Counting Actor Screen Time in Movies
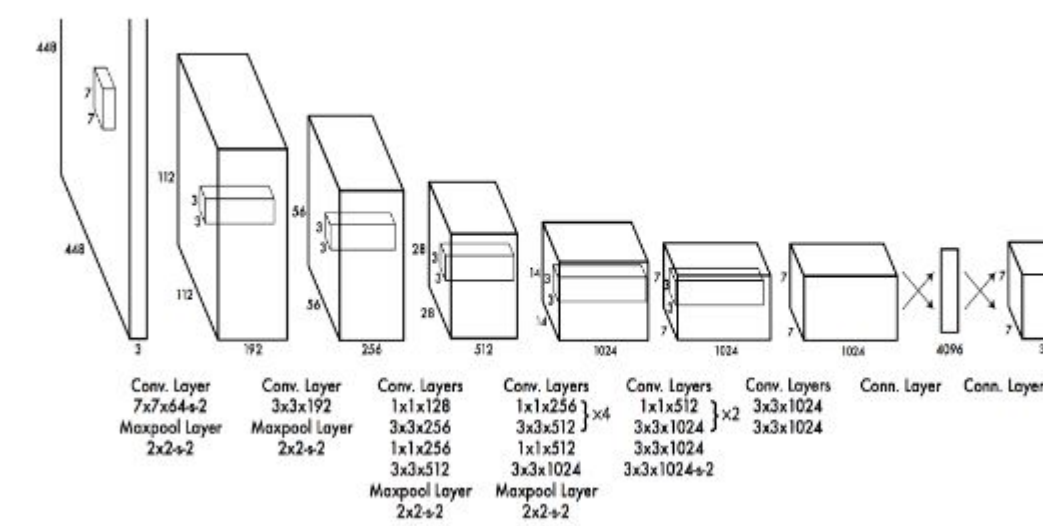
Christina Lin, Stephany Liu, Samuel Wong
chrlin@stanford.edu, liu48185@stanford.edu, samwwong@stanford.edu
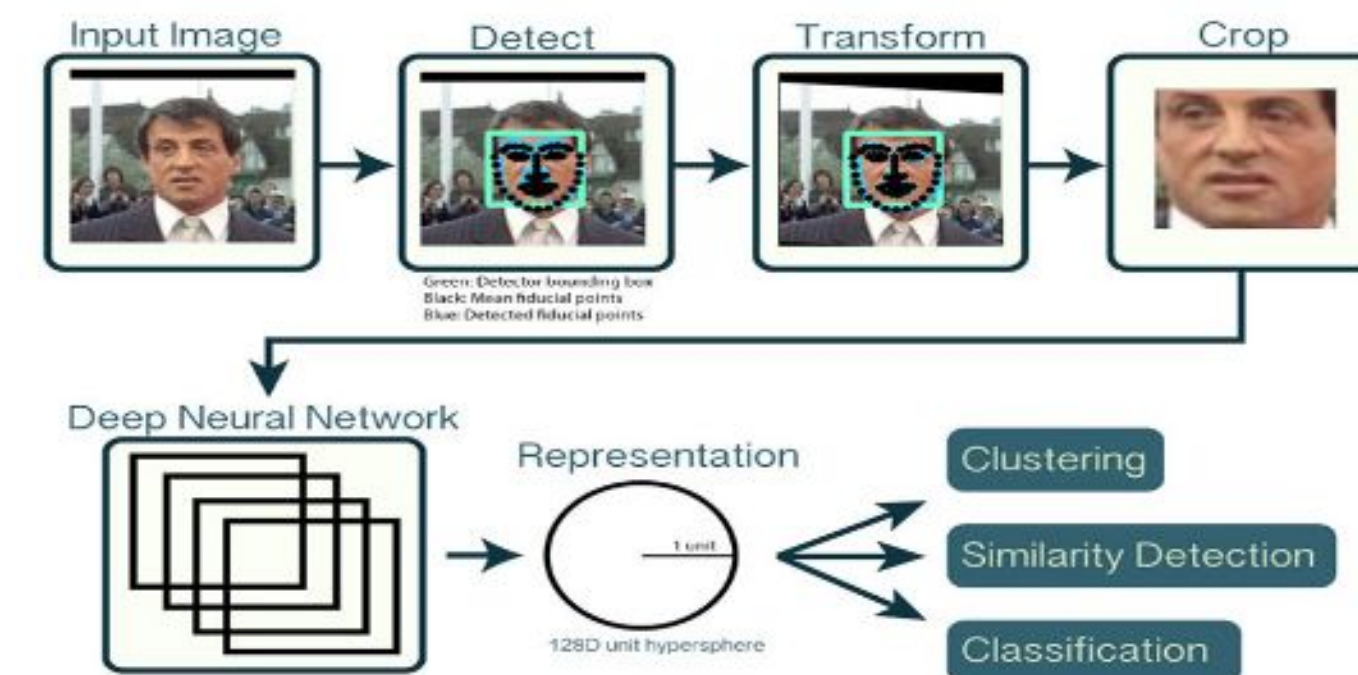
Introduction

- In Avengers: Endgame, the highest grossing blockbuster of all time, which characters were given the most minutes?
- By incorporating deep learning and CNNs to Avengers movie trailers, we can compute the total time on screen
- Use two modules: face detection (bounding box) and face recognition (face labels)

Data

Source
- Avengers movie trailers: Avengers 2012, Age of Ultron, Endgame, Infinity War

Processing
- Alignment
- Augmentation

Split
- 70/30 split (train/test)

Model/Features

YOLO

OpenFace

Results

Table 2: Predicted screen time (in seconds) for Avengers 2012

	Human	OpenFace	MTCNN + FaceNet	YOLO + FaceNet
Screen time - Loki	2	4	5	7
Screen time - Fury	3	0	1	1
Screen time - Hawkeye	2	3	4	3
Screen time - Black Widow	5	3	5	4
Screen time - Iron Man	15	11	17	16
Screen time - Thor	2	3	8	9
Screen time - Captain America	5	6	13	9
Screen time - Hulk	12	4	3	
Screen time - Scarlet Witch	0	0	0	5
Screen time - Quicksilver	0	0	0	1

Table 3: Predicted screen time (in seconds) for Age of Ultron

	Human	OpenFace	MTCNN + FaceNet	YOLO + FaceNet
Screen time - Loki	0	0	0	1
Screen time - Fury	1	0	1	1
Screen time - Hawkeye	3	3	4	3
Screen time - Black Widow	6	7	4	6
Screen time - Iron Man	5	4	8	5
Screen time - Thor	3	3	3	3
Screen time - Captain America	8	3	10	9
Screen time - Hulk	7	6	6	4
Screen time - Scarlet Witch	4	0	0	0
Screen time - Quicksilver	1	0	0	5

Discussion

Future

- Gather more data with other Marvel movies
- Occlusion sensitivity maps
- Techniques for better detection of faces in motion

References

1. Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Open: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.
2. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 779-788, 2016.
3. Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 815-823, 2015.
4. Yuzhou Wang. Who is this person on your screen? Classification of celebrities using CNN models. Technical report, Stanford CS230 Winter 2019, 2019.