AVSS 2021
NOV 16-19, VIRTUAL

FlagDetSeg: Multi-Nation Flag Detection and Segmentation in the Wild

Shou-Fang Wu[1], Ming-Ching Chang[2], Cheng-Shih Wong[3], Abhineet Kumar Pandey[2], Siwei Lyu[4], Po-Chi Su[1]

[1] National Central University, Taiwan
[2] University at Albany, State University of New York, USA
[3] Academic Sinica, Taiwan
[4] University at Buffalo, State University of New York, USA

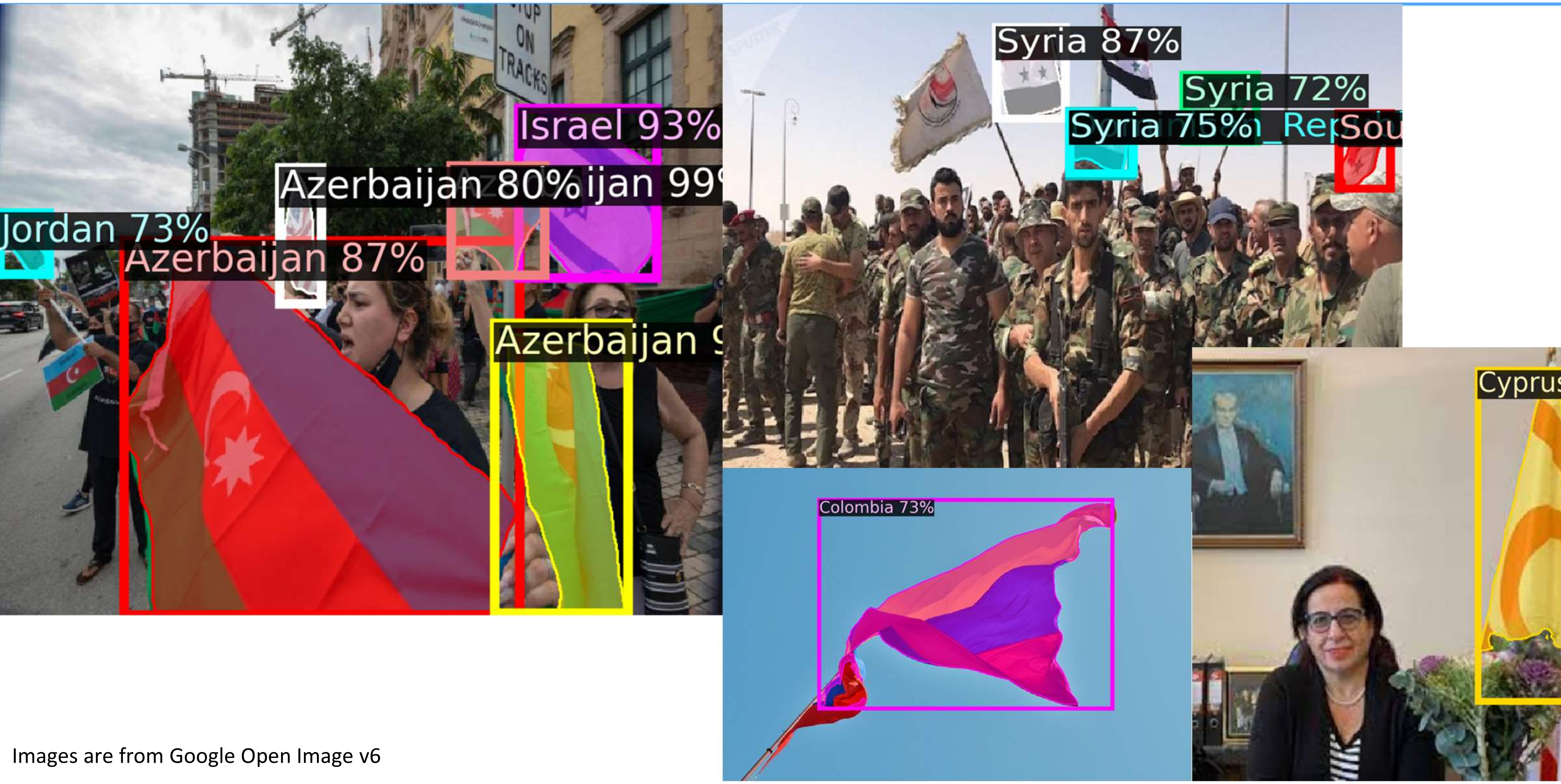
Outline

- Overview
- Prior Work
- Method
- Experiment
- Discussion

Overview

- Precise **instance segmentation for 200+ country flags**
- **Data-augmentation-based** methods for fine-tuning
- Experiments performed on several popular detectors
- *RELEASE* *(https://github.com/sfstefanwu/FlagDetSeg.git)*
  - Pre-trained multi-nation flag detector
  - Annotated multi-nation flag dataset (authentic images)
  - Synthetic multi-nation flag dataset

Overview - Quick Demo



Images are from Google Open Image v6

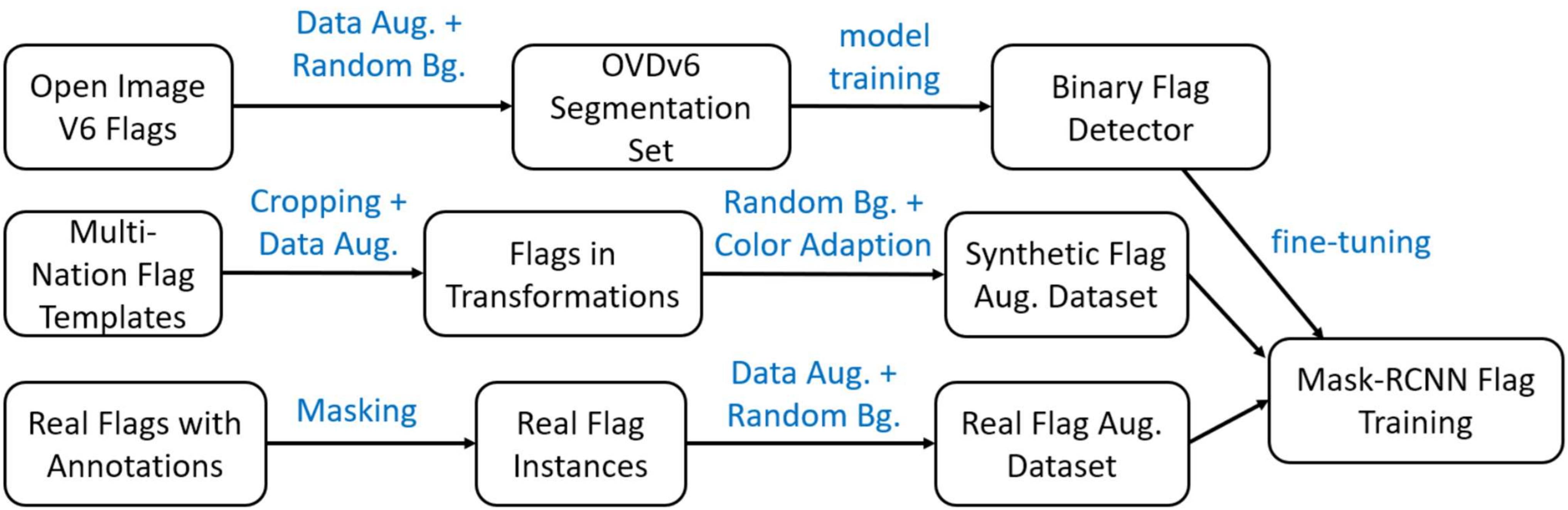
Overview - Challenge

- **Non-rigid** - tiled, rotated, elastified, etc.
- **Heavy occlusion** in many cases
- **Lack of data** - labor-intensive for production
    - 193 member states in UN

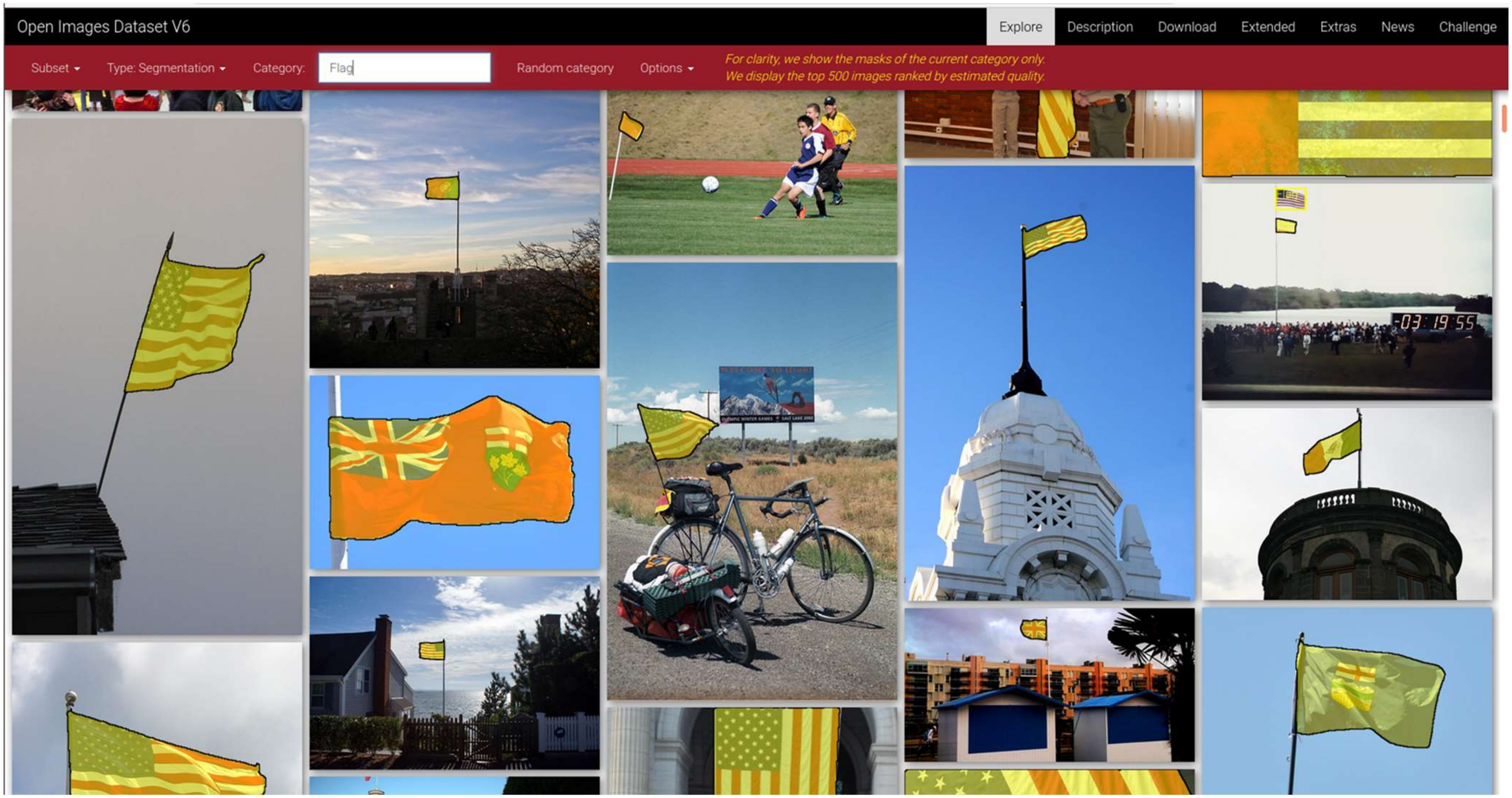
Prior Work

- Binary Flag Detection
  - **HSV color texture analysis** and **gradient features** (S. Jetley, et al.)
  - **Color features** and a **fuzzy-neural algorithm** with **kNN classifier** (E. Hart, et al.)
  - A 5-layer **CNN** but limited results (H. H. Duc, et al.)
- Multi-class Flag Detection
  - **Deep CNN**, yet is uncompetitive against RPN-based detectors (M. Gu, et al.)
  - Based on VGG16 FCN, local context network and Color-BRIEF features (T. Said, et al.)

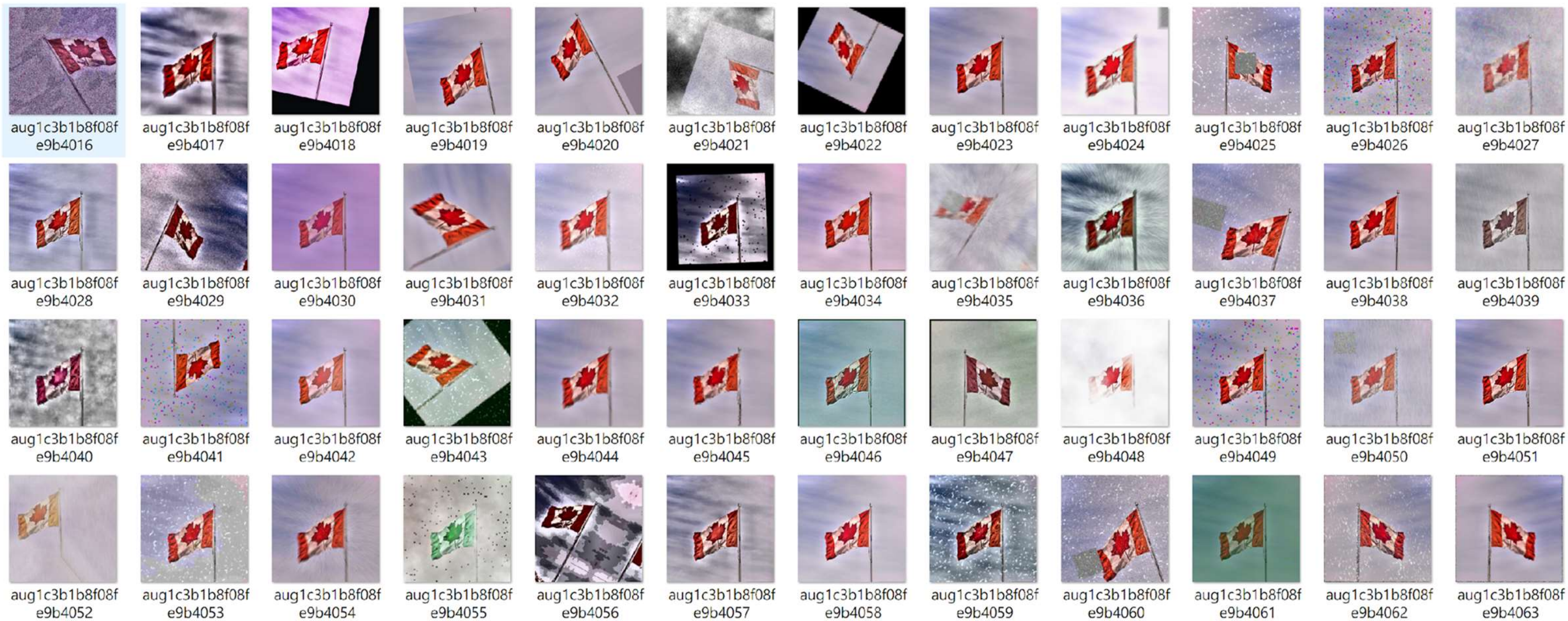
Method - Overview



Flags from Open Image Dataset v6



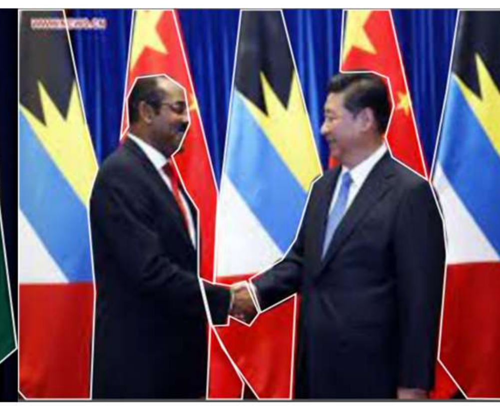
OIDv6 Segmentation Dataset



Real Image Set

- Collect **5** real images for each country
- Select flags with different poses to **maximize diversity**



Synthetic Image Set

- **Natural Backgrounds + (Template or Instance)**
- Generate **large** and **balanced** dataset with **ground truth**
- *"Simple copy-paste is a strong data augmentation method for instance segmentation"* (G. Ghiasi, et al., CVPR 2021)
- Source of background images
  - Human Made Scene Collection (Burge J., et al.)
  - Stanford Background Dataset (S. Gould, et al.)
  - In-house Collection

Samples of Background Image

Template in SVG

Samples of Transformed Instance

- **Safe Transformations** are applied to
  - **(a) Templates in SVG format**
  - **(b) Segmented instances from our *Real Image Set***

to hallucinate realistic-looking flag images

(a)

(b)

Unsafe Transformation

- Vertical flipping

Indonesia

Poland

- RGB channel shifting or aggressive hue change

Romania

Ireland

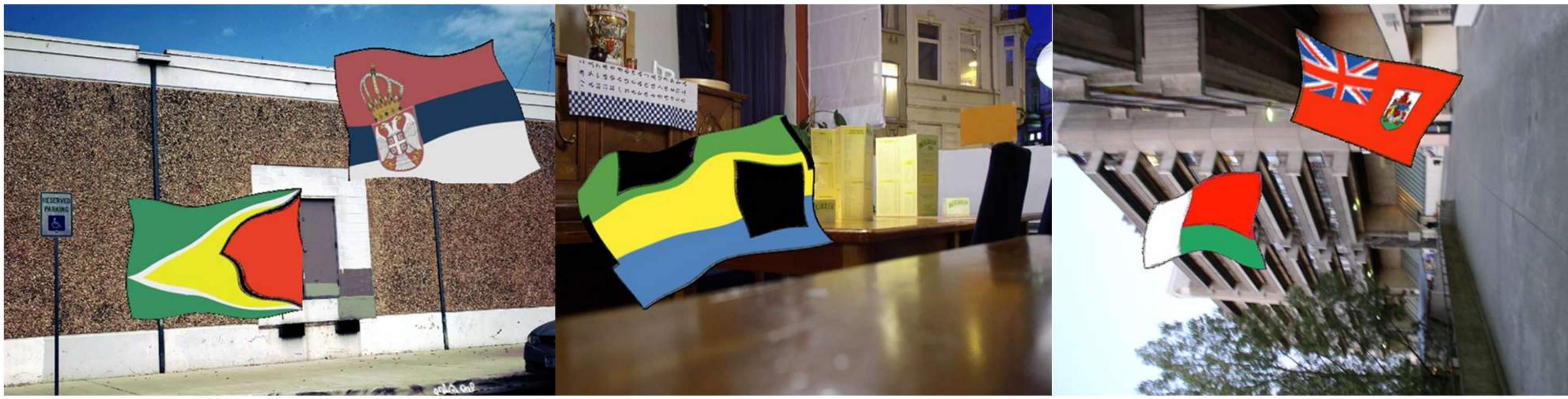
Italy

Nigeria

Mali

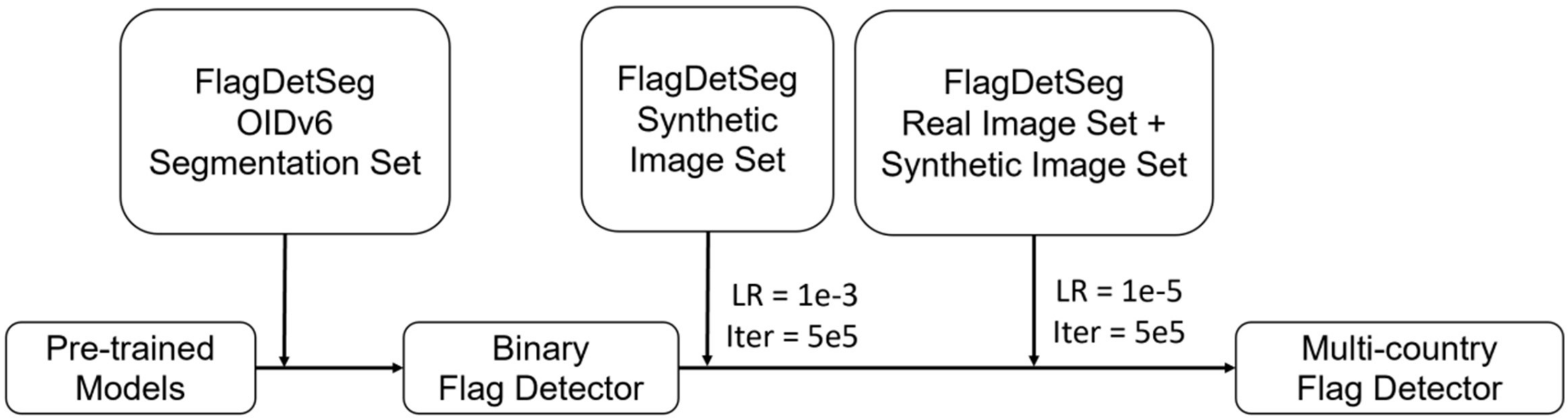
Samples of Synthetic Image Set

(a) template



(b) real

Training Pipeline



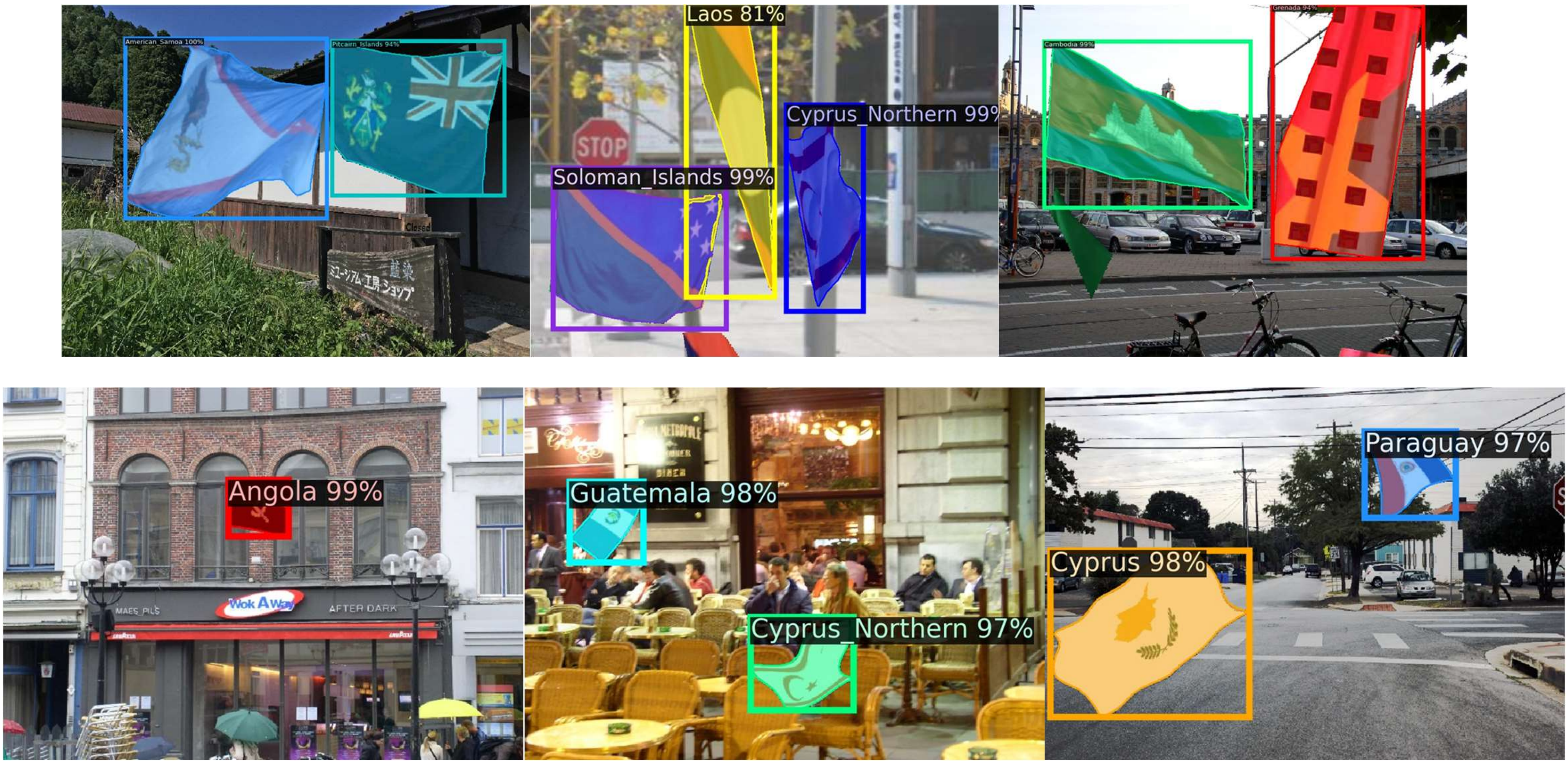
Visual Result - Binary

Experiment

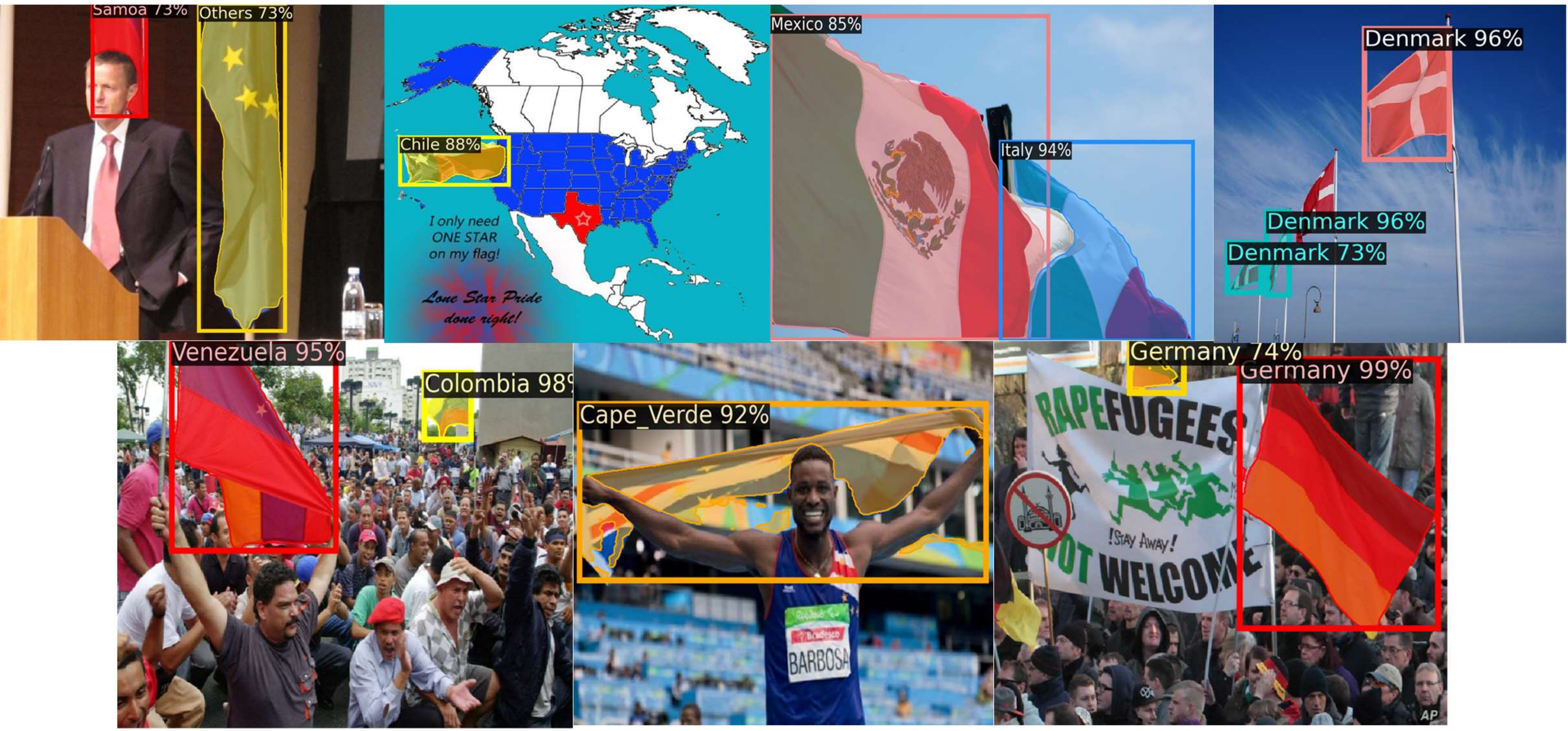
- **Multi-nation Flag Detector**

	Backbone	AP	AP_{50}	AP_{75}	AP_s	AP_m	AP_l
Mask-RCNN	ResNet-101-FPN	**87.92**	**93.1**	**92.36**	40.8	80.03	**91.03**
Mask-RCNN	ResNeXt-101-FPN	85.81	90.75	90.11	**44.71**	**81.58**	88.84
PointRend	ResNet-101-FPN	83.91	87.68	87.13	31.49	77.82	88.04
PointRend	ResNeXt-101-FPN	82.05	85.23	84.94	32.96	79.22	85.34

Visual Result of Synthetic Images



Visual Result of Real Images

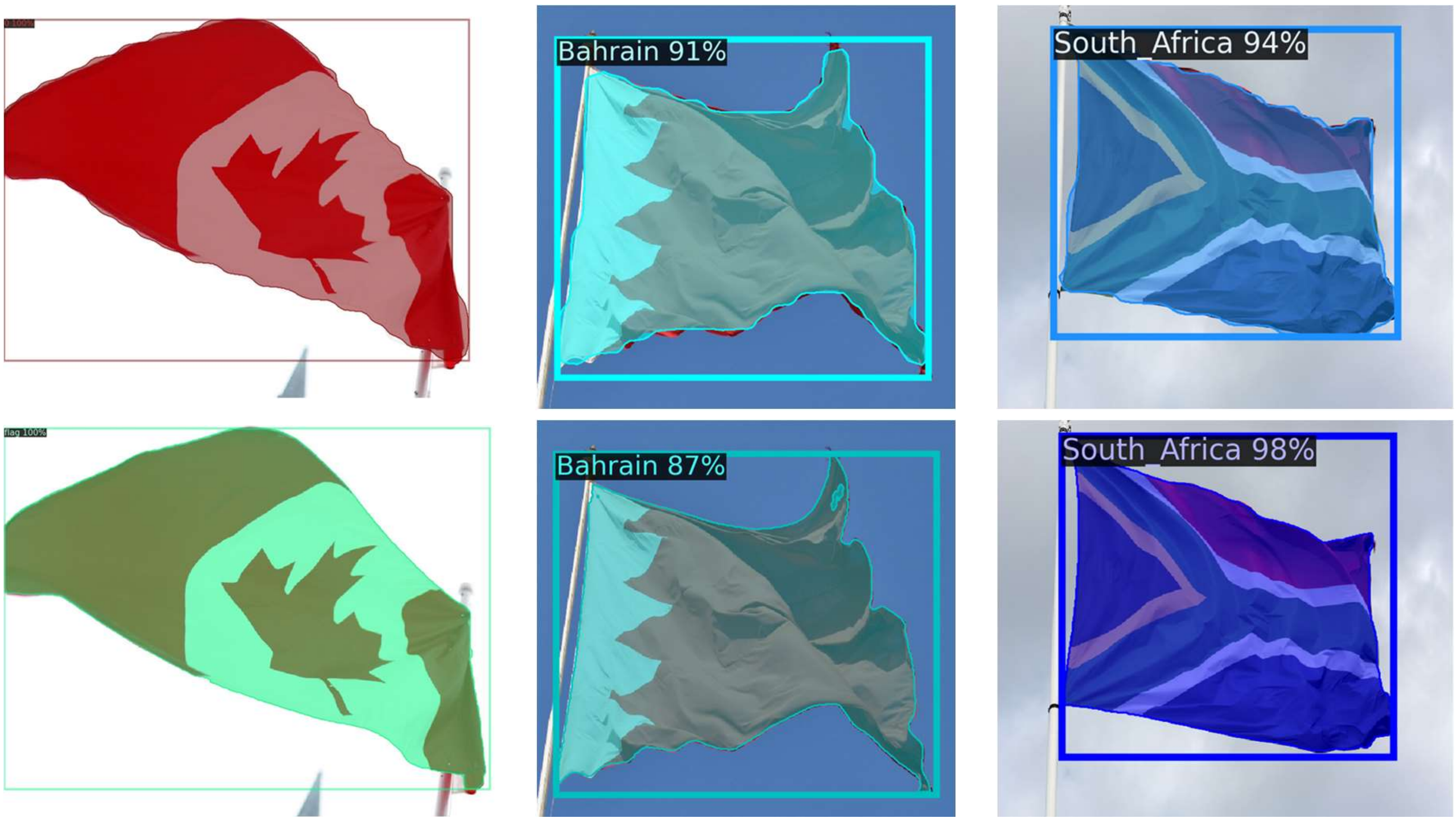


Visual Result

- PointRend predicts clearer segmentation masks

Mask-RCNN

PointRend



Discussion and Future Work

- **Limitation** : heavy occlusion causes mis-classification

(a)

(b)

- **3D engine** (b) to create realistic template and simulate deformation
- **Teacher-Student network** (semi-supervised learning) from flag images without annotation

Thank you for your listening.

Experiment

- **Binary Flag Detection**
  - *YOLACT++*
    - not ideal for delineating complex flag instances
  - *PointRend* has slightly better performance than *Mask-RCNN*

	Backbone	AP	AP_{50}	AP_{75}	AP_s	AP_m	AP_l
YOLACT++	ResNet-101-FPN	17.26	20.32	18.27	-	-	-
Mask-RCNN	ResNet-50-FPN	69.90	87.72	78.16	18.00	45.62	77.07
Mask-RCNN	ResNet-101-FPN	73.85	89.13	80.35	17.02	47.32	81.96
Mask-RCNN	ResNeXt-101-FPN	72.30	88.81	79.02	**32.65**	48.76	79.36
PointRend	ResNet-101-FPN	**76.30**	**90.24**	81.90	20.40	47.68	**84.68**
PointRend	ResNeXt-101-FPN	76.21	89.63	**82.06**	25.91	**49.75**	84.30

Binary flag detection - YOLACT++

- Miss
- False alerts, i.e. parachutes, birds