DeepRob

Discussion 5
Overview of Final Project Topics II
University of Michigan and University of Minnesota

Self-Supervised Learning

Core List

1. Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks, Lee et al., 2019
2. VICRegL: Self-Supervised Learning of Local Visual Features, Bardes et al., 2022
3. Fully Self-Supervised Class Awareness in Dense Object Descriptors, Hadjivelichkov and Kanoulas, 2022
4. Self-Supervised Geometric Correspondence for Category-Level 6D Object Pose Estimation in the Wild, Zhang et al., 2022

Tactile Perception for Grasping and Manipulation

Core List

1. More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch, Calandra et al., 2018
2. Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering, Bauza et al., 2020
3. Visuotactile Affordances for Cloth Manipulation with Local Control, Sunil et al., 2022
4. ShapeMap 3-D: Efficient shape mapping through dense touch and vision, Suresh et al., 2022

Dataset

Core List

1. Deep Learn
2. Isaac Gym:
3. Grounding
4. All You Nee

Pre-training for Robot Manipulation and Transformer Architectures

Core List

1. SORNet: Spatial Object-Centric Representations for Sequential Manipulation, Yuan et al., 2021
2. CLIPort: What and Where Pathways for Robotic Manipulation, Shridhar et al., 2021
3. Masked Visual Pre-training for Motor Control, Xiao et al., 2022
4. R3M: A Universal Visual Representation for Robot Manipulation, Nair et al., 2022
5. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances, Ahn et al., 2022
6. RT-1: Robotics Transformer for Real-World Control at Scale, Brohan et al., 2022

Grasp Pose Detection

Core List

1. Real-Time Grasp Detection Using Convolutional Neural Networks, Redmon and Angelova, 2015
2. Using Geometry to Detect Grasps in 3D Point Clouds, ten Pas and Platt, 2015
3. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics, Mahler et al., 2017
4. Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes, Sundermeyer et al., 2021
5. Sample Efficient Grasp Learning Using Equivariant Models, Zhu et al., 2022

More Frontiers

Interpreting Deep Learning Models

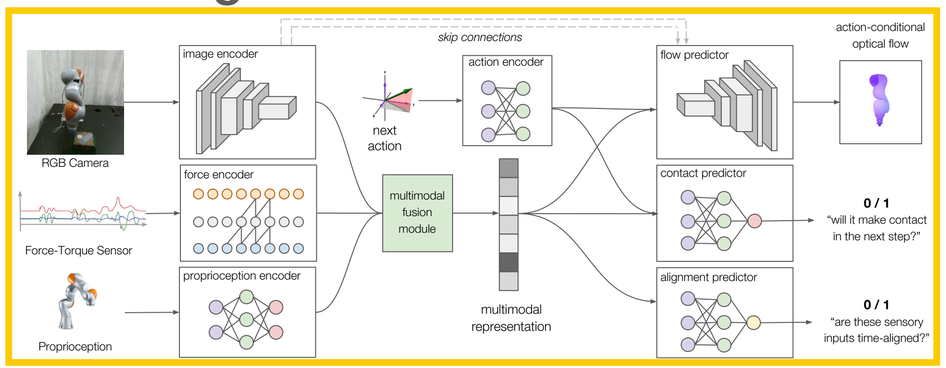
- Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps, Simonyan et al., 2013
- Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, Selvaraju et al., 2016
- The Building Blocks of Interpretability, Olah et al., 2018
- Multimodal Neurons in Artificial Neural Networks, Goh et al., 2021

Fairness and Ethics

- Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, Buolamwini and Gebru, 2018
- Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing, Raji et al., 2020

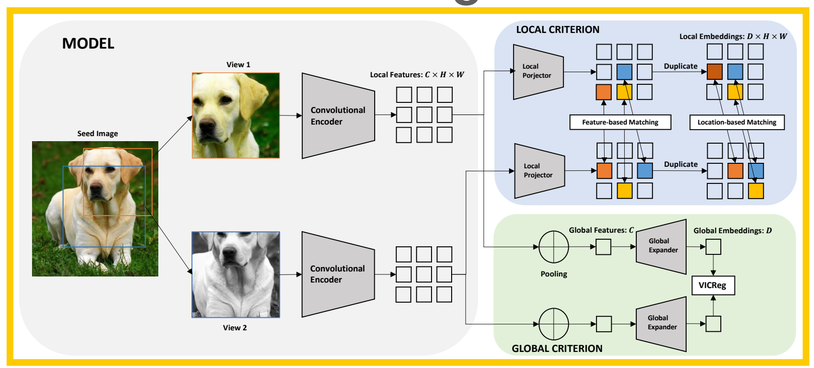
Self-Supervised Learning

Making Sense of Vision and Touch



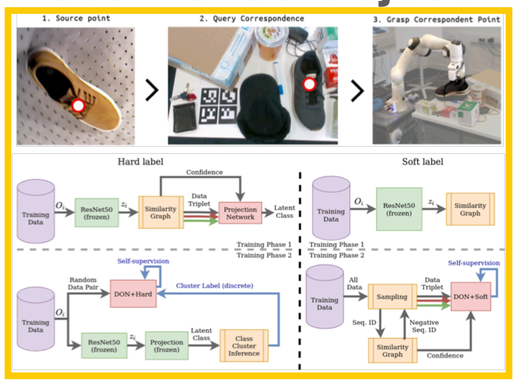
[1] Lee et al., ICRA 2019

VICRegL



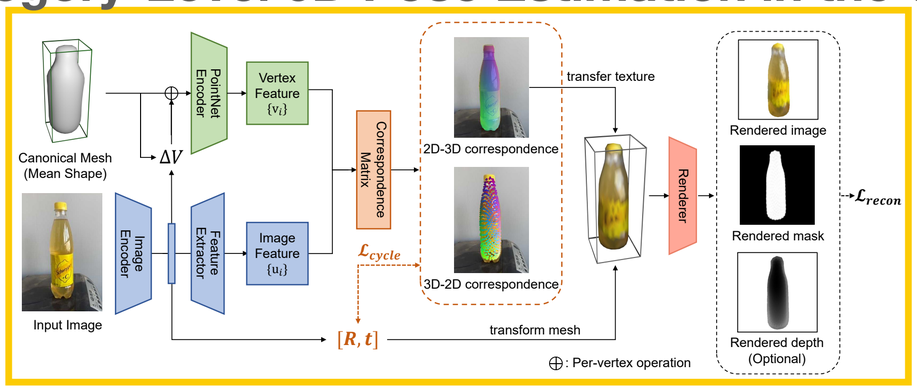
[2] Bardes et al., 2022

Self-Supervised Dense Object Descriptors



[3] Hadjivelichkov and Kanoulas, CoRL 2022

Category-Level 6D Pose Estimation in the Wild



[4] Zhang et al., 2022

[1] Michelle A. Lee, Yuke Zhu, Krishnan Srinivasan, Parth Shah, Silvio Savarese, Li Fei-Fei, Animesh Garg, Jeannette Bohg. "Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks" ICRA, 2019

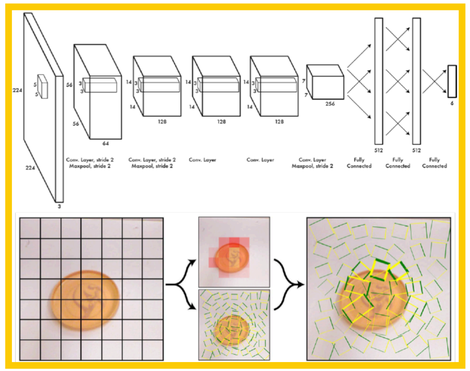
[2] Adrien Bardes, Jean Ponce, Yann LeCun. "VICRegL: Self-Supervised Learning of Local Visual Features" arXiv, 2022.

[3] Denis Hadjivelichkov, Dimitrios Kanoulas. "Fully Self-Supervised Class Awareness in Dense Object Descriptors" CoRL, 2022.

[4] Kaifeng Zhang, Yang Fu, Shubhankar Borse, Hong Cai, Fatih Porikli, Xiaolong Wang. "Self-Supervised Geometric Correspondence for Category-Level 6D Object Pose Estimation in the Wild" arXiv, 2022.

Grasp Pose Detection

Real-Time Grasp Detection

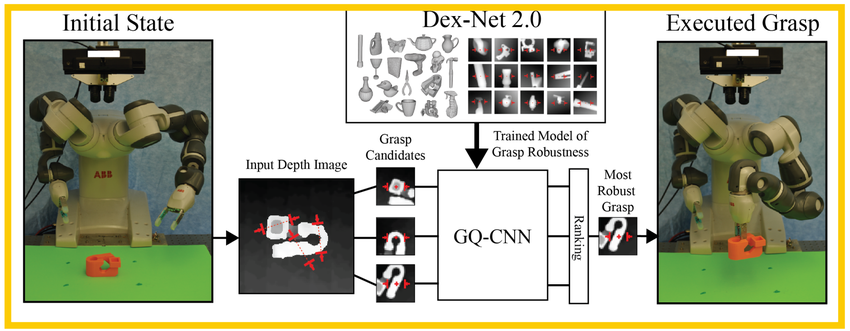


[1] Redmon and Angelova, ICRA 2015

Using Geometry to Detect Grasps

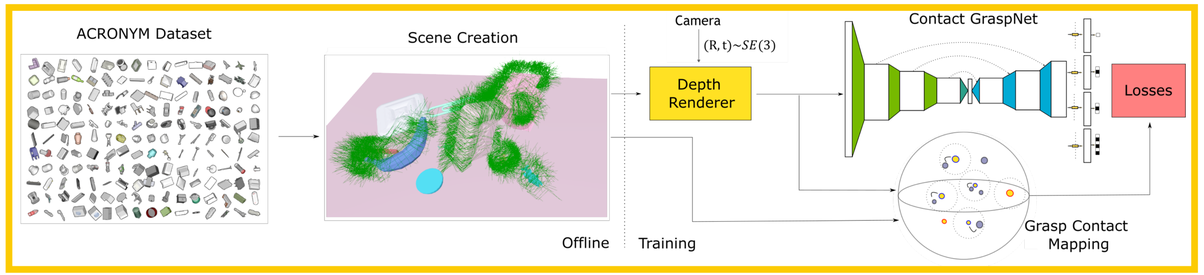
[2] ten Pas and Platt, 2015

Dex-Net 2.0



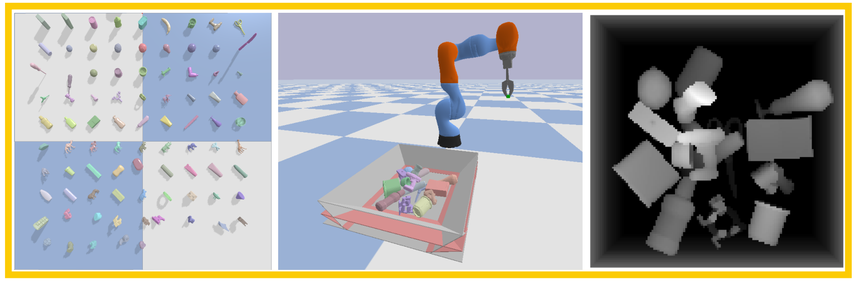
[3] Mahler et al., RSS 2017

Contact-GraspNet



[4] Sundermeyer et al., ICRA 2021

Sample Efficient Grasp Learning

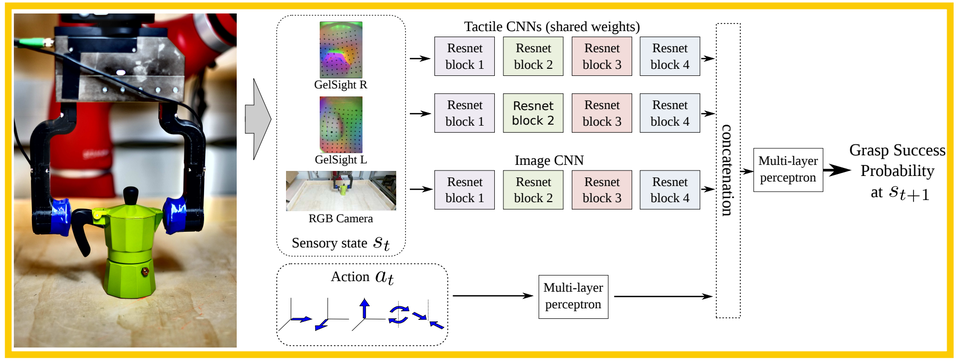


[5] Zhu et al., RSS 2022

[1] Joseph Redmon, Anelia Angelova. “Real-Time Grasp Detection Using Convolutional Neural Networks” ICRA, 2015
[2] Andreas ten Pas, Robert Platt. “Using Geometry to Detect Grasps in 3D Point Clouds” arXiv, 2015.
[3] Jeffrey Mahler, Jacky Liang, Sherdil Niyaz, Michael Laskey, Richard Doan, Xinyu Liu, Juan Aparicio Ojea, Ken Goldberg. “Dex-Net 2.0” RSS, 2017.
[4] Martin Sundermeyer; Arsalan Mousavian; Rudolph Triebel; Dieter Fox. “Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes” ICRA, 2021.
[5] Xupeng Zhu, Dian Wang, Ondrej Biza, Guanang Su, Robin Walters, Robert Platt. “Sample Efficient Grasp Learning Using Equivariant Models” RSS, 2022.

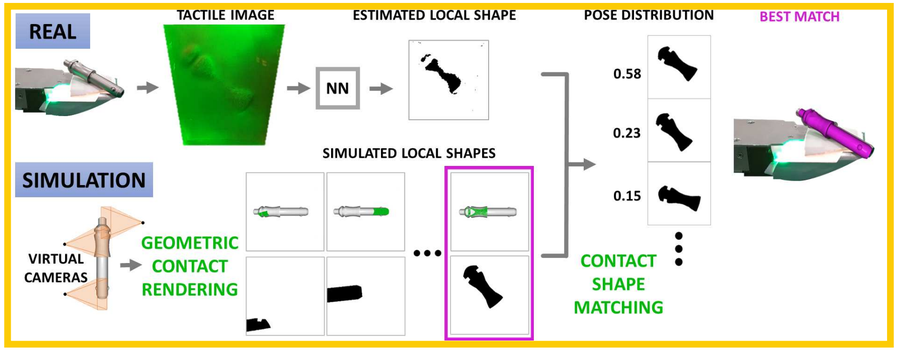
Tactile Perception for Grasping and Manipulation

More Than a Feeling



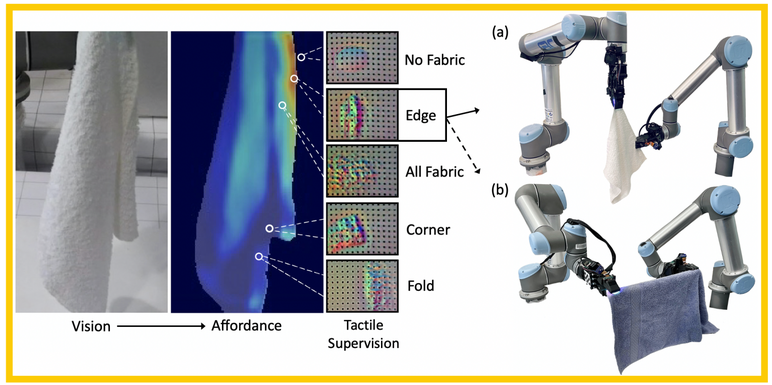
[1] Calandra et al., RAL 2018

Tactile Object Pose Estimation



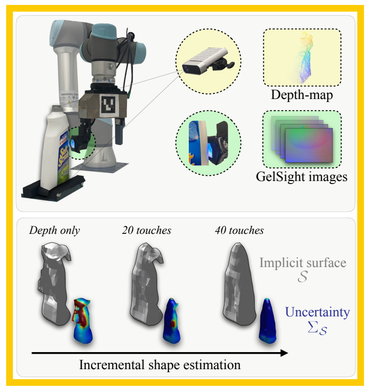
[2] Bauza et al., CoRL 2020

Visuotactile Affordances



[3] Sunil et al., CoRL 2022

ShapeMap 3-D



[4] Suresh et al., ICRA 2022

[1] Roberto Calandra, Andrew Owens, Dinesh Jayaraman, Justin Lin, Wenzhen Yuan, Jitendra Malik, Edward H. Adelson, Sergey Levine. “More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch” RAL 2018

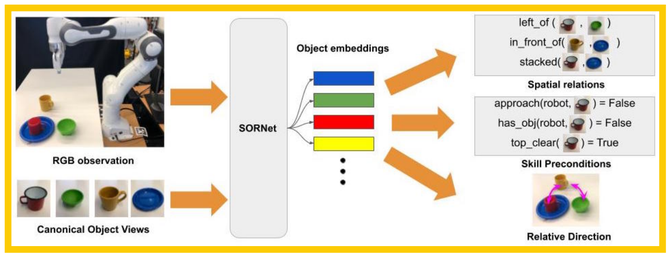
[2] Maria Bauza Villalonga, Eric Valls, Bryan Lim, Theo Sechopoulos, Alberto Rodriguez. “Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering” CoRL, 2020.

[3] Neha Sunil, Shaoxiong Wang, Yu She, Edward Adelson, Alberto Rodriguez. “Visuotactile Affordances for Cloth Manipulation with Local Control” CoRL, 2022.

[4] Sudharshan Suresh, Zilin Si, Joshua G. Mangelson, Wenzhen Yuan, Michael Kaess. “ShapeMap 3-D: Efficient shape mapping through dense touch and vision” ICRA, 2022.

Pre-training for Robot Manipulation and Transformer Architectures

SORNet

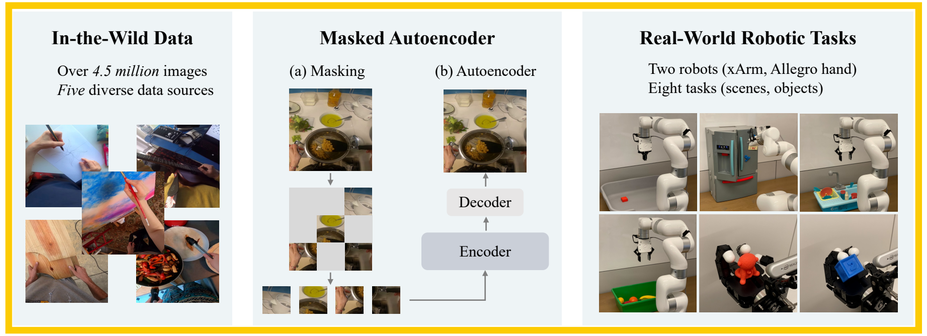


[1] Yuan et al., CoRL 2021

CLIPort

[2] Shridhar et al. CoRL 2021

Masked Visual Pre-training

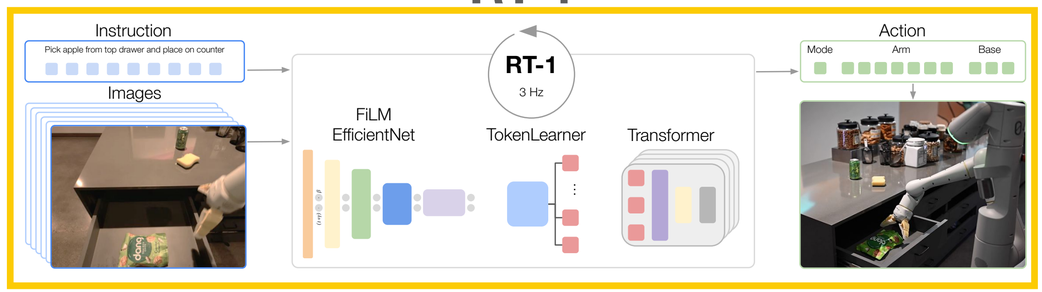


[3] Radosavovic et al., CoRL 2022

SayCan

[4] Ahn et al., 2022

RT-1



[5] Brohan et al., 2022

[1] Wentao Yuan, Chris Paxton, Karthik Desingh, Dieter Fox. “SORNet: Spatial Object-Centric Representations for Sequential Manipulation” CoRL, 2021
[2] Mohit Shridhar, Lucas Manuelli, Dieter Fox. “CLIPort: What and Where Pathways for Robotic Manipulation” CoRL, 2021.
[3] Ilija Radosavovic, Tete Xiao, Stephen James, Pieter Abbeel, Jitendra Malik, Trevor Darrell. “Real-World Robot Learning with Masked Visual Pre-training” CoRL, 2022.
[4] Michael Ahn et al. “Do As I Can, Not As I Say: Grounding Language in Robotic Affordances” arXiv, 2022.
[5] Anthony Brohan et al. “RT-1: Robotics Transformer for Real-World Control at Scale” arXiv, 2022.

Even More!

- Interpretable Models
- Fairness and Ethics
- Articulated and Deformable Objects
- Transparent and Reflective Objects
- Dynamic Scenes

DeepRob

Discussion 5
Overview of Final Project Topics II
University of Michigan and University of Minnesota

Self-Supervised Learning

Core List

1. Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks, Lee et al., 2019
2. VICRegL: Self-Supervised Learning of Local Visual Features, Bardes et al., 2022
3. Fully Self-Supervised Class Awareness in Dense Object Descriptors, Hadjivelichkov and Kanoulas, 2022
4. Self-Supervised Geometric Correspondence for Category-Level 6D Object Pose Estimation in the Wild, Zhang et al., 2022

Tactile Perception for Grasping and Manipulation

Core List

1. More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch, Calandra et al., 2018
2. Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering, Bauza et al., 2020
3. Visuotactile Affordances for Cloth Manipulation with Local Control, Sunil et al., 2022
4. ShapeMap 3-D: Efficient shape mapping through dense touch and vision, Suresh et al., 2022

Dataset

Core List

1. Deep Learn
2. Isaac Gym:
3. Grounding
4. All You Nee

Pre-training for Robot Manipulation and Transformer Architectures

Core List

1. SORNet: Spatial Object-Centric Representations for Sequential Manipulation, Yuan et al., 2021
2. CLIPort: What and Where Pathways for Robotic Manipulation, Shridhar et al., 2021
3. Masked Visual Pre-training for Motor Control, Xiao et al., 2022
4. R3M: A Universal Visual Representation for Robot Manipulation, Nair et al., 2022
5. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances, Ahn et al., 2022
6. RT-1: Robotics Transformer for Real-World Control at Scale, Brohan et al., 2022

Grasp Pose Detection

Core List

1. Real-Time Grasp Detection Using Convolutional Neural Networks, Redmon and Angelova, 2015
2. Using Geometry to Detect Grasps in 3D Point Clouds, ten Pas and Platt, 2015
3. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics, Mahler et al., 2017
4. Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes, Sundermeyer et al., 2021
5. Sample Efficient Grasp Learning Using Equivariant Models, Zhu et al., 2022

More Frontiers

Interpreting Deep Learning Models

- Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps, Simonyan et al., 2013
- Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, Selvaraju et al., 2016
- The Building Blocks of Interpretability, Olah et al., 2018
- Multimodal Neurons in Artificial Neural Networks, Goh et al., 2021

Fairness and Ethics

- Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, Buolamwini and Gebru, 2018
- Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing, Raji et al., 2020