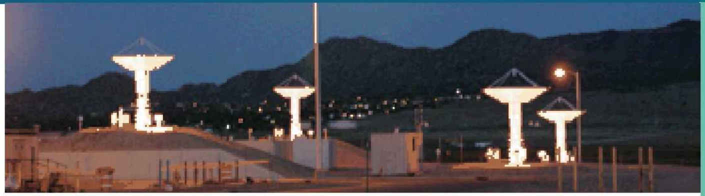
SAND2018-13073PE

Fragment Simulation and Characterization



PRESENTED BY

Emily Donahue

Sandia National Laboratories

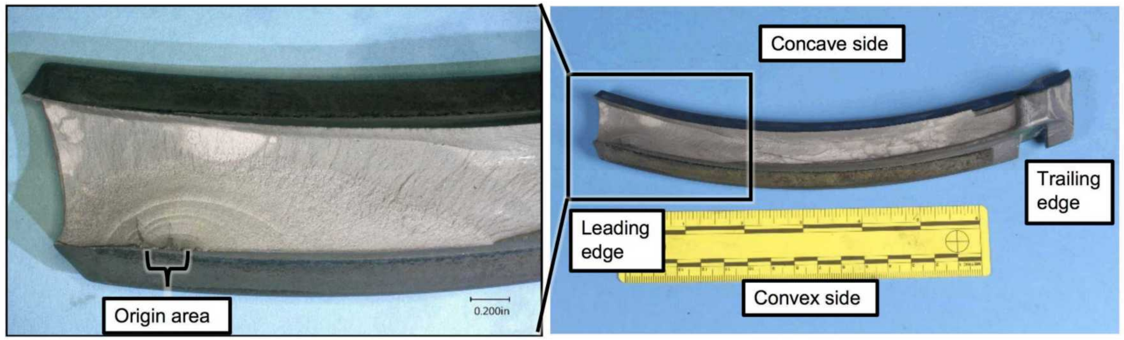
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

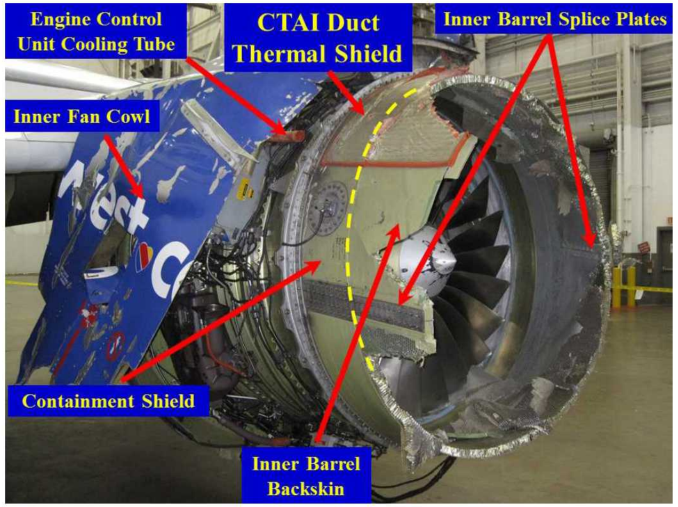
Case Study: Southwest Airlines Engine Explosion

Southwest flight 1380, April 17th, 2018

- Engine failure after takeoff from New York LaGuardia
- Metal fragments from explosion punctured fuselage
- 1 fatality, several injuries

How can we understand fragment flight to prevent future safety incidents?





https://www.nytimes.com/2018/04/17/us/southwest-airlines-explosion.html
https://en.wikipedia.org/wiki/Southwest_Airlines_Flight_1380

Outline

Fragment characterization

- Case Study
- Problem Scope

Current Solutions for Detection & Tracking

- Challenges
- Future ideas from Deep Learning & Computer Vision

Conclusion

- Applications and Impact

Problem Scope

Definition of the problem and experimental setups

Problem Statement

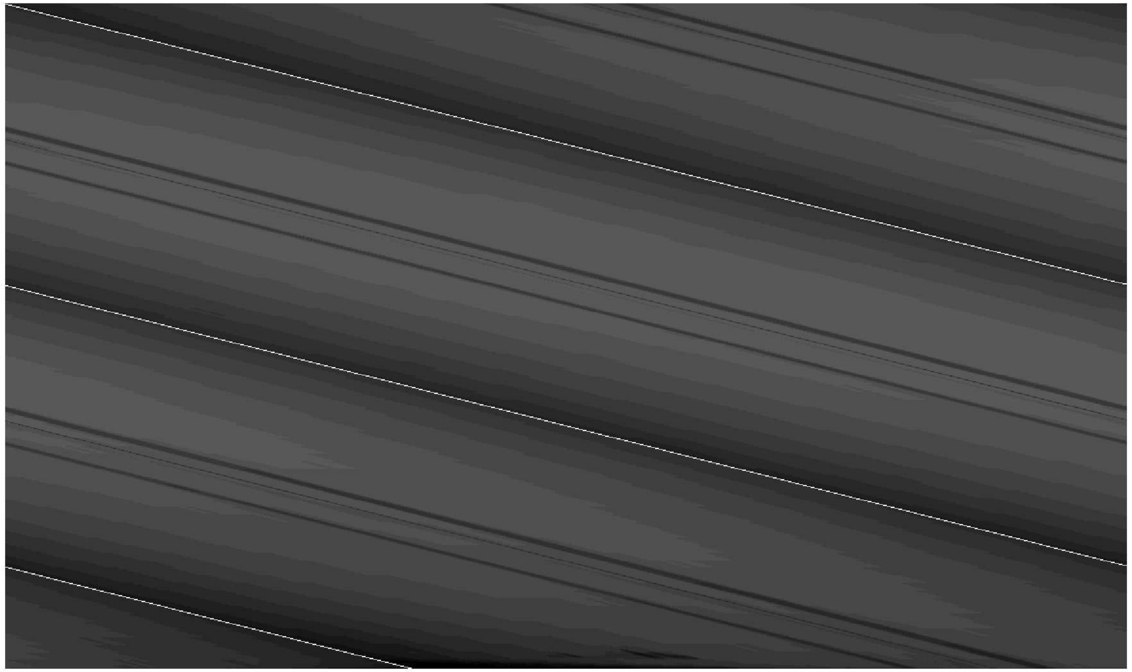
Multi-camera stereo system with synchronized videos

Fragments from experiment fly across field-of-view

Using machine learning & computer vision techniques, can we answer:

- How do fragments form?
- Where are the fragments located in 3d space?
- What are their velocities?
- What are their masses?

Stereo camera system



Fragment experiment

http://www.1vision.co.il/Stereo-3D

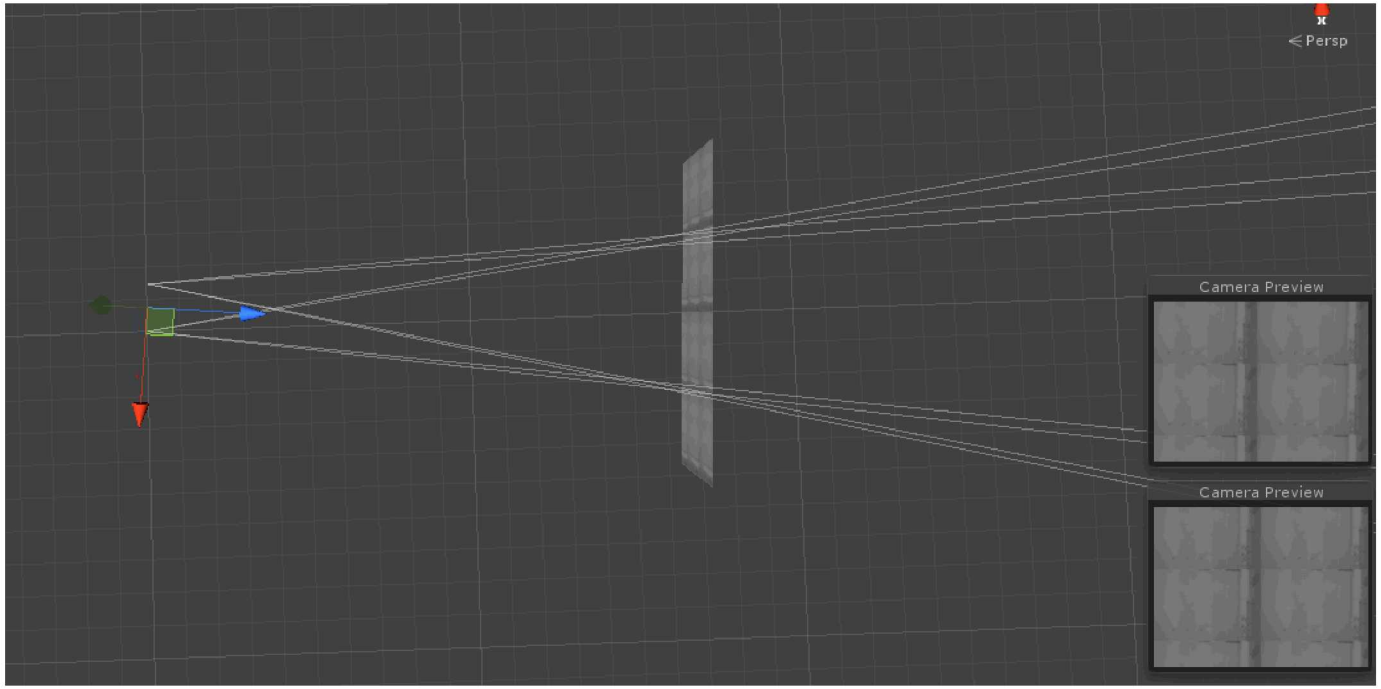
Simulated experiments

Simulate an experimental setup with a stereo camera system

- Verifiable against known ground truth
- Avoid confusers and noisy data

Allow any positioning of cameras, informing experiment design

Can quickly and cheaply simulate a ton of experiments



Bird’s eye view of simulation environment

Current Progress

Detection and tracking of fragments from simulated videos

Fragment characterization method

Simulate fragment experiments

Use deep learning to segment fragment locations

Track fragments

Stereo matching between fragments

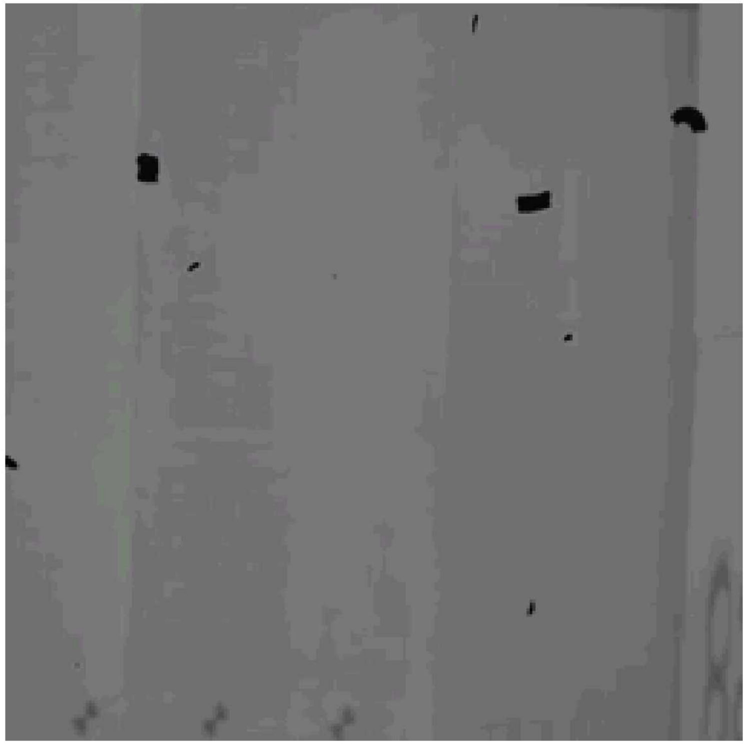
Characterize each fragment
- Positions
- Velocity
- 3D reconstruction

Simulated experiment

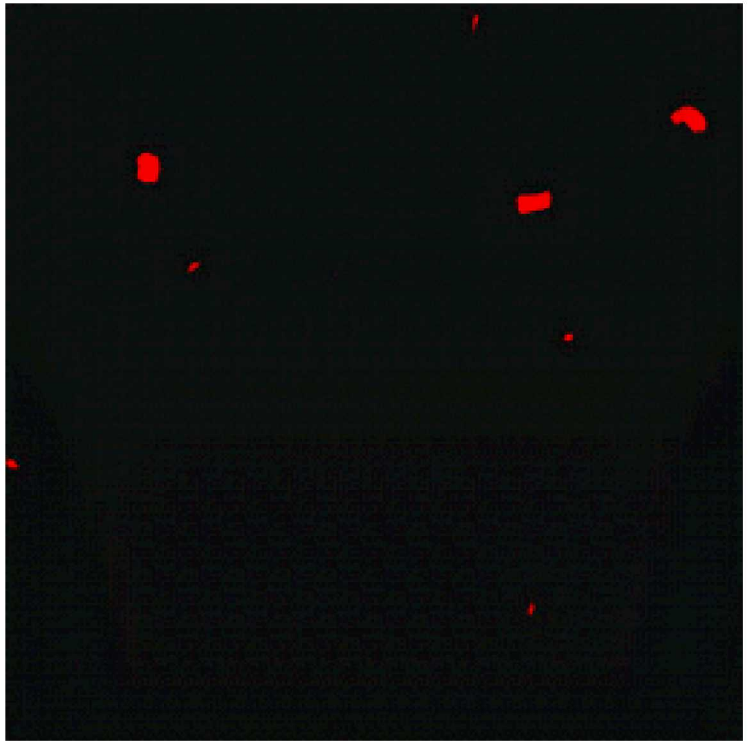
Detection & segmentation

Used pix2pix on simulated video frames

- Have "ground truth" positions of fragments, used that as a transformation target



Simulated Input

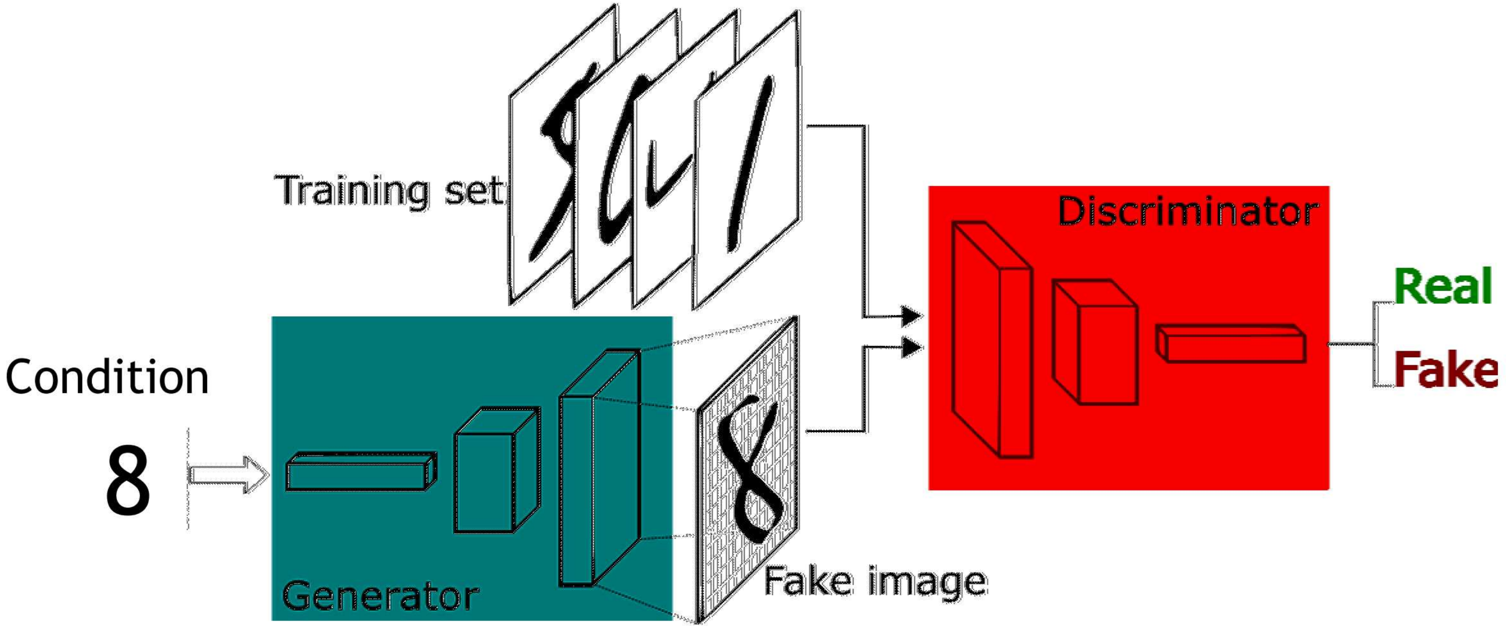


ML segmentation

Generative Adversarial Networks: a game theoretic approach to machine learning

Generative Adversarial Networks (GANs) pit two competing neural networks against each other

- **The generator**, tries to mimic real results
- **The discriminator**, tries to identify mimicked results from real results



Pix2Pix and Generative Adversarial Networks

Pix2pix model

- Condition is an image instead of a label
  - E.g. color segmentation of a scene
- GAN has to learn how to fill in segmentations convincingly
- Training goal is to fool the discriminator

Input labels

Synthesized image

Wang, et. Al. “High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANS”. CVPR 2018

Algorithmic tracking

Idea: frame the tracking problem as a graph optimization problem

- Nodes are detection locations for all frames from pix2pix
- Edges are all possible tracks from fragments in one frame to fragments in next frame

3-way matching algorithm

- Given edges between 3 successive frames
- Find the edges that make physical sense:
  - Approximately constant velocity
  - Fragment shapes are similar

Fragment correlation

Tracking

CSRT algorithm:

- User specifies first bounding box (or automatic segmentation algorithm)
- Algorithm learns a filter from affine transformations of the image patch
- Filter correlated with next frame, max response is the detected location

Future research: improve tracking by incorporating 3-way matching

CSRT algorithm applied to simulated video

Challenges

Stereo Tradeoff

Stereo cameras that are close together (i.e. small baseline) have greater depth error

http://www.coolopticalillusions.com/chalk-art/globe-sidewalk-chalk-paintings.htm

Stereo Tradeoff

…but stereo cameras with a wide baseline have ambiguous matches

https://hiveminer.com/Tags/stereo%2Cwiggled
https://www.nycgo.com/museums-galleries/statue-of-liberty-national-monument

Stereo Reconstruction Tradeoff

A 2-camera (or multi-camera) system can mimic how the eyes perceive depth

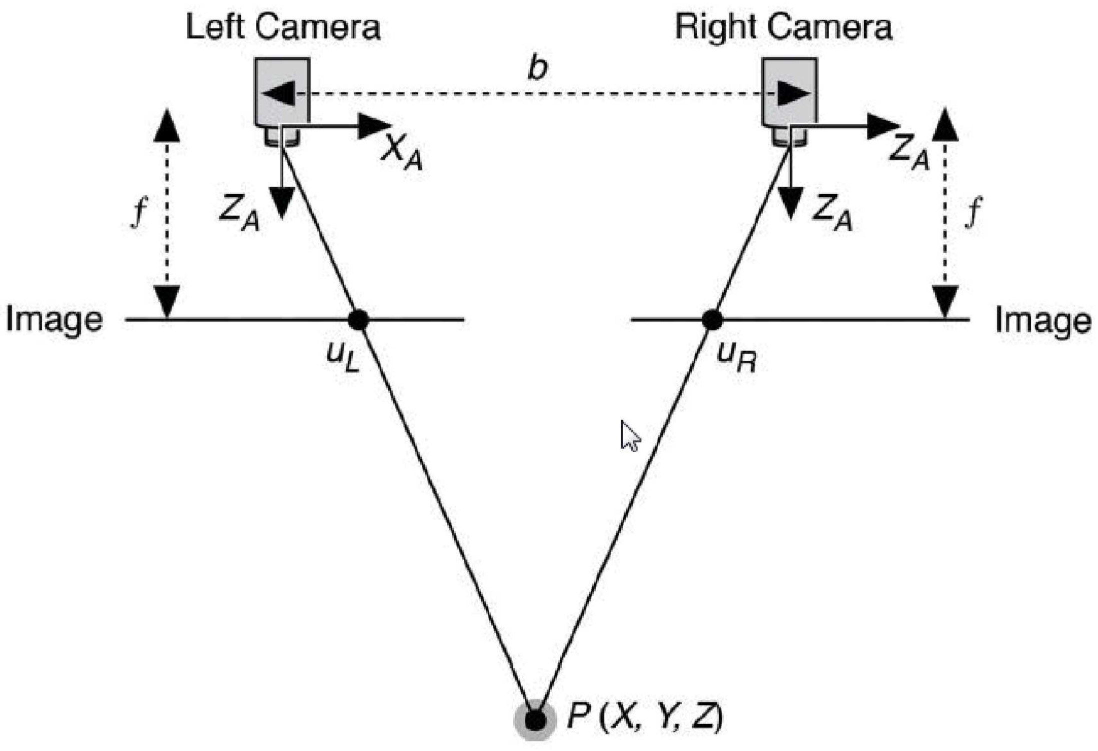
Can find distance to the matching point with triangulation

Stereo angle tradeoff:

- More distance between the cameras reduces depth error
- …but increases difficulty in matching image points

Can add more cameras, but this can become expensive

Requires careful design of number of cameras and their positions



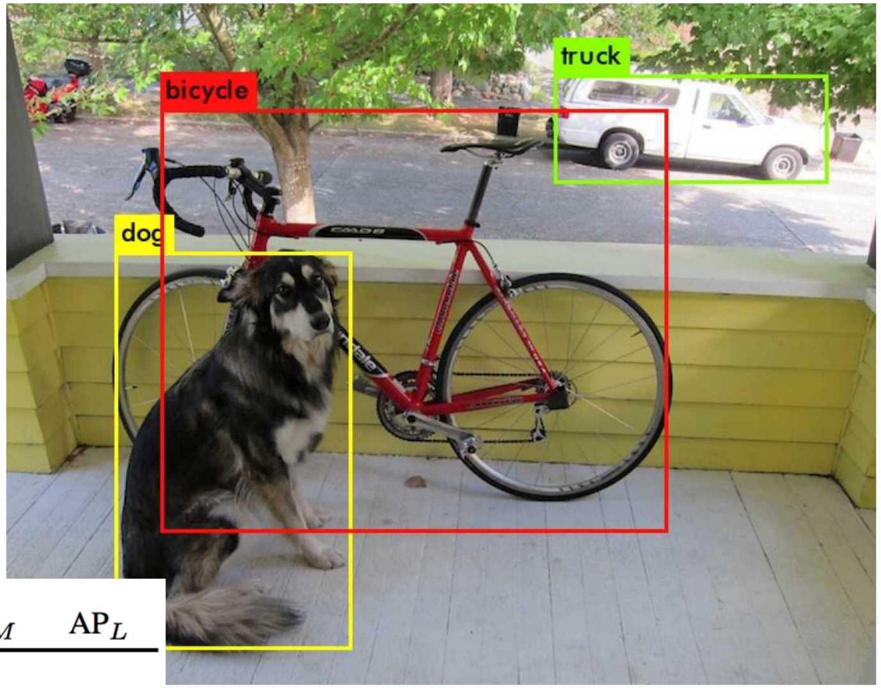
http://www.ni.com/white-paper/14103/en/

Conventional object detection

Recent computer vision techniques using deep neural networks have shown impressive results on object detection & tracking in pictures and videos

Challenge: datasets used for training these networks use objects from many different categories with labels and hand-drawn bounding boxes

- i.e. Supervised methods



Results from YOLO v3 network

	backbone	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Two-stage methods							
Faster R-CNN+++ [5]	ResNet-101-C4	34.9	55.7	37.4	15.6	38.7	50.9
Faster R-CNN w FPN [8]	ResNet-101-FPN	36.2	59.1	39.0	18.2	39.0	48.2
Faster R-CNN by G-RMI [6]	Inception-ResNet-v2 [21]	34.7	55.5	36.7	13.5	38.1	52.0
Faster R-CNN w TDM [20]	Inception-ResNet-v2-TDM	36.8	57.7	39.2	16.2	39.8	**52.1**
One-stage methods							
YOLOv2 [15]	DarkNet-19 [15]	21.6	44.0	19.2	5.0	22.4	35.5
SSD513 [11, 3]	ResNet-101-SSD	31.2	50.4	33.3	10.2	34.5	49.8
DSSD513 [3]	ResNet-101-DSSD	33.2	53.3	35.2	13.0	35.4	51.1
RetinaNet [9]	ResNet-101-FPN	39.1	59.1	42.3	21.8	42.7	50.2
RetinaNet [9]	ResNeXt-101-FPN	**40.8**	**61.1**	**44.1**	**24.1**	**44.2**	51.2
YOLOv3 608 × 608	Darknet-53	33.0	57.9	34.4	18.3	35.4	41.9

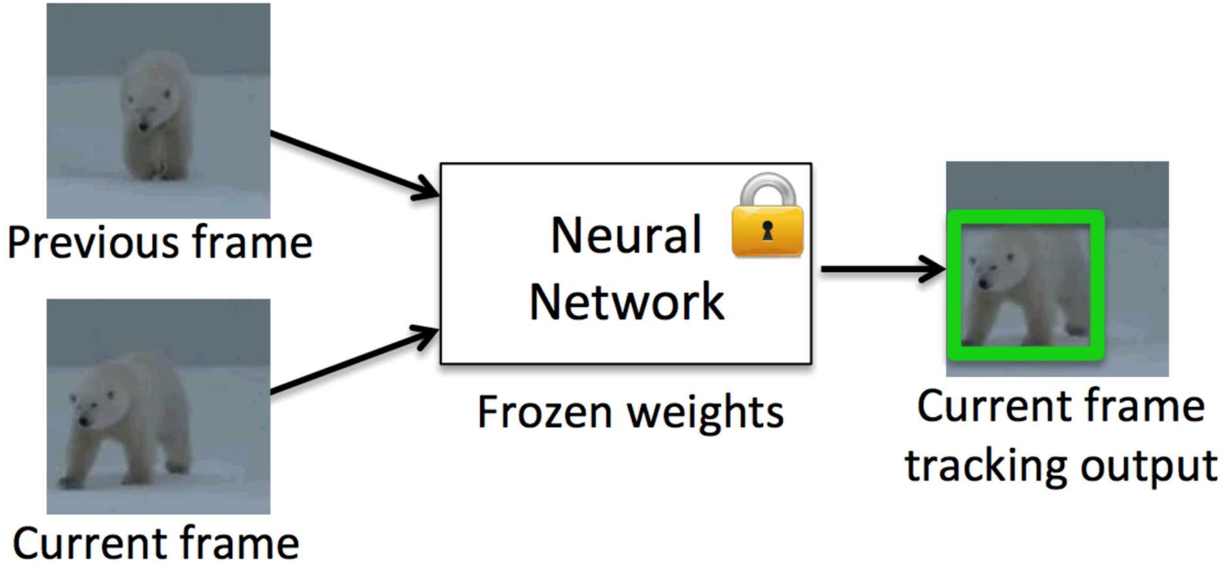
https://pjreddie.com/media/files/papers/YOLOv3.pdf

Deep Learning-based Tracking

Deep learning has made huge strides in many computer vision challenges, including object tracking

e.g. GOTURN

- Use a pre-trained CNN on videos
- Learns how to track an object in a video given a single image crop of the object
- Downside: Supervised training
- Downside :Trained on conventional color videos



D Held, et al. "Learning to track at 100 FPS with Deep Regression Networks. ECCV 2016

Future directions

State-of-the-art techniques for solving fragment characterization

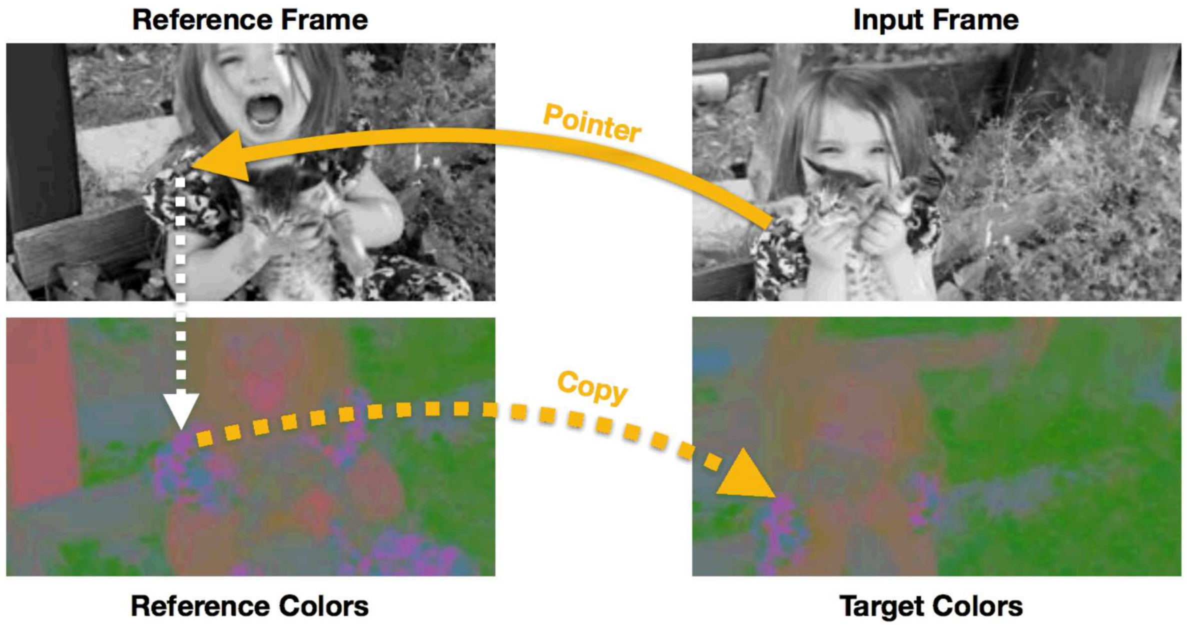
Object tracking via color

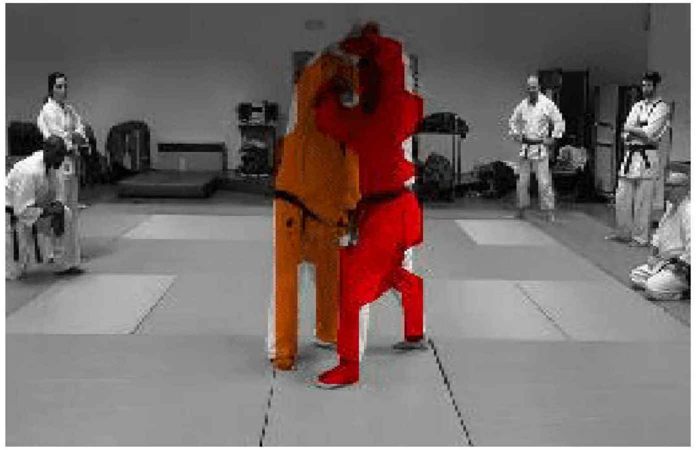
Vondrick, et. al. "Tracking emerges by Colorizing Videos", Google Research

Train a CNN to colorize grayscale images from video frames

Since the trained CNN can color an object, it can also track that object via color between sequential video frames

Unsupervised technique



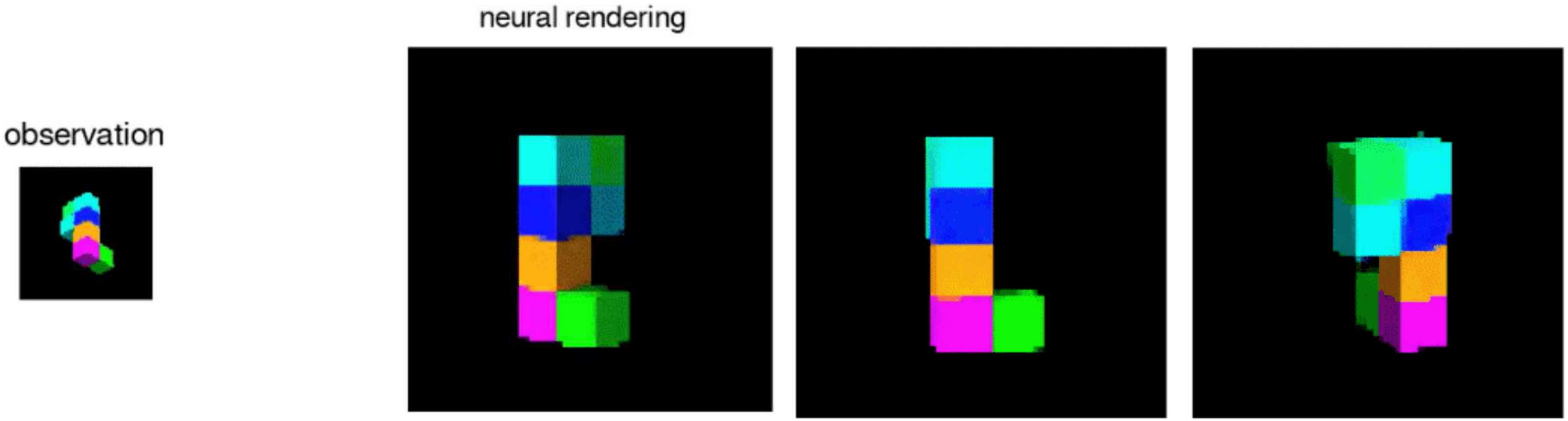


Object reconstruction

DeepMind: Neural scene representation and rendering

Generative Query Network (GQN):

- Learns to perceive its “world” by training only on observed data during exploration
- Makes 3-dimensional sense of objects with many 2-dimensional examples

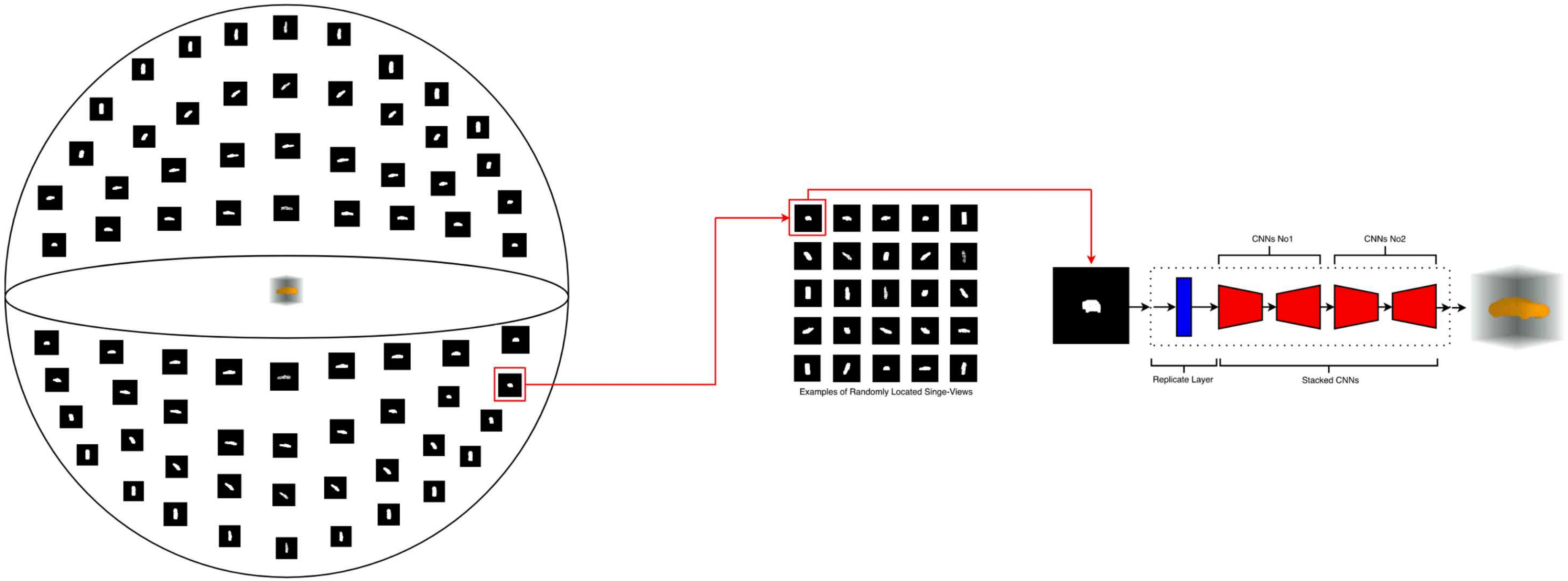


Object reconstruction from silhouettes

Di, Yu 2016: 3D Reconstruction of Simple Objects from A Single View Silhouette Image

Idea:

- Use 3D models of shapes to generate many silhouette images
- Train a CNN to generate a 3D model by training on 2D silhouettes



Conclusion

Applications and Impact

National security is Sandia's business

- Take on high-risk (i.e. too expensive for industry) projects with a focus on improving safety and security

Understanding fragment flight could reduce risk for many applications

- Airplane engines
- Industrial environments

Thank you!