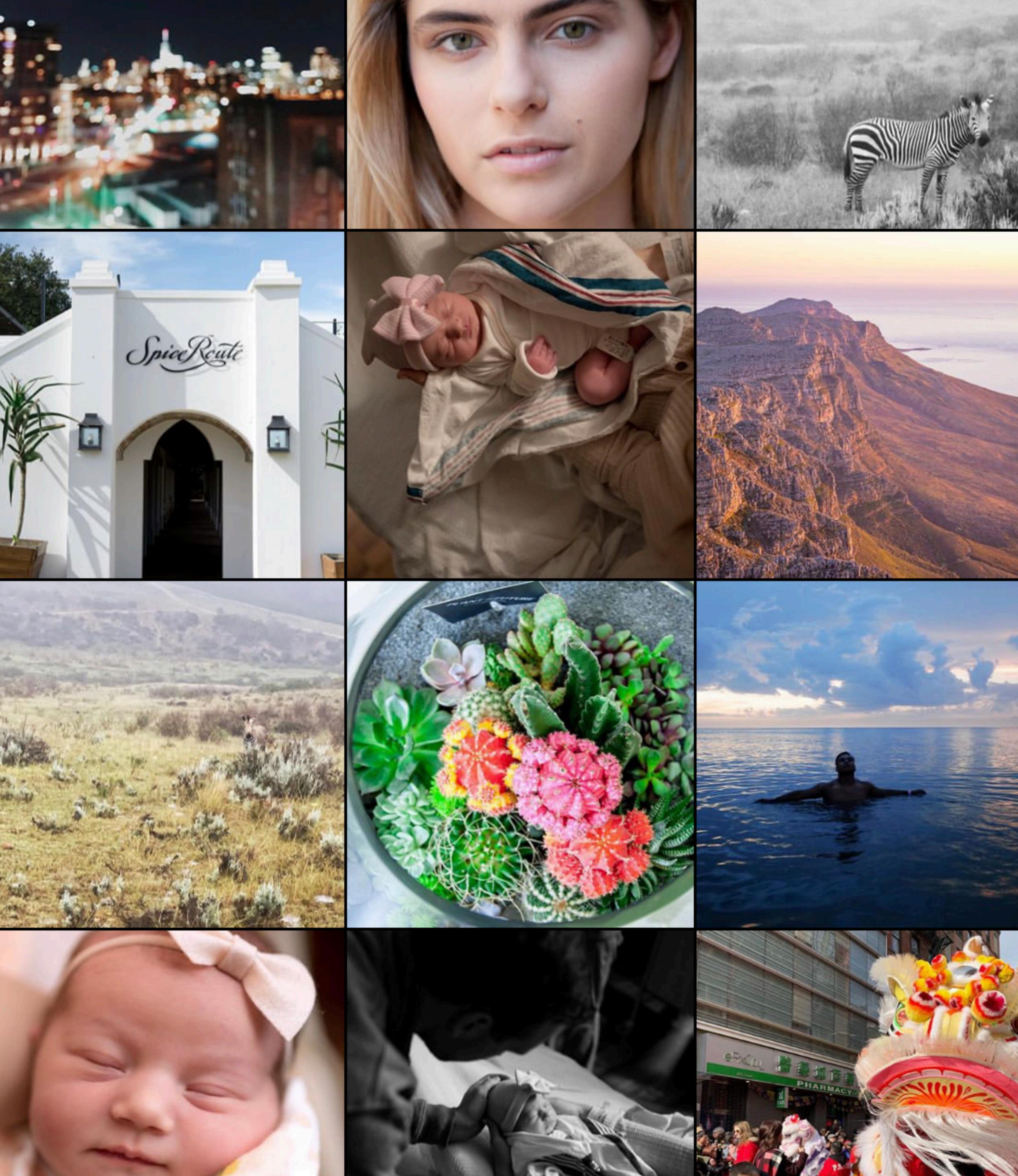
Using CNNs for image classification
in the post processing of photographs

Dolci Key
September 2020

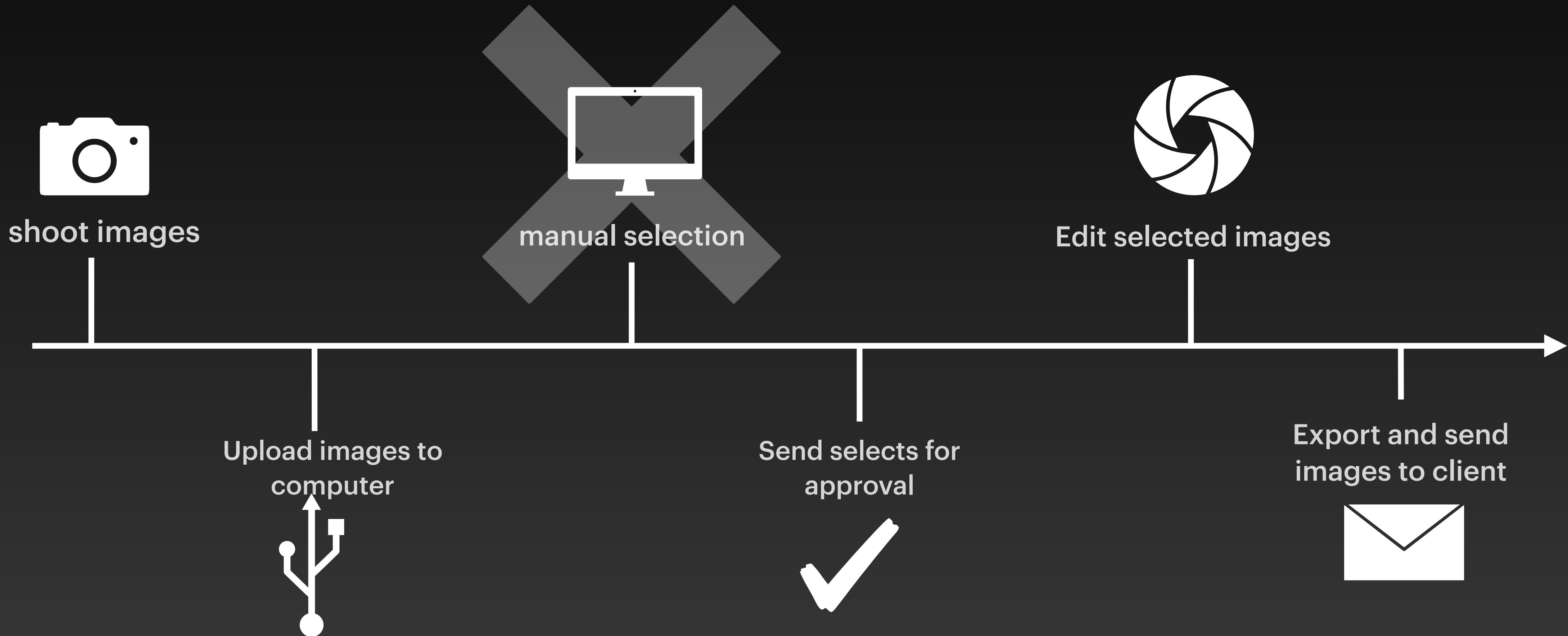
Content is today's marketing currency.

Photography

- Today's marketing relies heavily on a flow of image content.
- Good photos are called Selects.
- Sorting through 1000 photos for the best images can hours.



Photographers' Work Flow



Sorting through Unedited Images

images by @dolcikeyphotography



Reject

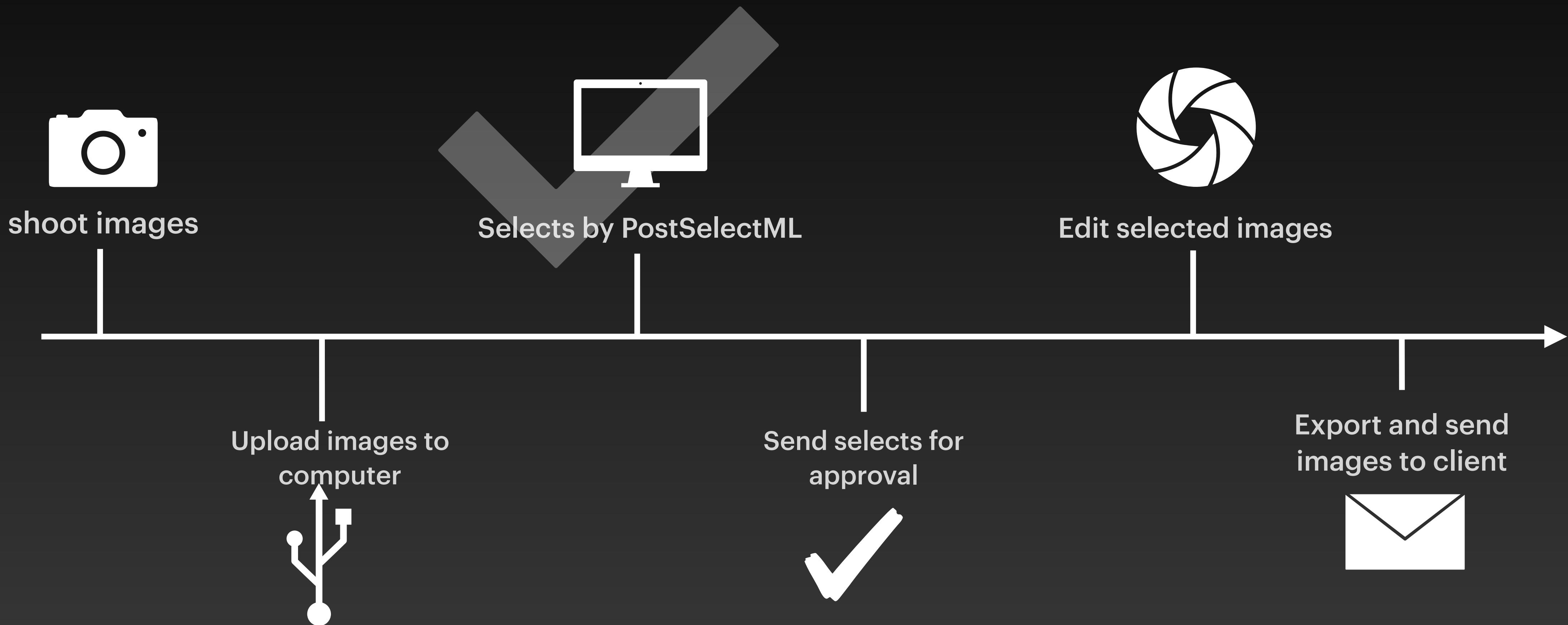


Select

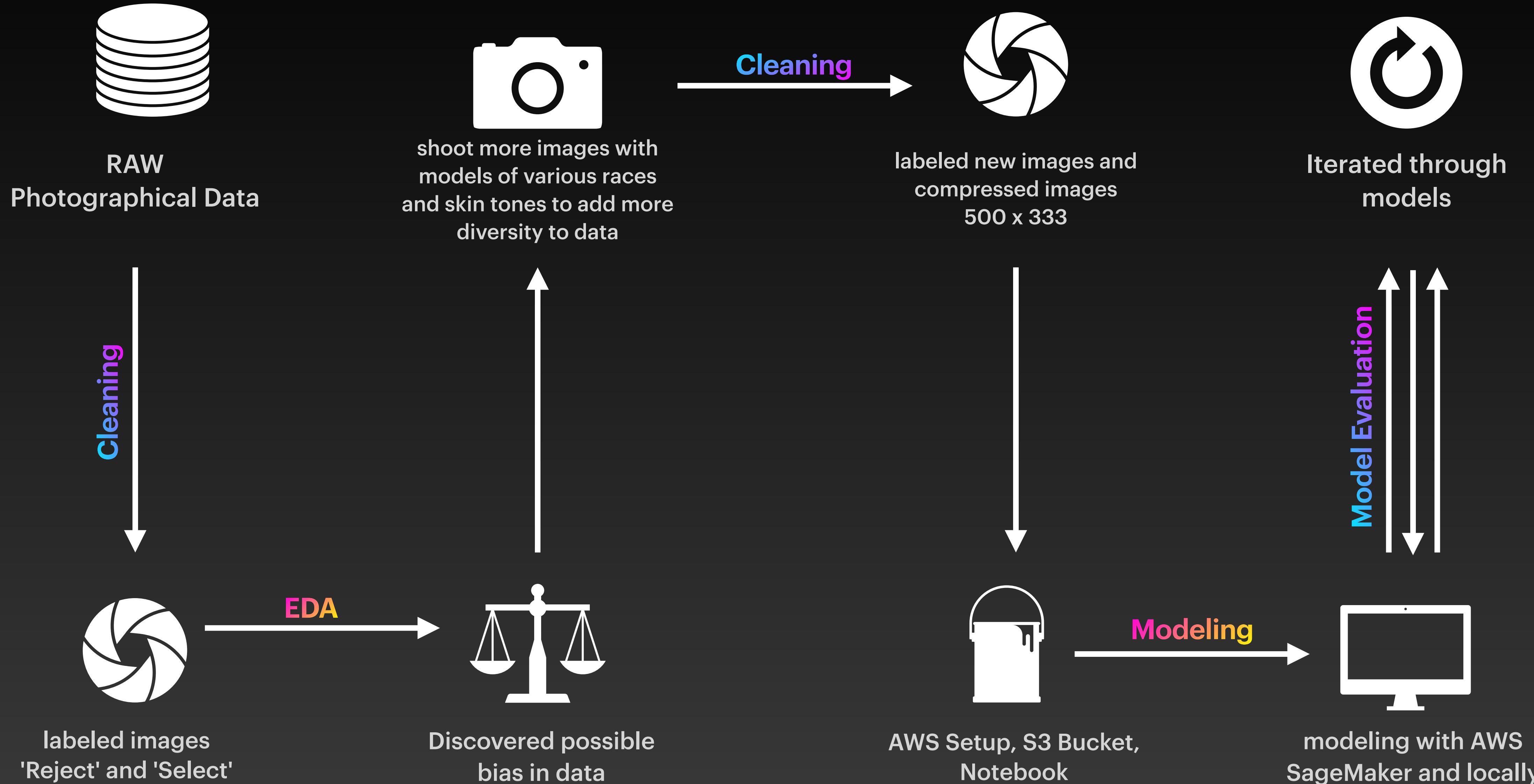


Reject

With PostSelectML



CRISP-DM Data Journey



3,421

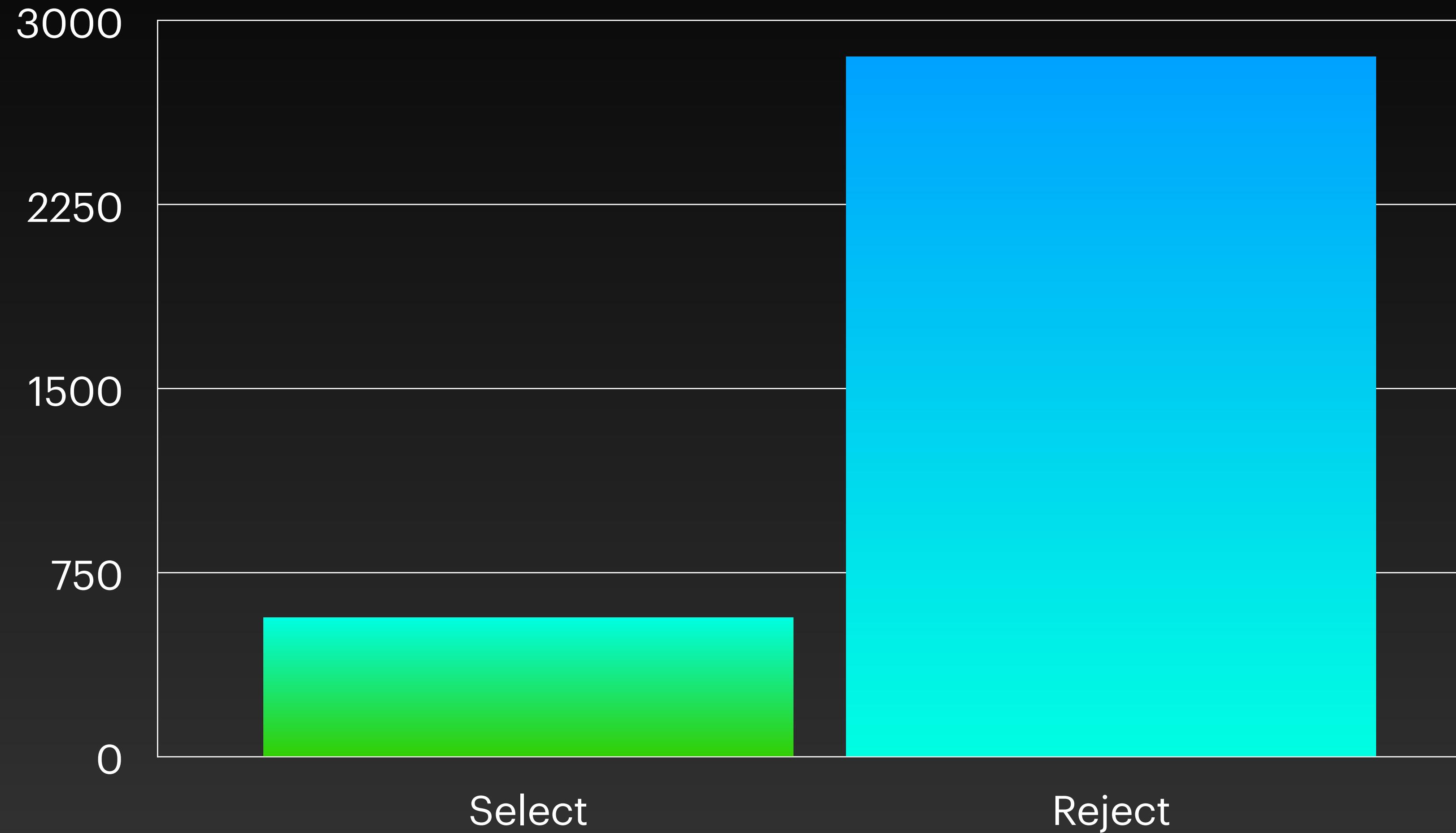
total images

Data Considerations

- Dealing with Class Imbalance
- Minimizing Racial Bias
- Acknowledging Artistic Bias



Classification of Data



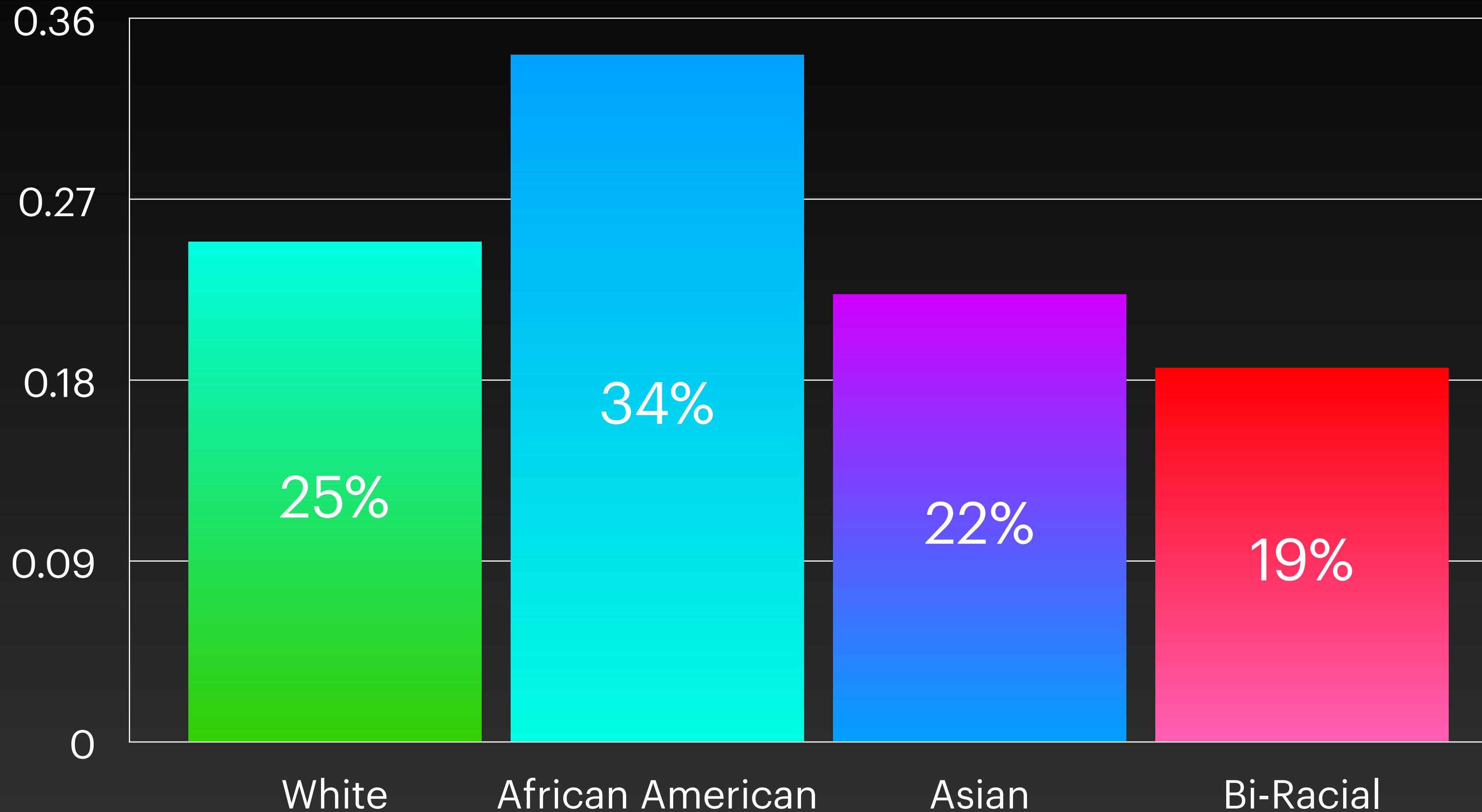
Select being the positive class (good images), Reject being the negative class (bad images)

Minority Class Considerations

- Horizontal/ other data augmentation
- Weighting Classes
- Metrics: Recall



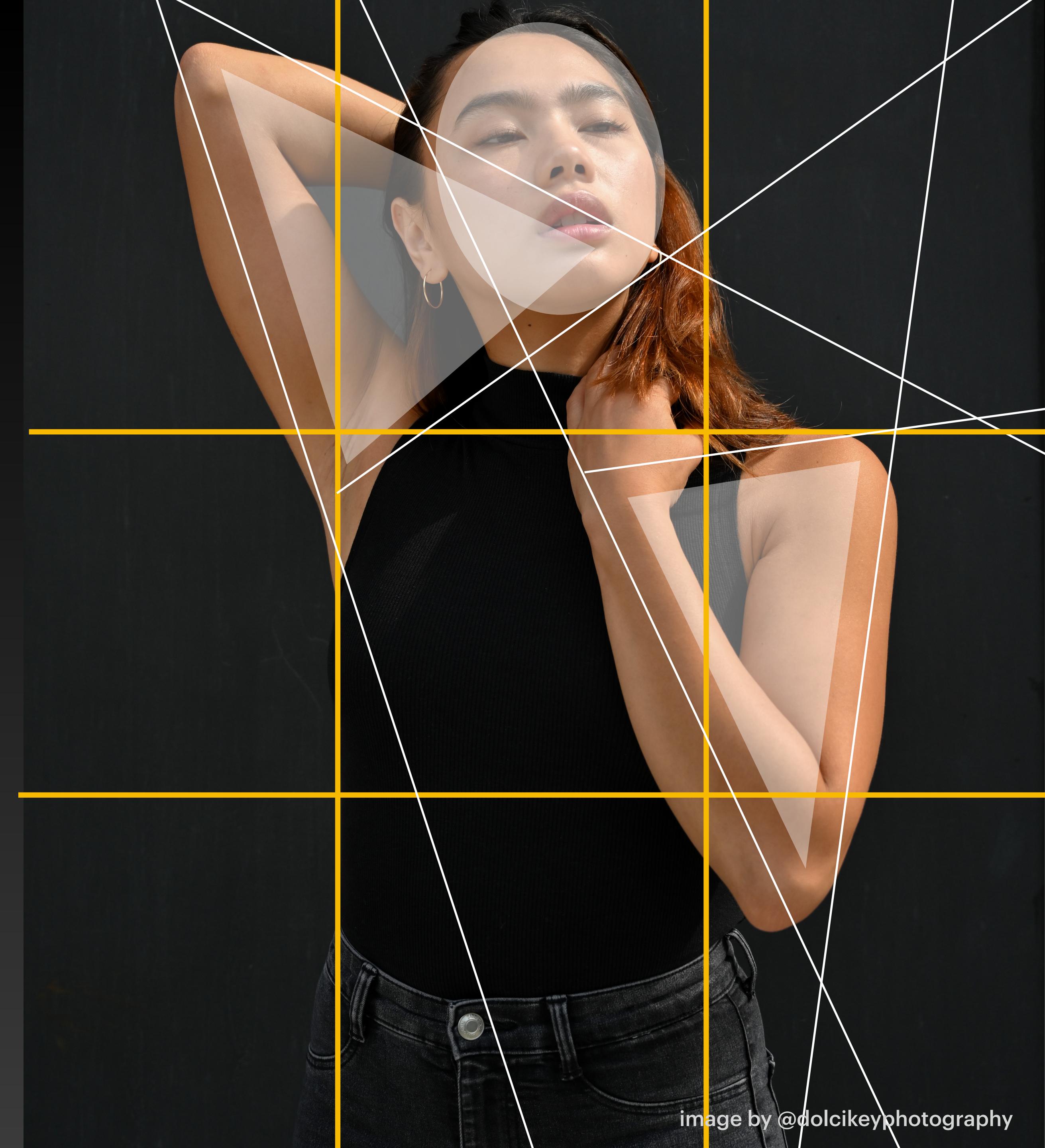
Exploring the Racial Breakdown of Overall Data



CNNs can become racially bias when there is not enough representation of a races in a training data set. Including images with diverse subjects will help make sure the model is not racist and can determine a good or bad photo regardless of the race of the subject.

Artistic Bias

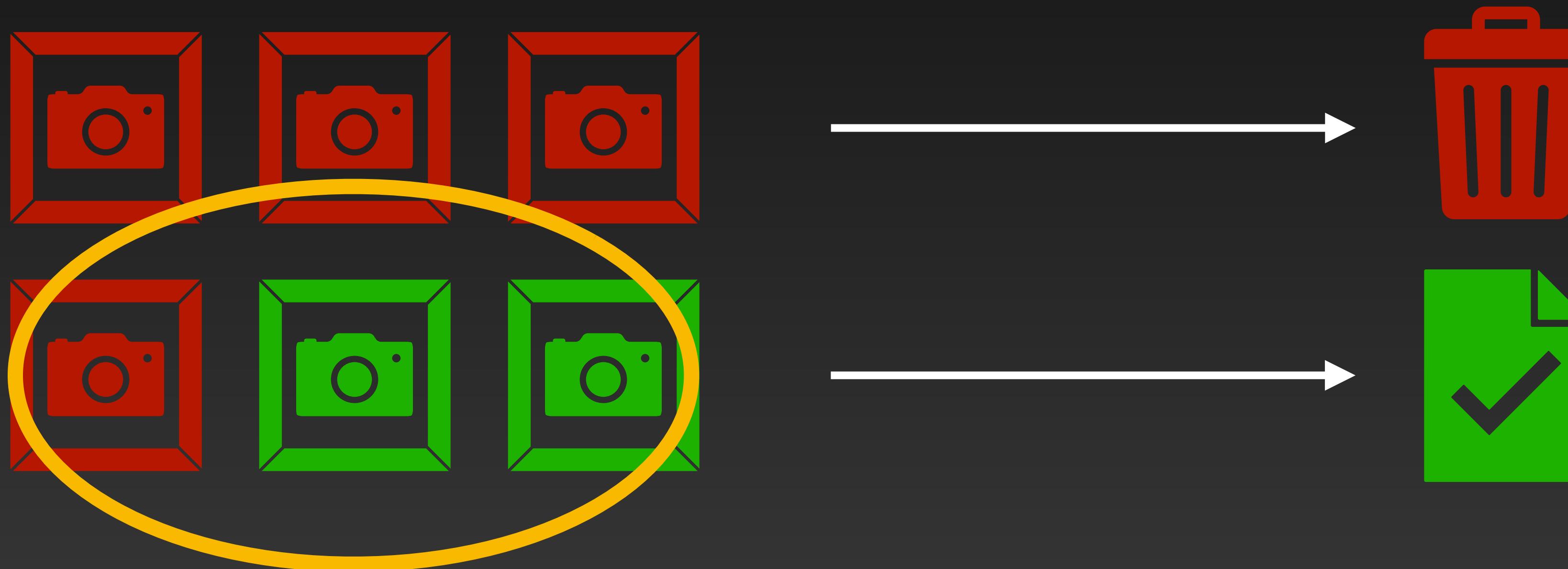
When selecting images, I relied on what I've learned over years of being both behind and in front of the camera. Prioritizing focused eyes/faces, balanced light, framing rules (rule of thirds), consideration of geometric shapes.



Metrics

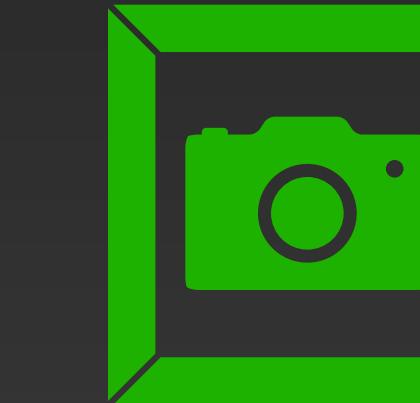
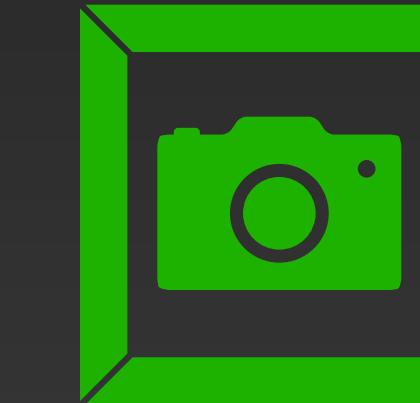
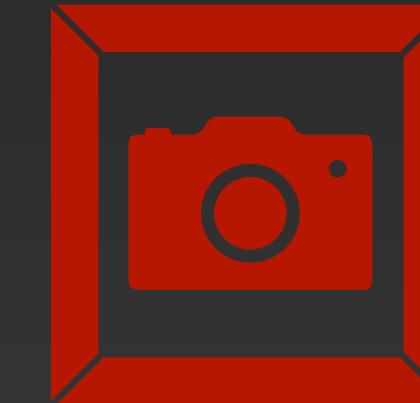
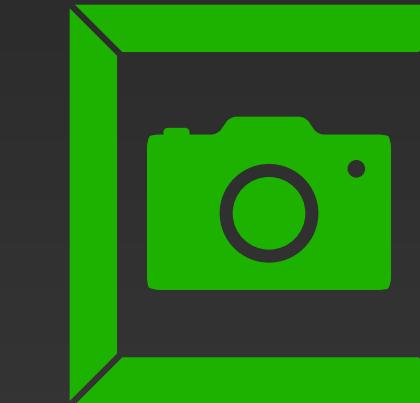
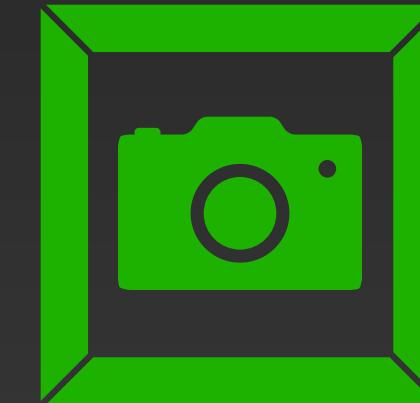
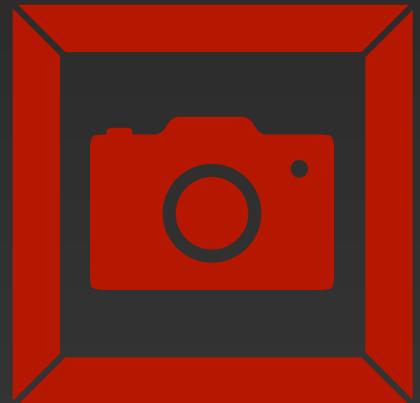
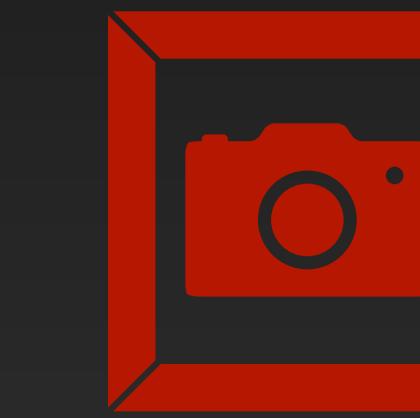
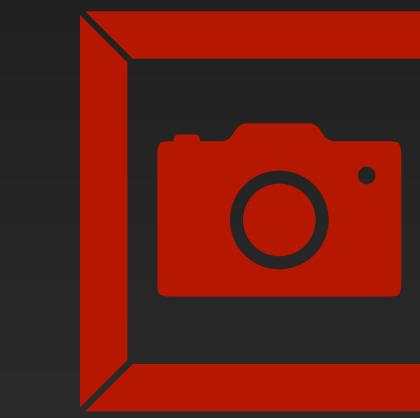
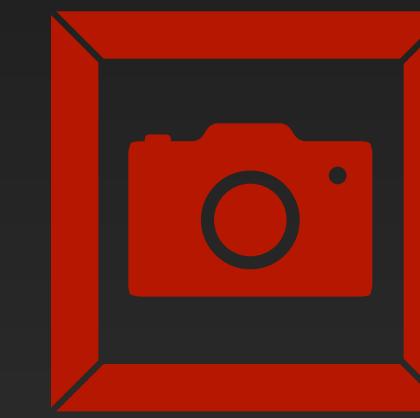
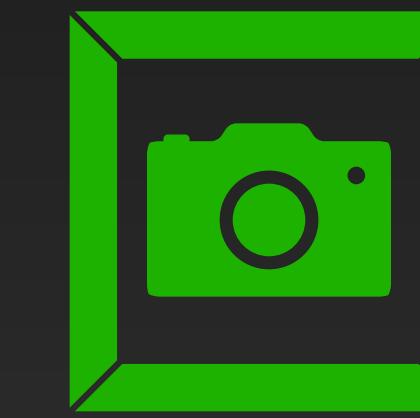
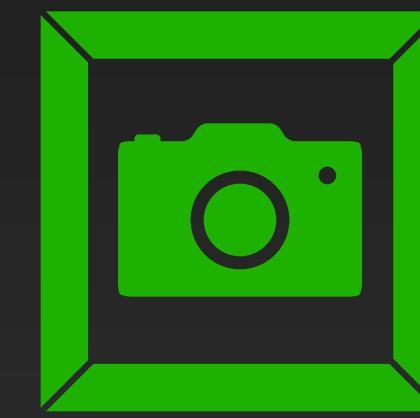
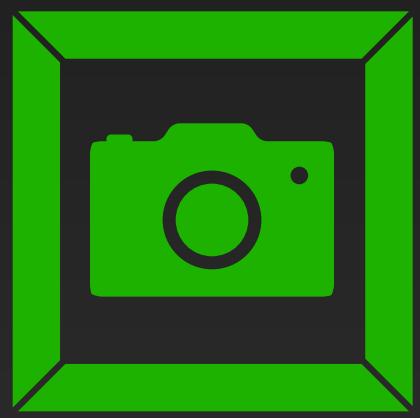
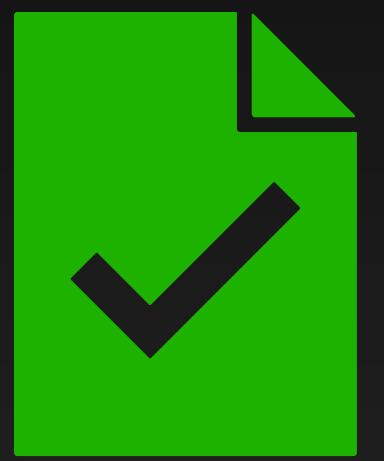
Recall

I would rather get a **bad image** predicted as a **good image**
rather than
get a **good image** predicted as a **bad image**

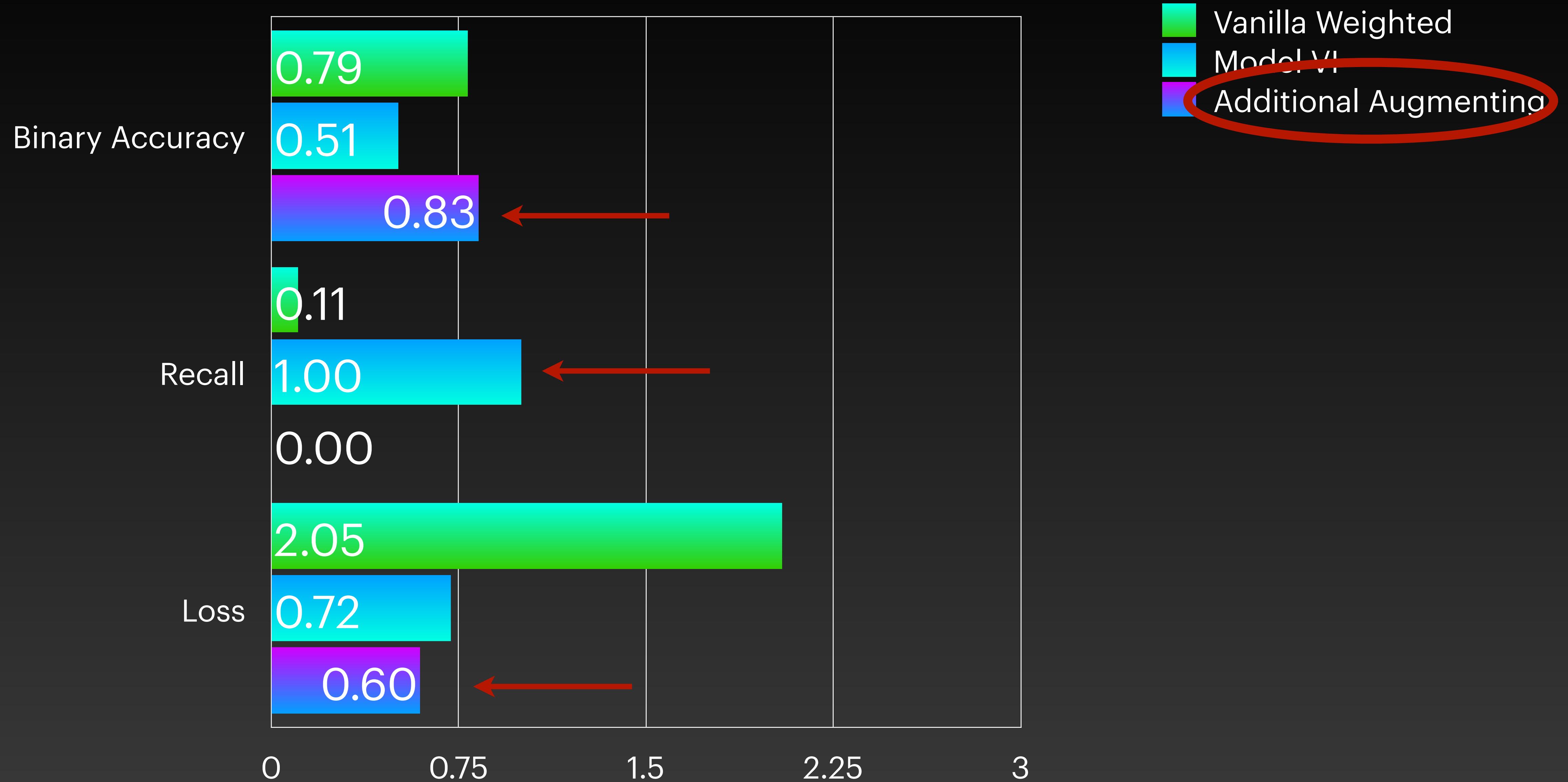


Binary Accuracy

How accurate are the predictions?



Model Performance



Testing the Unseen Data

100%



SELECT

99%



REJECT

99%



REJECT

Conclusion

Concluding Observations

- This model as a proof of concept shows promise.
- Predicting photography is complex
- This model stands to improve on both accuracy and recall when it comes to predicting on unseen data and would likely benefit from more data.



Further Steps

- More data (5-10k images)
- More models (subjects) with focus on a wide spectrum of diverse skintones and body types and sizes
- More randomized batch testing



image by @dolcikeyphotography

Future Business Plans

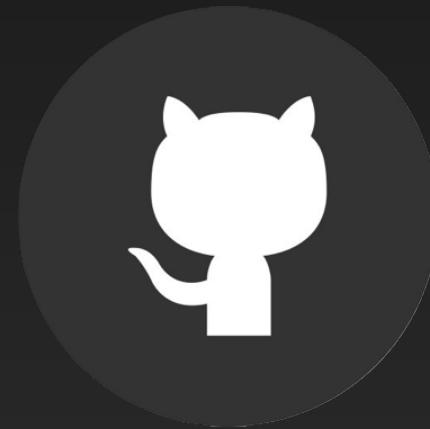
- Additional types of images like families, weddings, animals, food, products, e-commerce
- An integrated Script for straight NEF or JPG compression for photographers
- User friendly application allowing for unsupervised selection process

Where to Find Me

Thanks for watching/reading this presentation.
Follow me on the socials, and feel free to send me any questions or
feedback.



medium



github



instagram

PostSelectML GitHub Repository

dolci@dolcikey.com

Sources and Credits

- modeling coding references
 - Paul Breton AWS SageMaker Tutorial on [Medium.com](#)
- Troubleshooting and code debugging help from [Aren Carpenter](#)
- models (A huge thanks to all these talented models who made themselves available for data collection*)
 - Samayah Jaramillo [@samayahjaramillo](#)
 - Kristen Heavey [@kristenheavey](#)
 - Beth Chasteen [@officialchasteen](#)
 - Joana Pauline [@thejoannapauline](#)
- Photography by [Dolci Key Photography](#) (that's me)

*I traded edited portfolio images to be delivered to each model after my final presentation for usage of this image data in my data set/presentation. This data set will remain private.