



Synthetic Training Data For Rare Object Detection in Computer Vision

Max Klein^{1,2}, George Weinert¹, Michael Zelinski¹



¹ Lawrence Livermore National Lab ² GroupLens Research, University of Minnesota

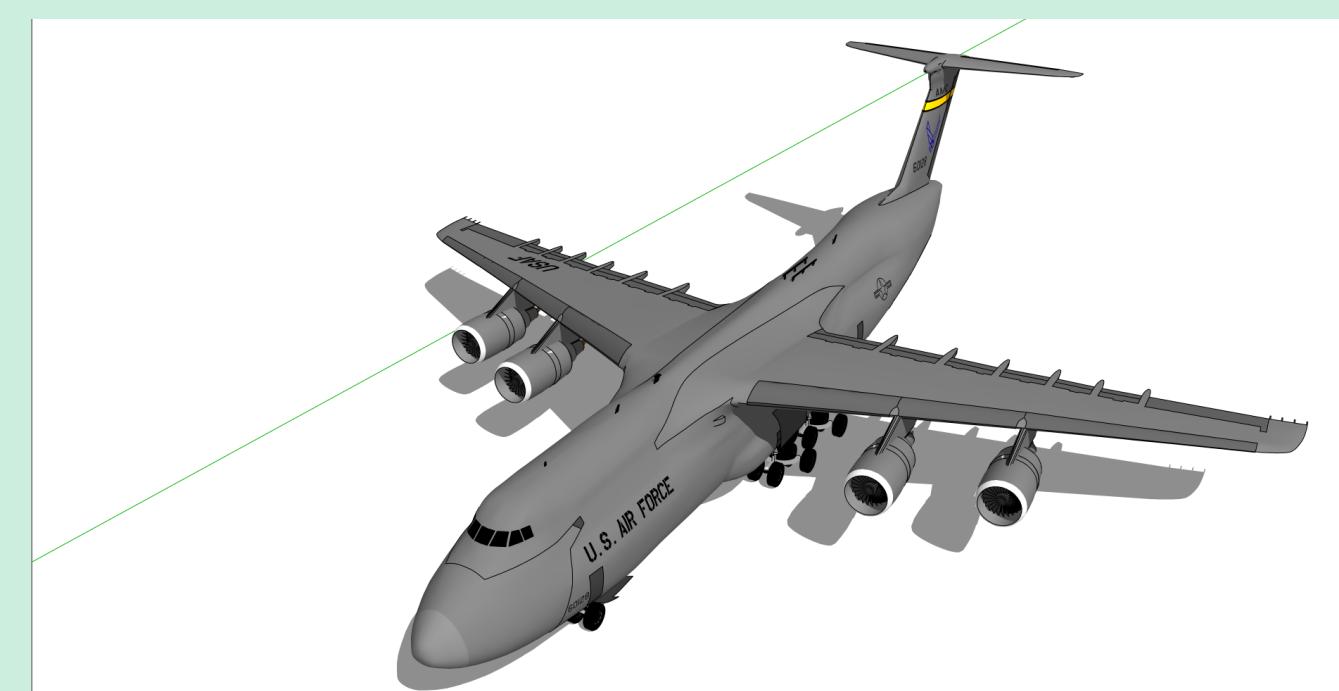
Convolutional Neural Networks do well at detecting objects for which there are many examples, but what can we do about rare objects of which there are *no natural examples*? We experimented with a using *synthetic training dataset* made from CAD models.

Introduction

- Rare objects, where only 1 to 100 natural examples exist present a problem for deep neural network object detection algorithms.
- For instance, normally it would be hard to **detect rare Airplanes in satellite imagery**.
- Idea:** use Computer Assisted Design (CAD) models to make “synthetic” training data instead of natural training data.
- RQ:** How well can **synthetic training data** classify **natural examples**?



Less than 100 C5 aircraft

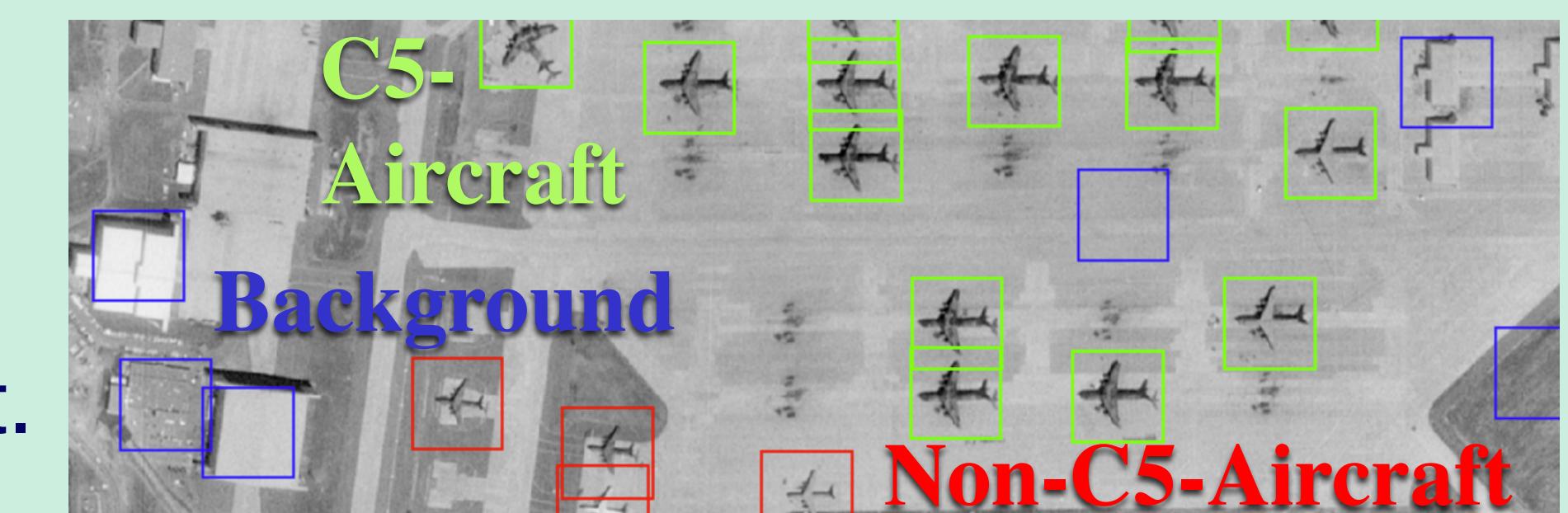


...but its easy to make images of them in CAD.

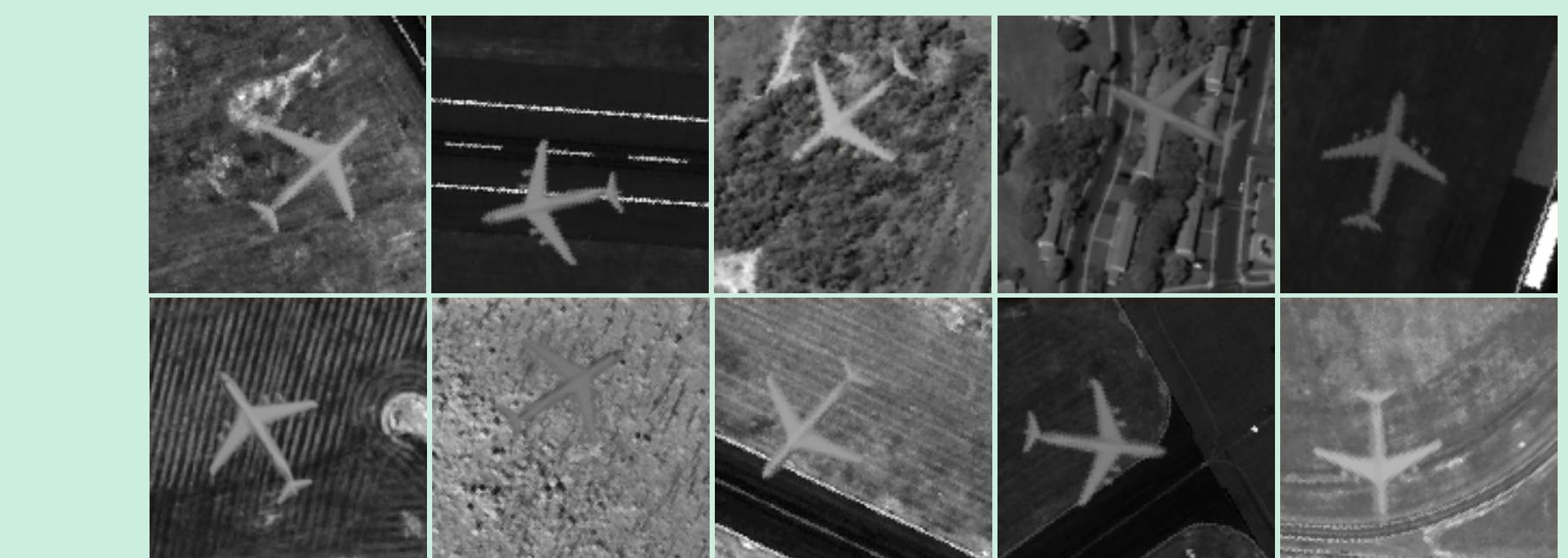
Methods

- Annotating natural satellite images.
- Compositing synthetic CAD models on natural backgrounds.
- “VGG” Very-Deep Conv. Neural Net.

Source	# background	# c5	# confusers
Natural	4568	281	405
Synthetic	0	1184	0



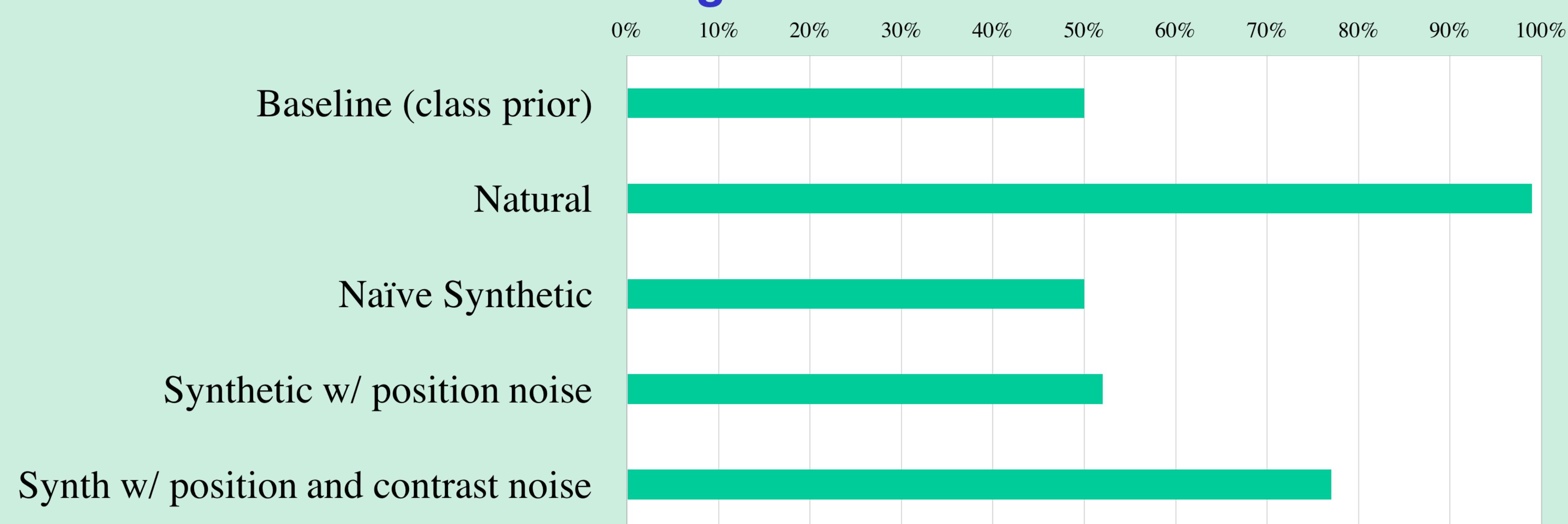
Natural training data, annotated



Synthetic data, composited

Results

Evaluation of Training Sets Predicting Natural Test Set



Discussion

- This “0-shot” learning problem is **very hard**.
- Synthetic** performance shows **some promise** of predicting natural object, **but only with** a lot of **noise** in generating the synthetic data.
- Future** work should focus on **optimizing the CAD-compositing pipeline** to generate more realistic images with high variation.



Detect objects which we've never manually identified in satellite images?