Objects of Violence

Synthetic data for practical ML in human rights investigations



THE PAPER

https://lachlankermode.com /objectsofviolence_neurips_2019.pdf

Eoropoio Architecture

Lachlan Kermode

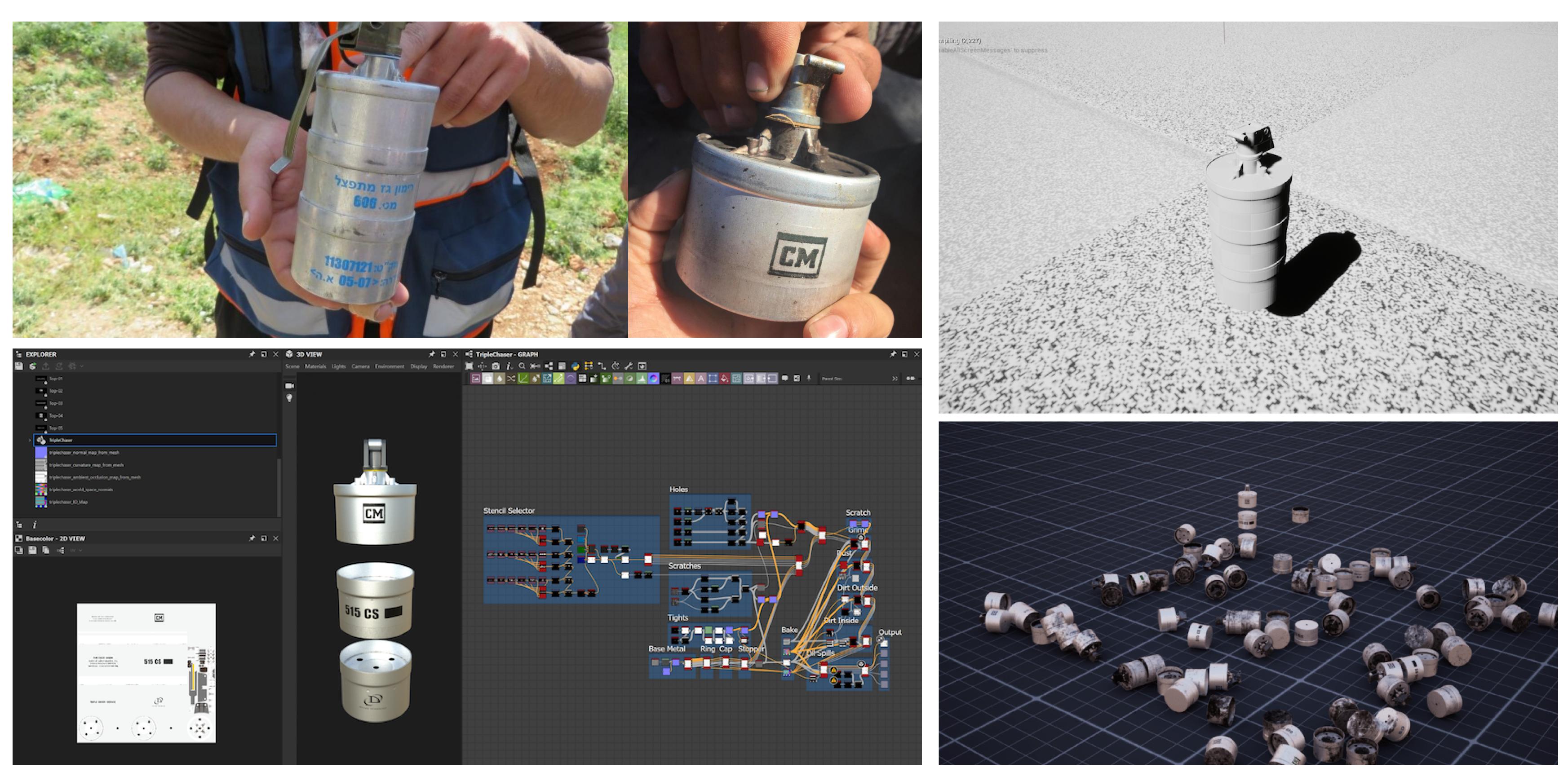
lk@forensic-architecture.org

Alican Akturk Denis Kochetkev Julien Cornebise

Robert Trafford Rafael Pardinas Eyal Weizman

MODEL CREATION

Modelled on the basis of real images. Parametric textures created with Substance.

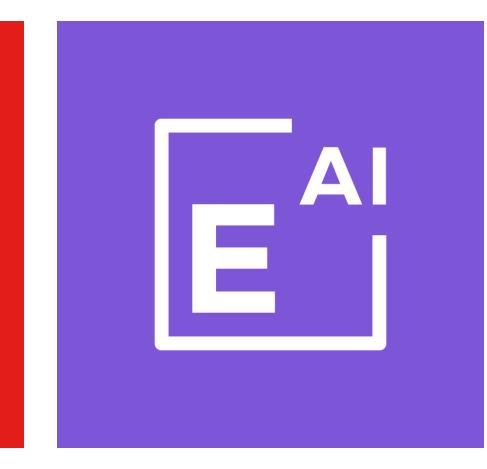


SYNTHETIC DATA

Engines written to take 3D object as input, and generate synthetic image output. **Controlled** Variation





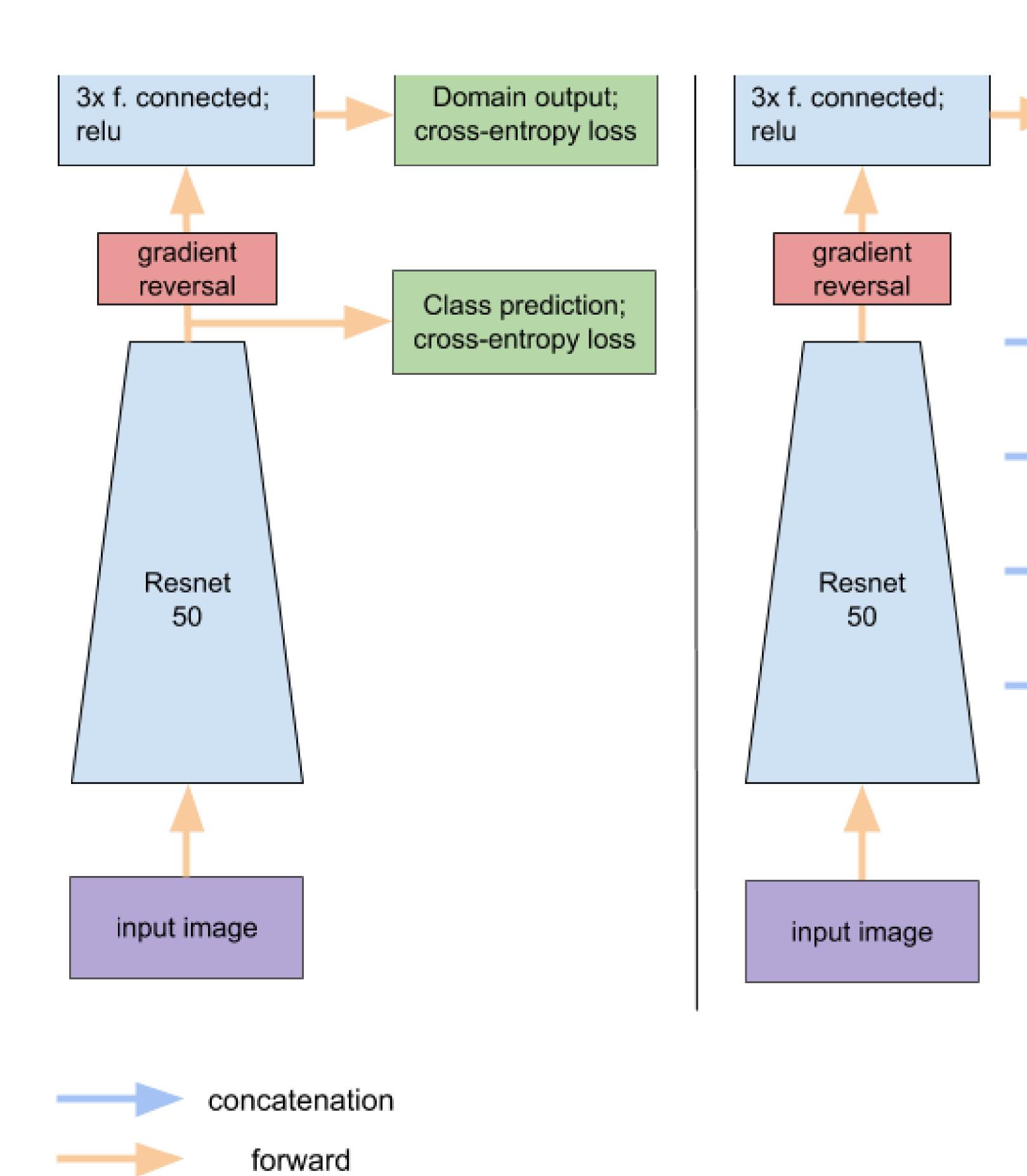


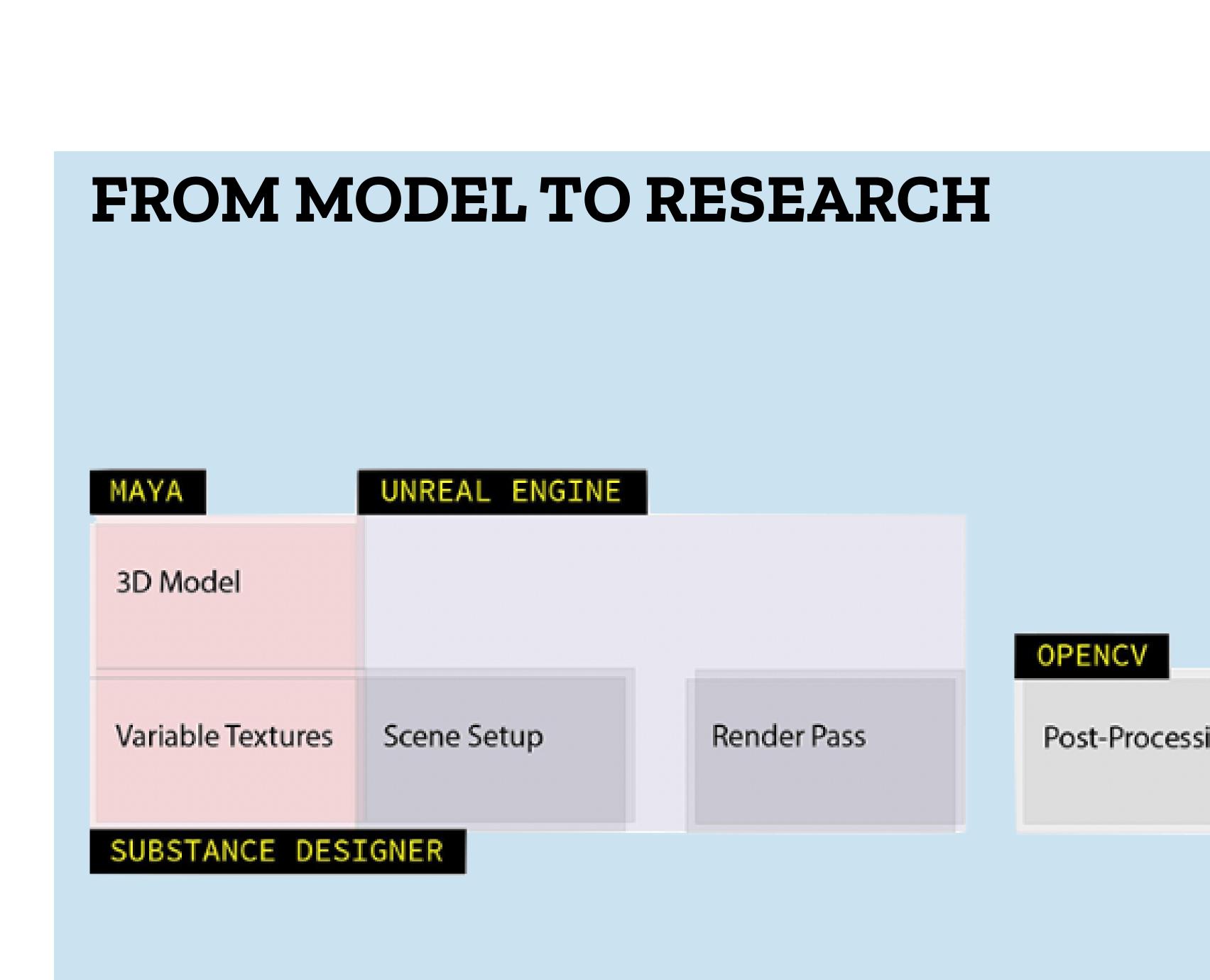
Jan Freyberg jan.freyberg@elementai.com

Flying Distractors

CLASSIFIER TRAINING

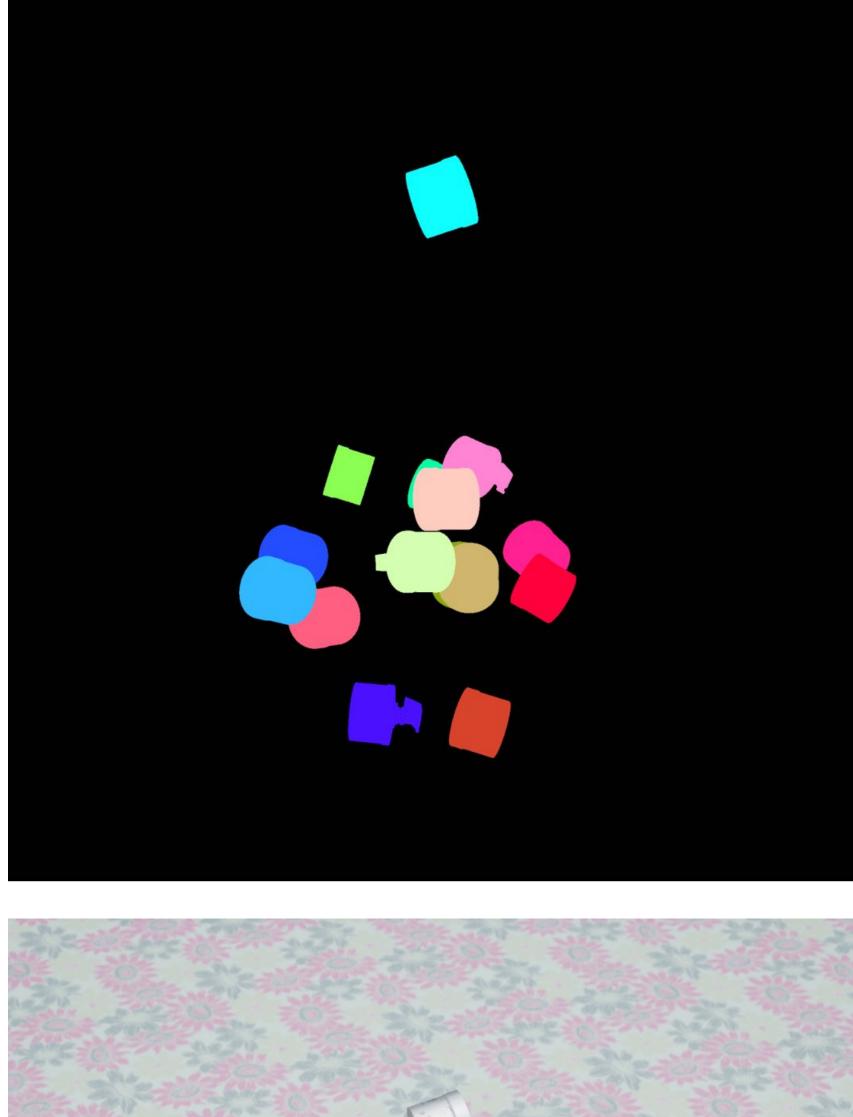
ResNet 50, UNet*, and Mask RCNN*. Trained to predict Triple-Chaser, as well as 20 "nuisance" classes from Pascal VOC. Adam optimisation and L2 regularisation. Optional domain adversarial module using reverse gradiants, sampling from VOC. * ImageNet pre-trained ResNet 50 as initial encoder.





AUTOMATIC ANNOTATION

Scripts to generate annotations by comparing render with masks.

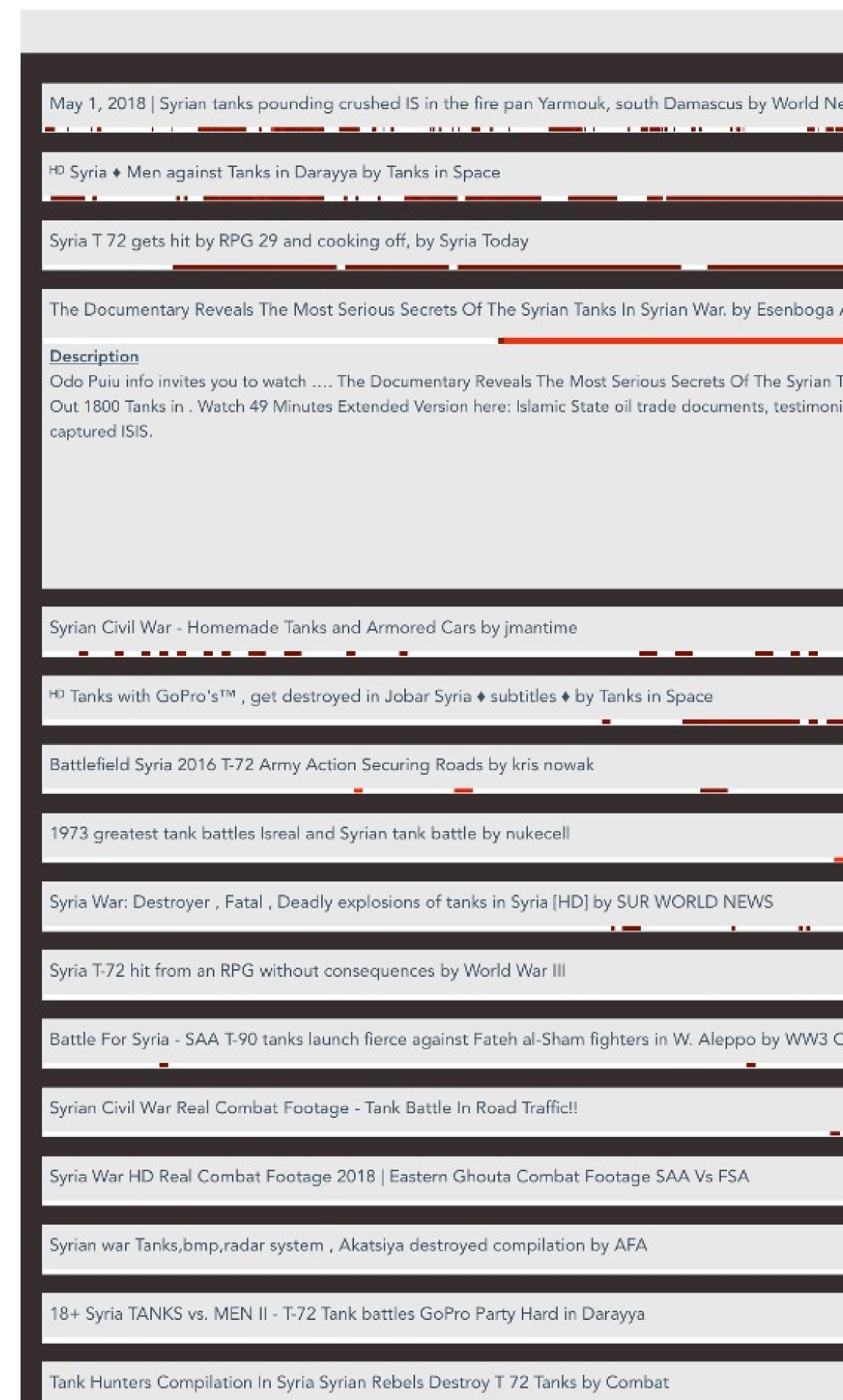


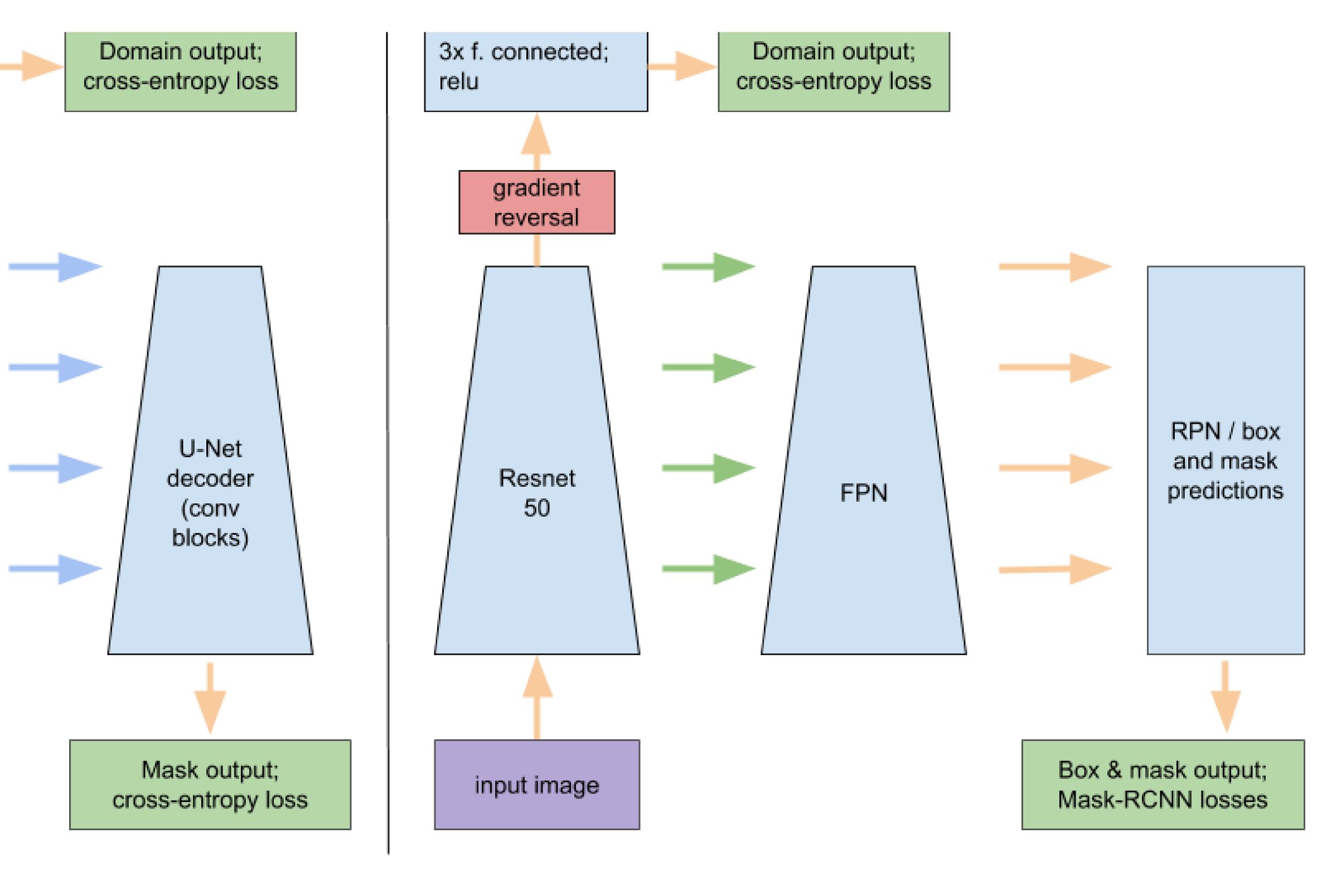


FRONTEND WORKSTATION

Mtriage presents a web interface to visualise results as a searchable interface for researchers.

CLI tool to deploy any logic in parallel on media. Streamlines scraping heuristics across platforms.



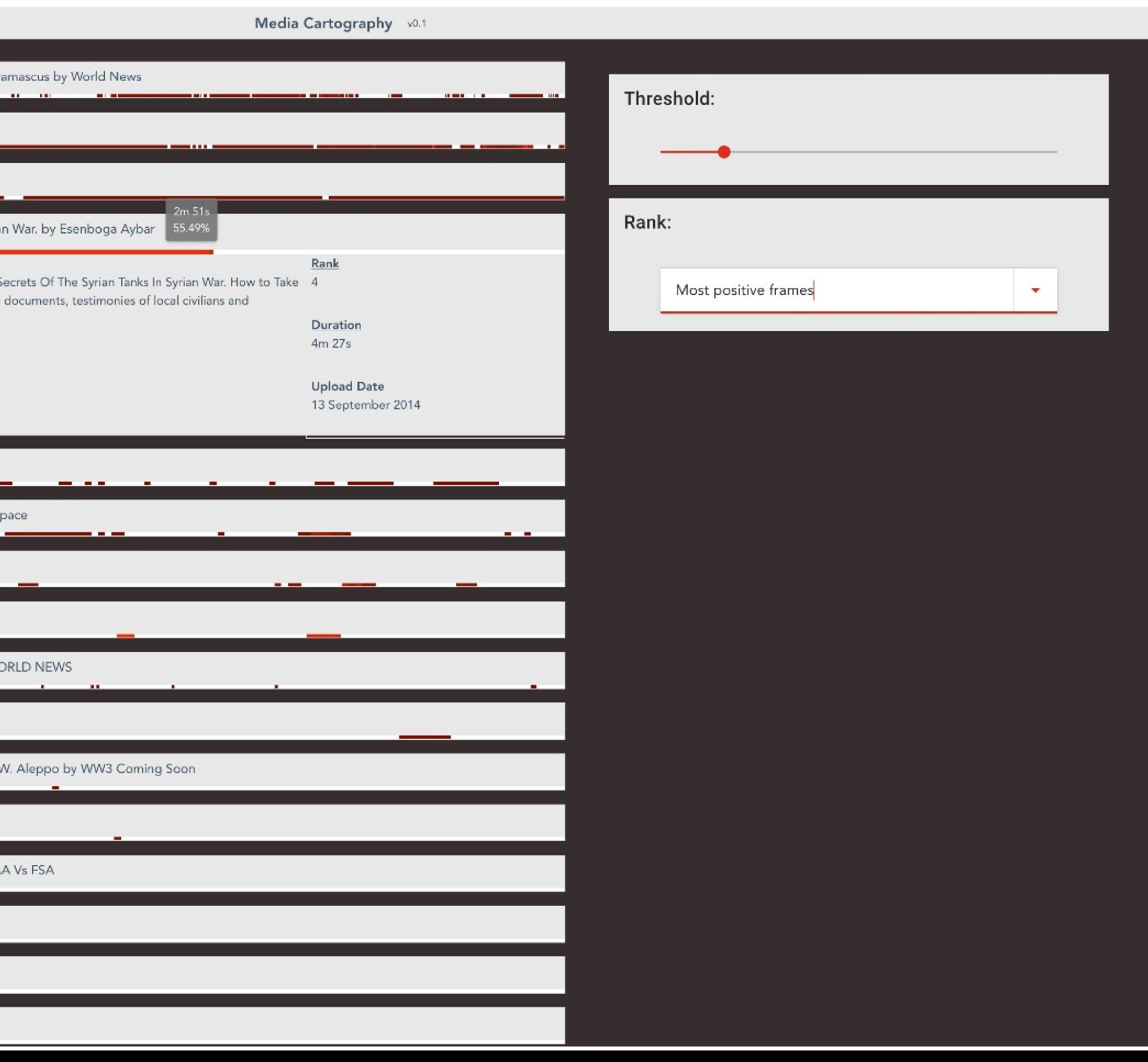


				Select	<section-header><section-header><section-header></section-header></section-header></section-header>
				Pre-Processor	
ing	Synthetic Data	PYTORCH Classifier Training	Trained Model	Analyser	
				Predictions	Web Platform

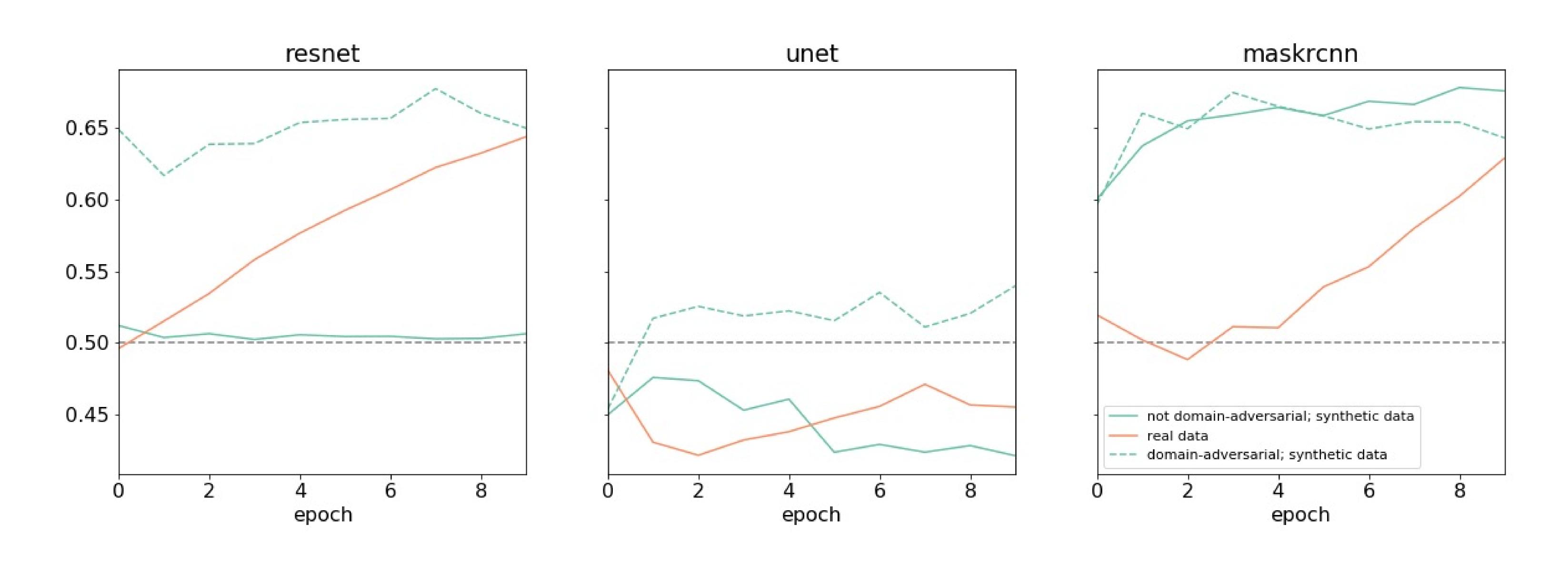


MTRIAGE

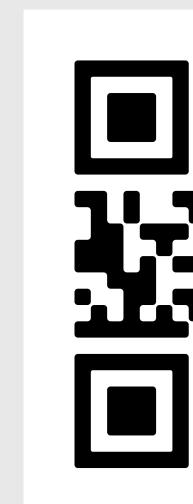
https://github.com /forensic-architect



RESULTS



Evaluation of this work purely with classifier accuracy does not assess theusefulness of the workflow in practice. Thresholds for network predictions are set dynamically: after ranking based on network scores, humans simply start at the top and work their way down until they realise the predictions are no longer accurate.



- Generate synthetic datasets and train classifiers for other objects. - Expand knowledge on what quality of synthetic data is necessary. - Add new platforms to Mtriage, and new visualizations.



Each architecture evaluated on: **1.** Real data benchmark (sample held out for evaluation). **2.** Synthetic data, sampled from low and high fidelity sets.* **3.** Synthetic data with domain adversarial module.* * Evalution on ~130 labelled images sourced online.

TAKEAWAYS

- Training with **synthetic data** - increased performance of the model - allowed models to train significantly faster.
- Training with **domain adversarial methods**
- significant performance increases with classification models
- smaller with semantic segmentation model
- nonexistent with mask model.

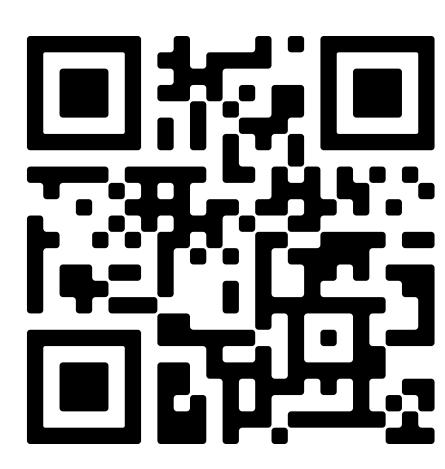


THE INVESTIGATION

https://forensic-architecture.org /investigation/triple-chaser

FUTURES

Help us develop open source, collaborative infrastructure for human rights research.



JOIN OUR DISCORD

https://discord.gg/FJ4XsCg