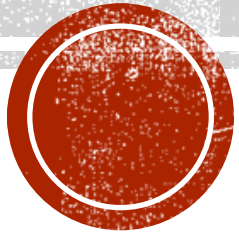
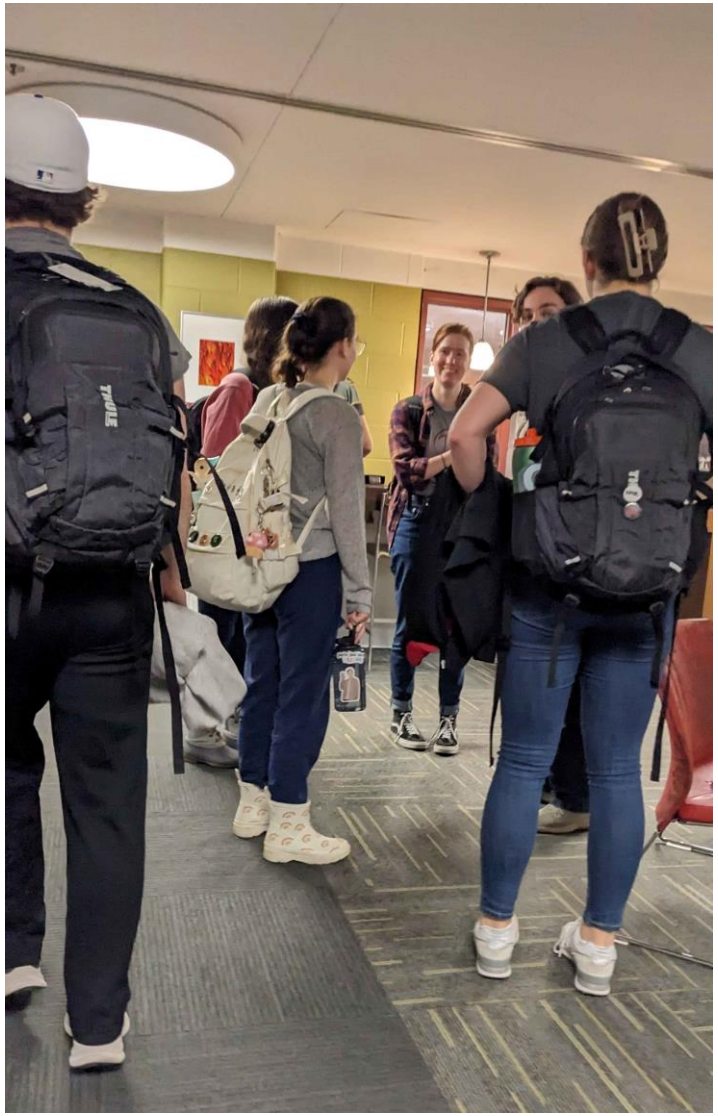
SNIPING PREDICTIONS

Katie Collins, Spencer Halsey, Allyn Loyd, Reilly Mooney



PROBLEM & MOTIVATION

- Large volume of snipes
- Lots of people to tag per picture
- Want to make sniping easier

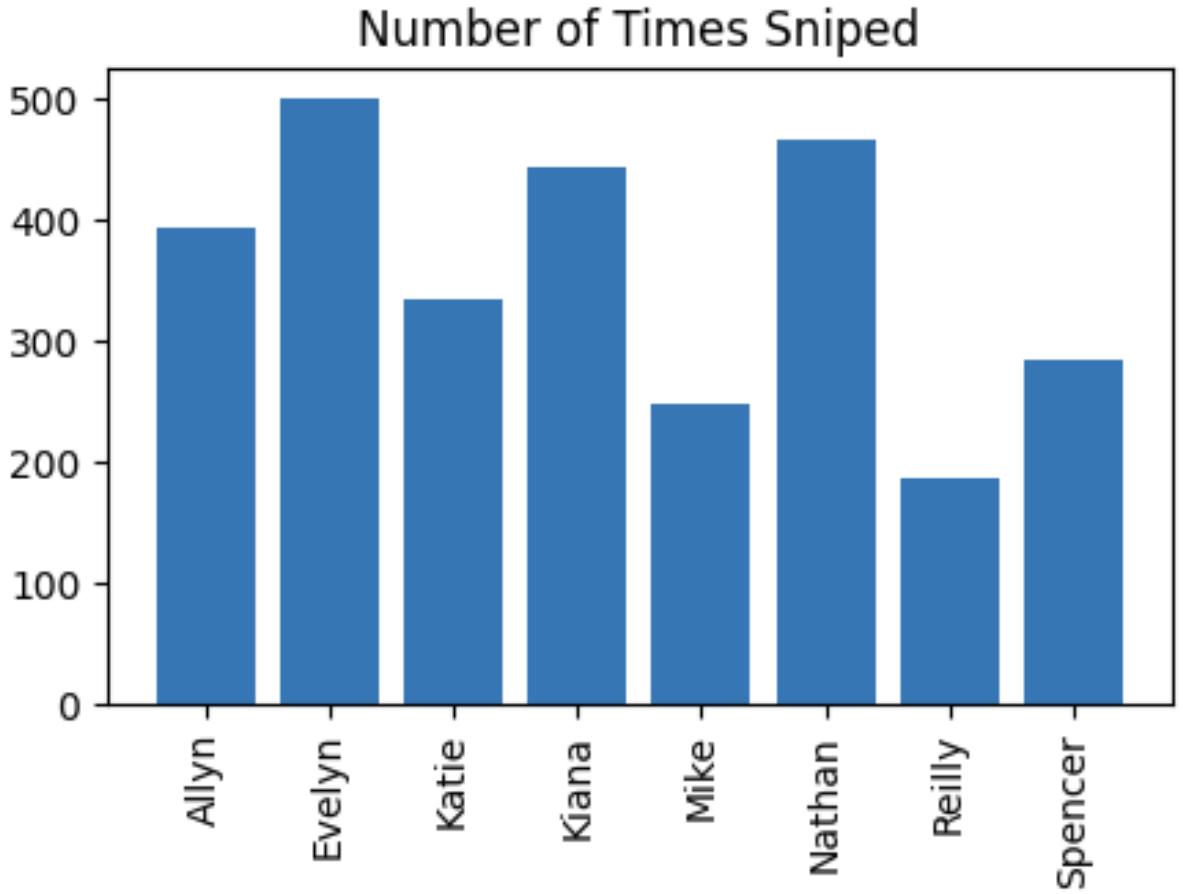


DATASET

- All photos taken by us or our friends over the span of a year
- Shared and labeled in our discord server
- Vary widely in quality

DATASET

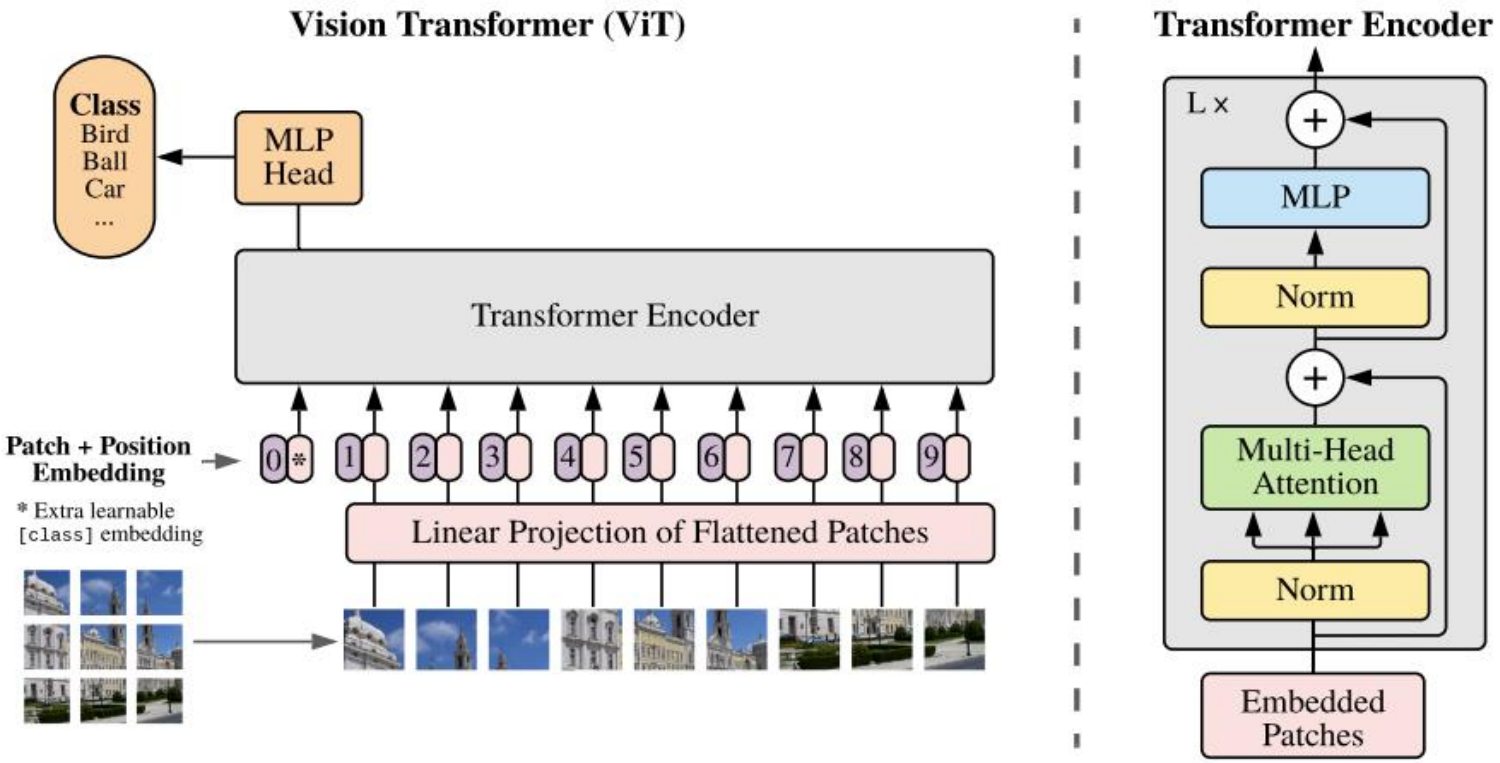
	Raw	Cleaned
# Images	6500+	2848
# Classes	26	8



DATA PIPELINE

- Cropped images
- 256 x 256 RGB
- Metadata
- Data Augmentation

MODEL



- Vision transformers (ViT)
  - Breaks images into patches and processes as sequences

https://arxiv.org/pdf/2010.11929

OTHER THINGS TRIED

Using the metadata

- Increasing the number of params led to slightly worse performance

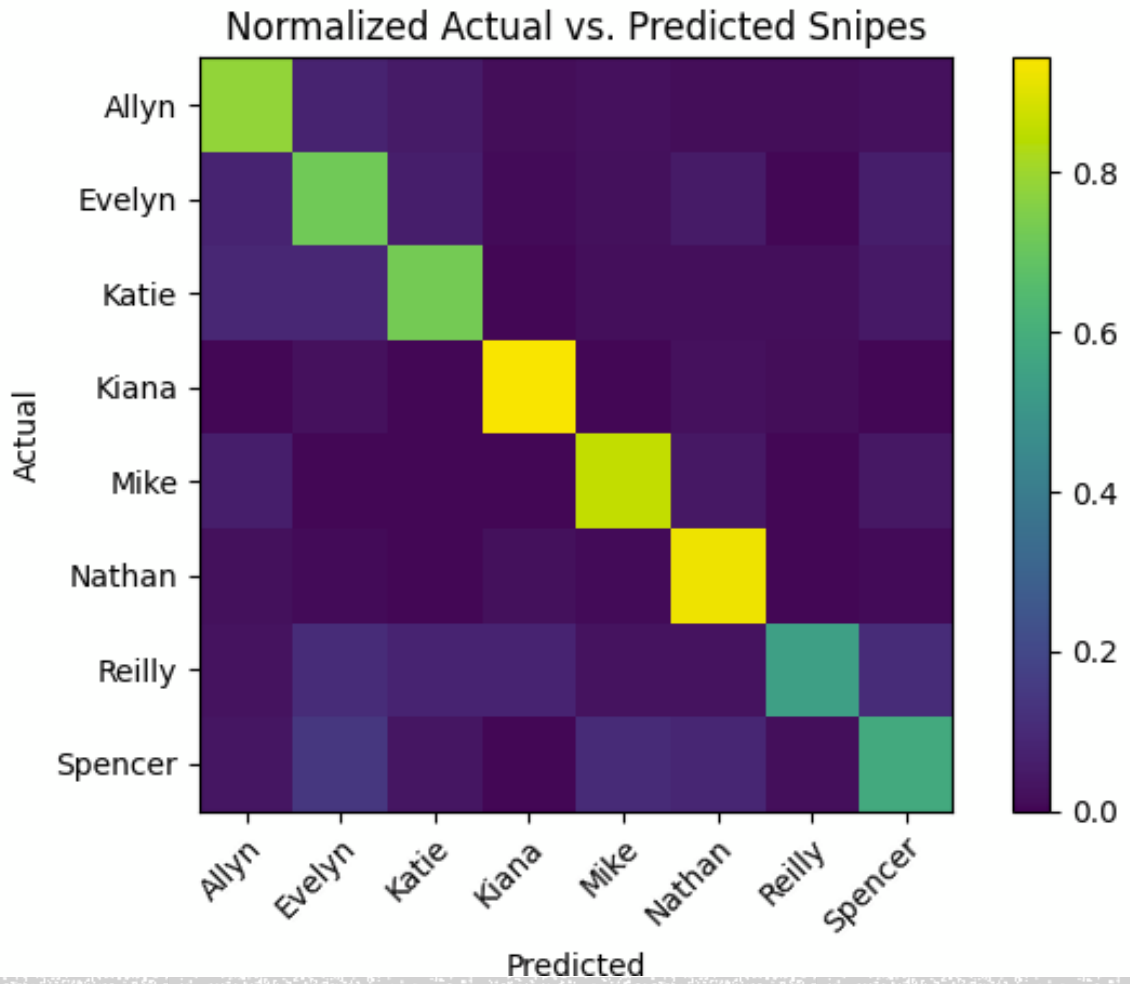
YOLO

ResNets

- Fine-tunes a ResNet-50, ~11% accuracy

VGG16

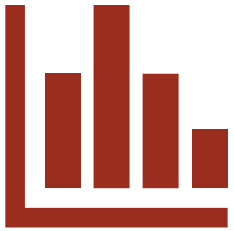
- Tried adding augmentation, attention, and fine-tuning but still only had ~12% accuracy



RESULTS

Baseline accuracy was **8%**
Best accuracy achieved was **78.42%**

STRENGTHS



Performed well for number of data points

Performs well given different environments

Backdrops

Lighting

Angles

- Dataset properties
- Not enough data
- Slow to train

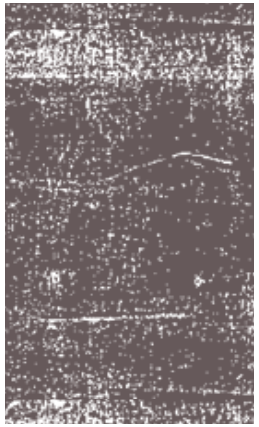
LESSONS LEARNED

Difficult to work with messy data

- Blurry images
- Targets are too far away
- Katie's car is not a trainable images

CNNs vs. Vision Transformers

YOLO vs. Cropping Tool



Literature Review

- Mostly focused on face recognition
- Our data mostly does not show faces clearly
- Had to further research other ideas
  - Visions Transformers
  - Multimodal learning
  - Alternative applications

CNN

- Research behind development of a facial recognition library
- Use CNNs for accurate identification
- Data augmentation to artificially expand the dataset

RELATED WORK

A. Dosovitskiy *et al.*, "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE," Jun. 2021. Available: https://arxiv.org/pdf/2010.11929

D. Shah, "Vision Transformer: What It Is & How It Works [2023 Guide]," *www.v7labs.com*, Dec. 15, 2022. https://www.v7labs.com/blog/vision-transformer-guide

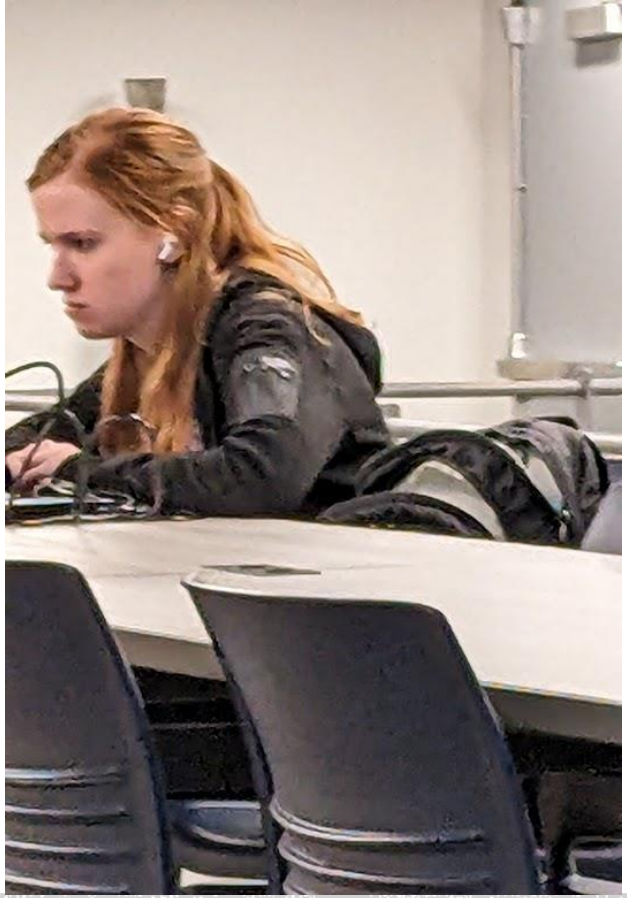
L. Blanger and A. R. Panisson, "A face recognition library using convolutional neural networks," *International Journal of Engineering Research and Science*, vol. 3, no. 8, pp. 84–92, Aug. 2017.25125/engineering-journal-ijoer-aug-2017-25 . doi:10

G. Guo and N. Zhang, "A survey on deep learning based face recognition," Computer Vision and Image Understanding, p. 102805, Aug. 2019, doi: https://doi.org/10.1016/j.cviu.2019.102805.

E. Ahmed, M. Jones and T. K. Marks, "An improved deep learning architecture for person re-identification," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 3908-3916, https://doi.org/10.1109/CVPR.2015.7299016.

W. Wang, J. Yang, J. Xiao, S. Li, and D. Zhou, 'Face Recognition Based on Deep Learning', in Human Centered Computing, 2015, pp.812–820. https://doi.org/10.1007/978-3-319-15554-8_73

Ou, C., et al. "A deep learning based multimodal fusion model for skin lesion diagnosis using smartphone collected clinical images and metadata," in Frontiers in Surgery, vol. 9, 2022.



DEMO

QUESTIONS?