



Developing and Deploying AI in Risk Adverse Industries



Christian Thurow
Searidge Technologies

Intro SRT



- ❖ Leading Provider of Remote Tower and Airport Surface Optimization Solutions
- ❖ Technology at 30+ sites in 17 countries
- ❖ ISO 9001:2015 Certified
- ❖ Co-owned owned by NAV CANADA and NATS



- What is Searidge doing?
- What issues are we trying to solve?

Intro SRT





Surface Optimizations

- what we mean by surface optimizations
- maximising movements, time = money
- these systems are helpful and might be powered by AI or not
- they might need to get certified
- lets get into some of the systems that do need to get certified for operational use



Presentation Structure

First part

Real world examples for certified ATC systems, with and without AI

Second part

Tools for understanding AI, as first step in certification

First European Multi RWY RT

- example system - vision only: Budapest, first multi runway remote tower
- show how the system looks, what it does
- requirements, basics
- Video Wall (24x fullHD screens) driven by 6 beefy machines with GTX 1080s

Non-AI example



Non-AI example



Certification Process non-AI



- regulator cared more about documentation and processes followed during implementation than anything else
- Technical:
 - Quality Management (Engineering Process), like ISO 9001
 - Technical Validation: FAT, SAT, Component level validation (Unit Tests & Reviews)
 - Performance, system and industry specific, ED240
- Safety:
 - Software Assurance, ED109, DO-278
 - Interoperability, EC-552, EC-482
 - Security ISO 27001

Certifying non-AI, but new tech

Certification Process non-AI



- TODO: image

Findings from Experience



- problem that needs solving -> new system as solution -> build trust and acceptance within industry -> first buyer willing to test it out(only possible if very tech savvy or general acceptance level is high) -> user acceptance (air traffic controllers) -> actual certification process can start
- future projects with same tech go much quicker: just problem, solution, buying, certify (if not same country)



How to build Trust - AI systems

“The more complex the information, and the more consequential its recommendation/prediction, the more critical it will be that its operators should understand how it decides things.” - Danny Bradbury

From Traditional Systems to AI



- TODO: insert image

How we use AI

- main focus on visual spectrum/thermal camera image classification
- but also non-vision applications, e.g. ASMGCS based:
 - go-around detection&prediction (missed approaches)
 - congestion detection&prediction
 - TODO

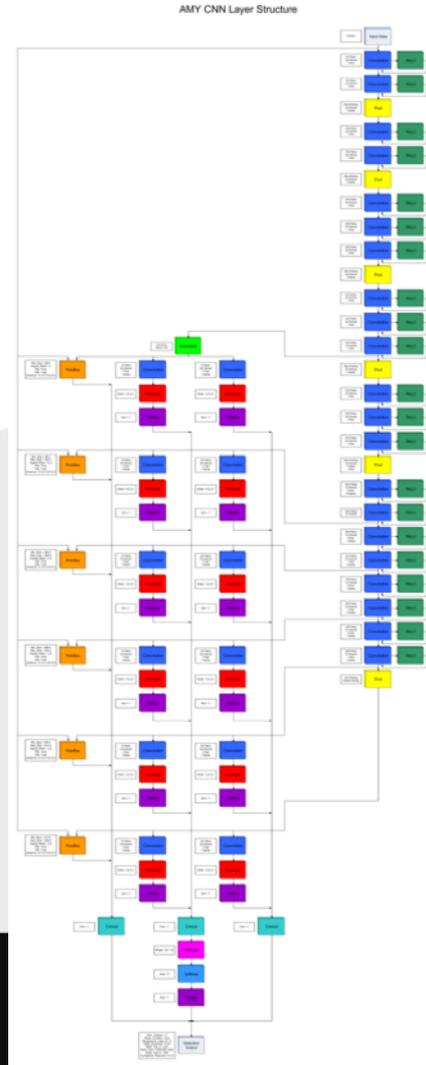
How we use AI

- Image Classification:
- ANN, SSD/Faster R-CNN
- on top: tracking, positioning
- use of Nvidia GPUs
- Quadro for server and server like WKS + TitanX for WKS
- caffe, TensorRT, NV GPU CLOUD (NGC -> Docker Images for Dev)

certifying AI

Certifying AI in ATC

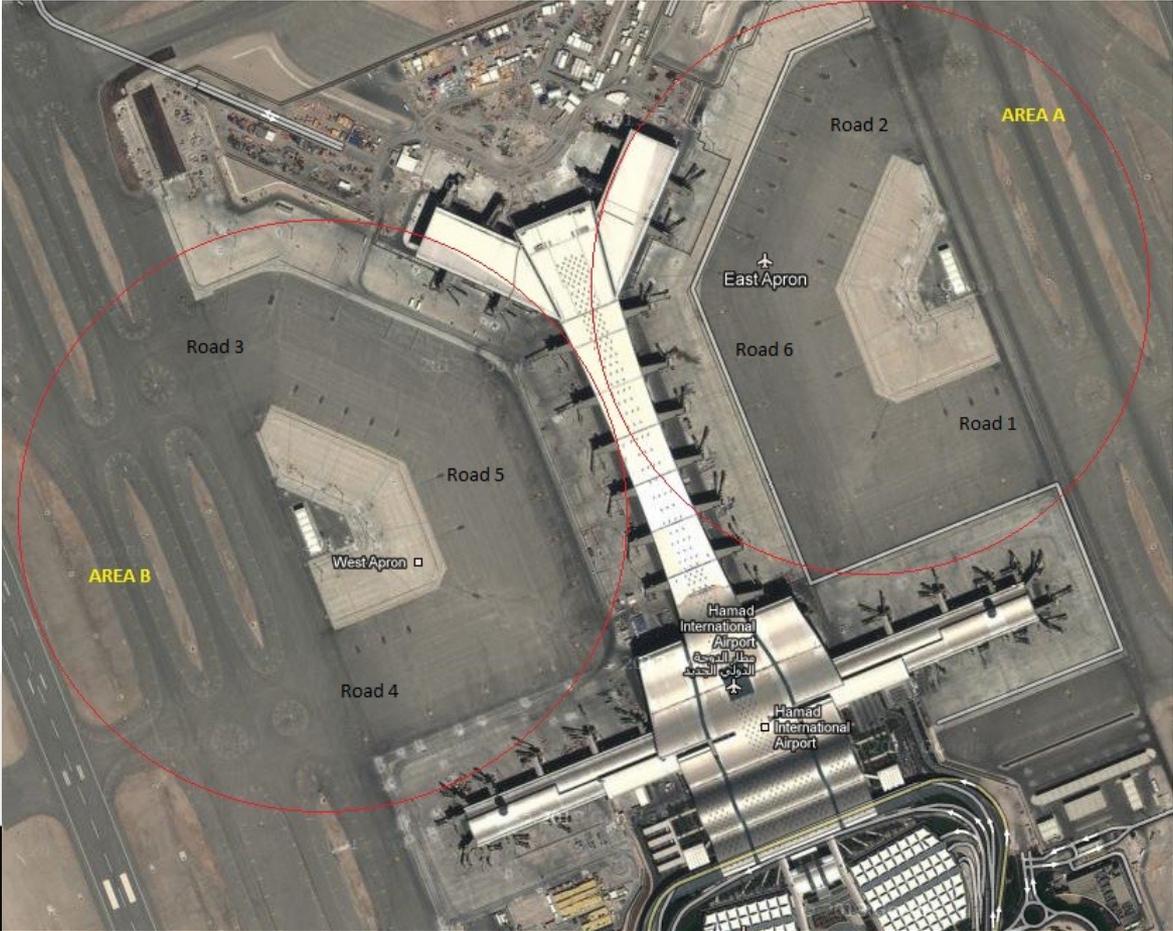
- why is certifying AI different from any other system?



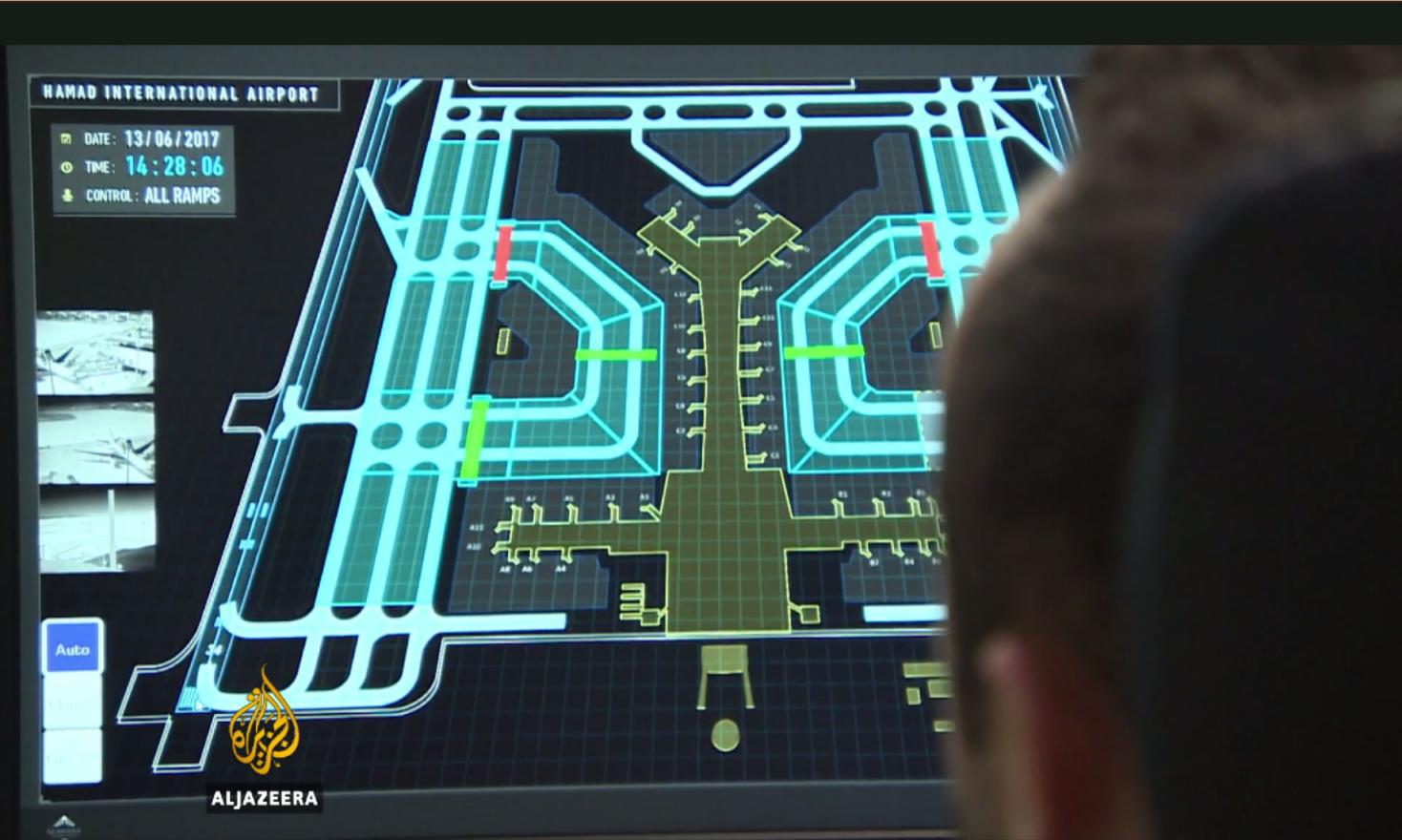


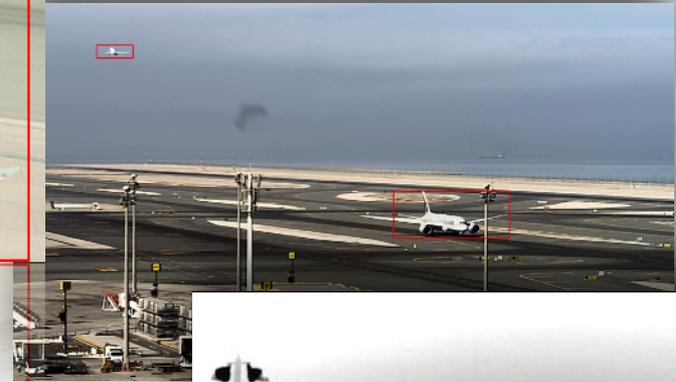
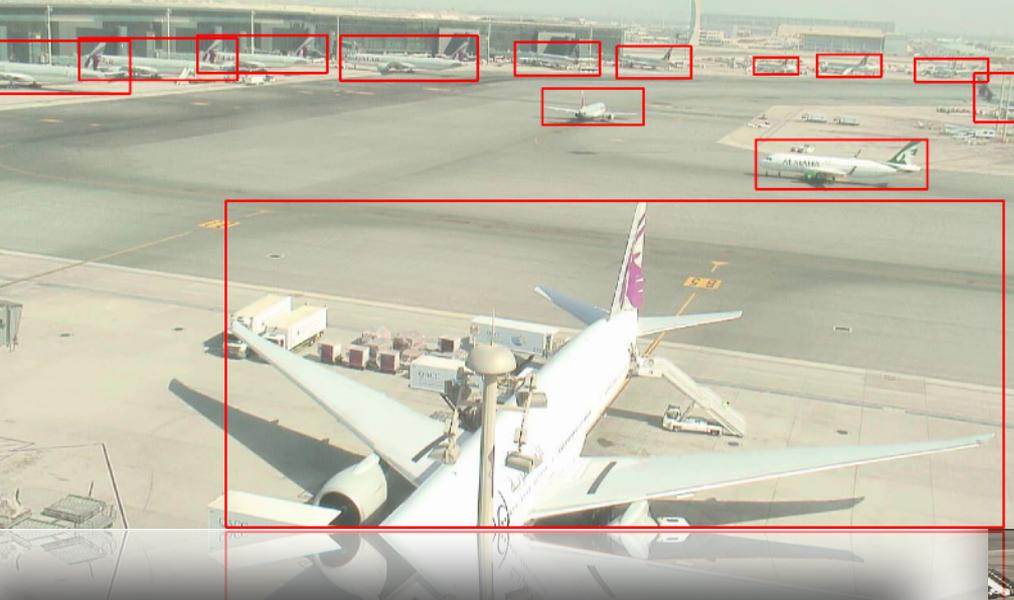
First ever AI system certified in ATC

AI specific example - DOHA



AI specific example - DOHA







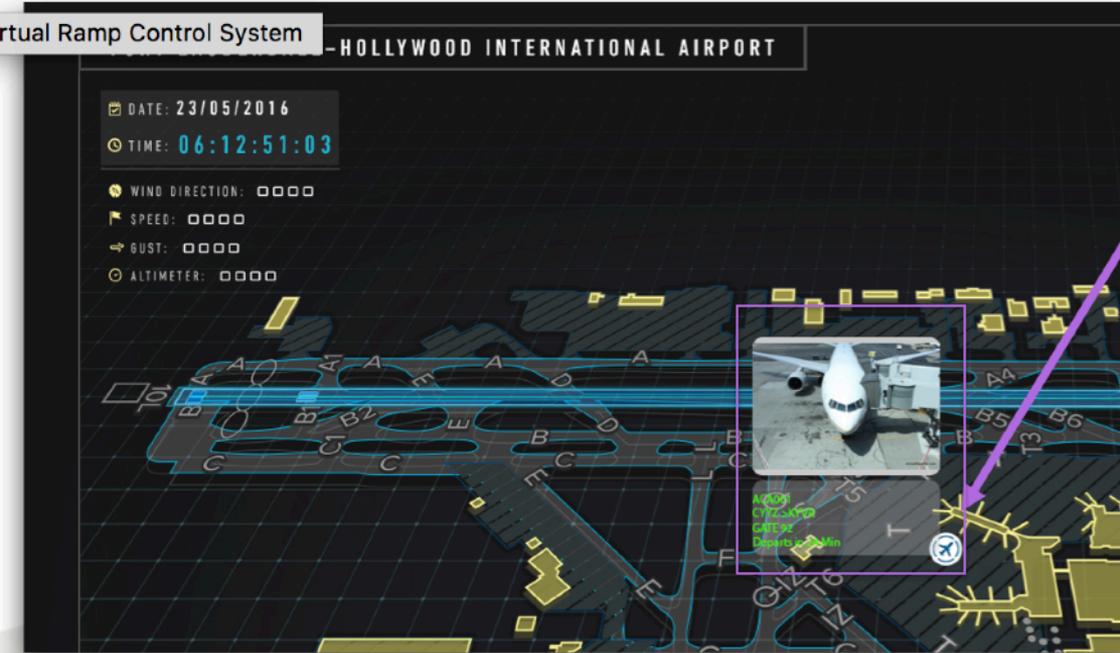
AI specific example - FLL



AI specific example - FLL

AODB + A-SMGCS - EAVD Integrated Map Display

Reference: Ft. Lauderdale VRCS



Click on AC Target or Gate

- Live Video
- Fused data from A-SMGCS + AODB
 - Scheduling
 - Live Status (e.g. on-time, delayed ,etc.)
 - Flight Plan
 - Counters: e.g. TOBT



Certifying AI in ATC

- Certification of non AI ATC systems same as for AI ATC
- Regulator only cares about safety within the framework of defined concept of operations
- as long as AI is tool and human makes the ultimate decision, not as a big issue
- good approach: don't start with it as primary system, have a fallback system just in case
- incremental approach/ phased approach

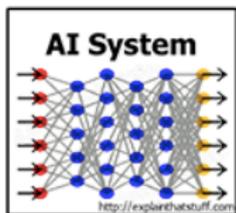


Certifying AI in ATC

- understanding AI, in the end, humans need to be confident in the system
- not enough that engineers are confident, need to have users and regulators on the same side
- identify where understanding lacks

Interpretability

- interpretability is building trust
- DARPA: XAI



- We are entering a new age of AI applications
- Machine learning is the core technology
- Machine learning models are opaque, non-intuitive, and difficult for people to understand



- Why did you do that?
- Why not something else?
- When do you succeed?
- When do you fail?
- When can I trust you?
- How do I correct an error?

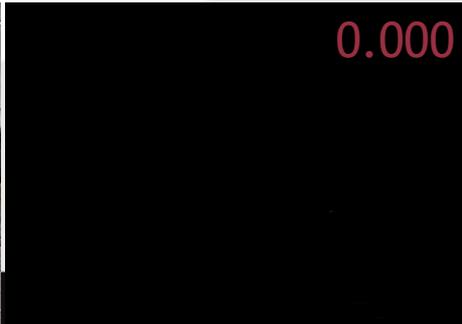
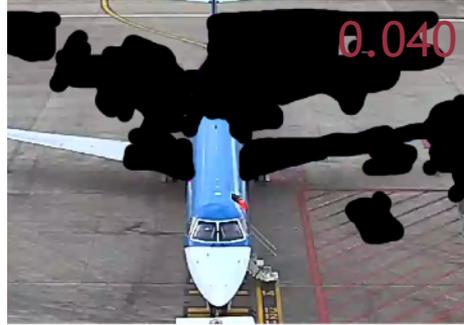
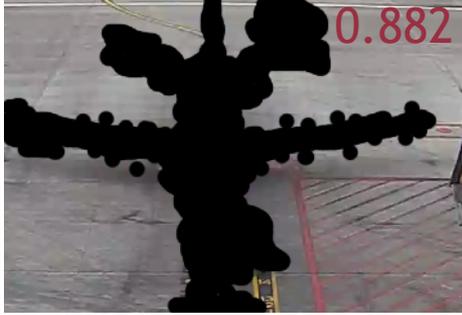
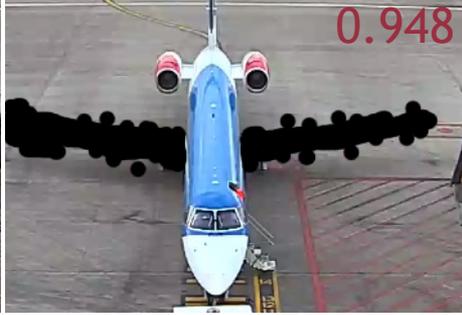


Tools for Understanding AI

Shining light into the black boxes.

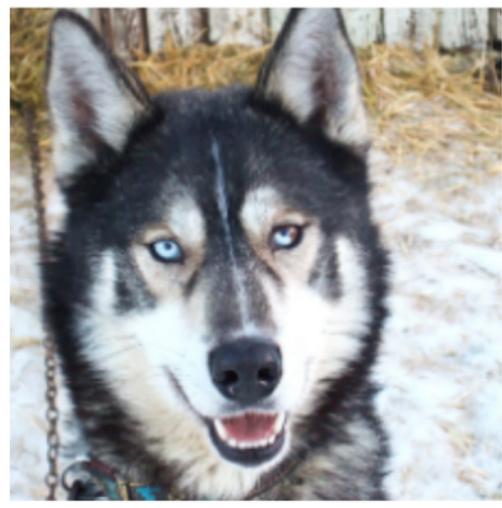
Sanity Checks





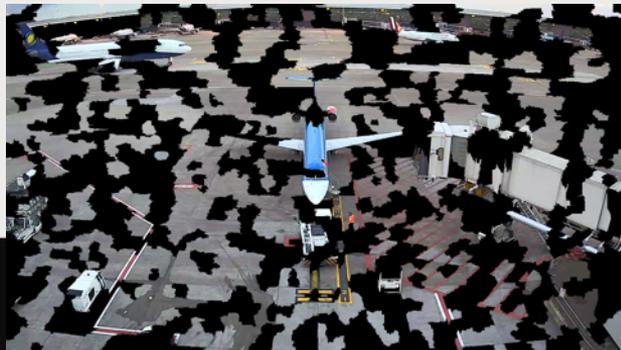
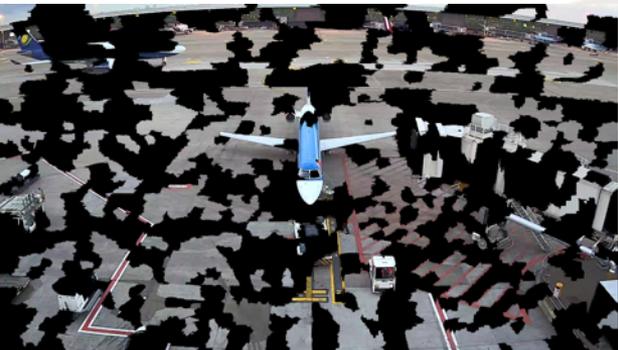
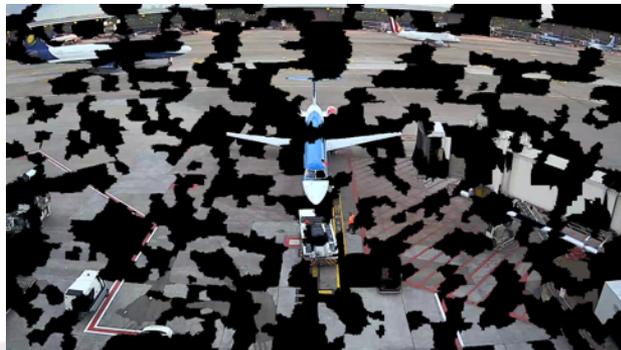
LIME - Local Interpretable Model-Agnostic Explanations

- “Why should I trust you?” Explaining the Predictions of Any Classifier, 2016 Ribeiro et al.
- Github: [marcotcr/lime](https://github.com/marcotcr/lime)



LIME





LIME



Input



Segments > 0.1



All Segments



Segments > 0.5

LIME



Input



Segments > 0.1



All Segments

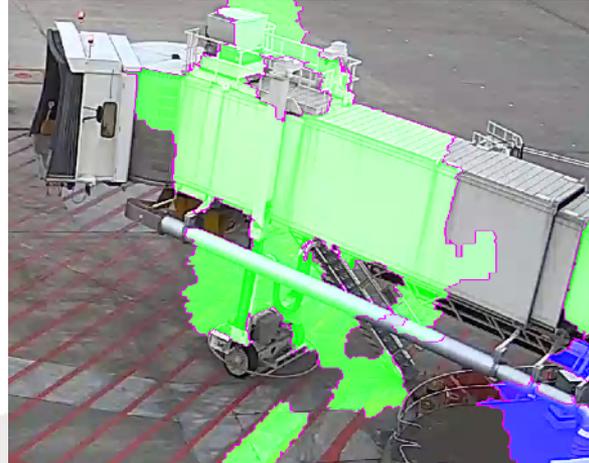


Segments > 0.5

LIME



Input



Segments > 0.01

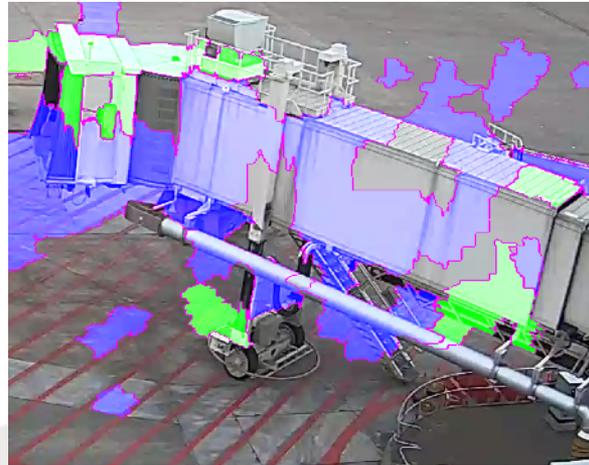


Segments > 0.1

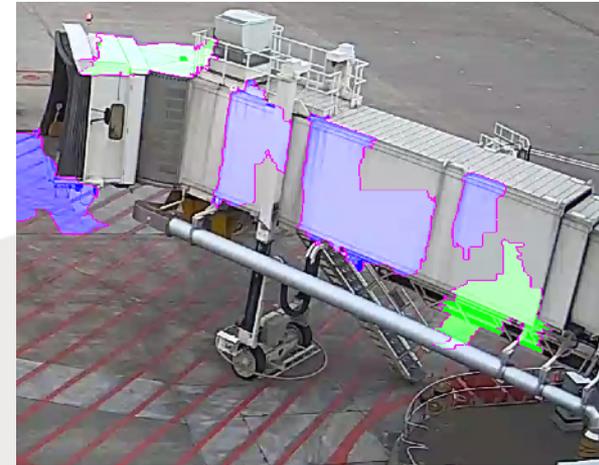
LIME



Input



Segments > 0.01

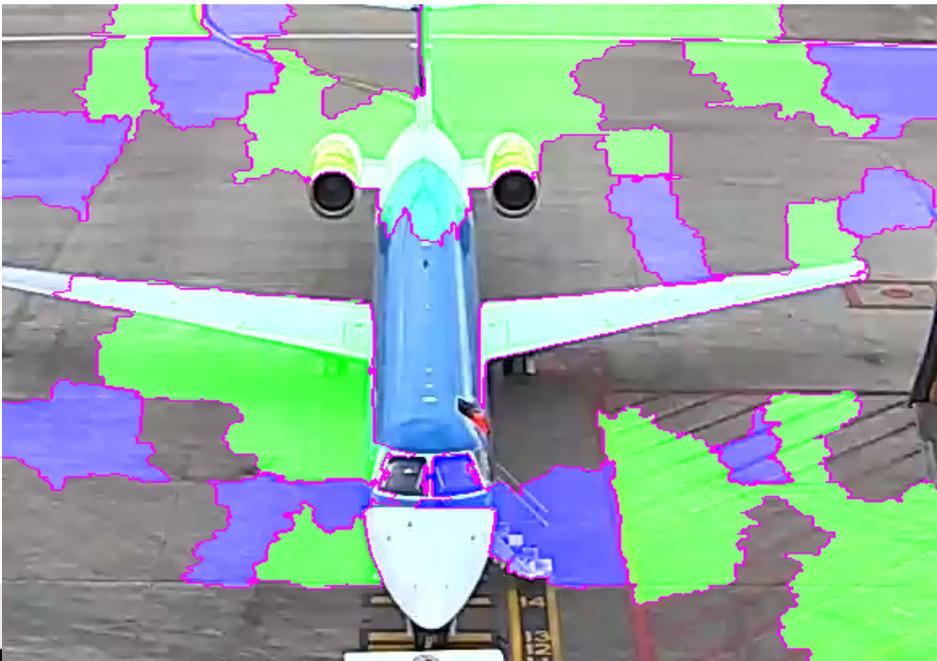


Segments > 0.1

LIME



LIME



Network in Reverse



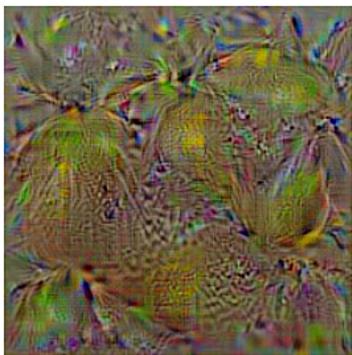
Network in Reverse



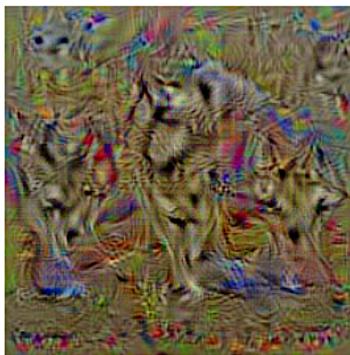
Network in Reverse



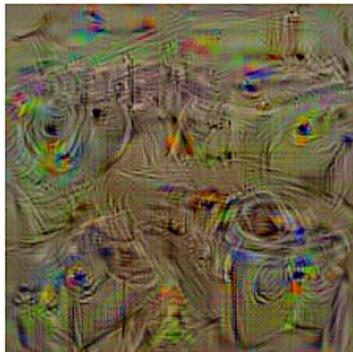
bell pepper



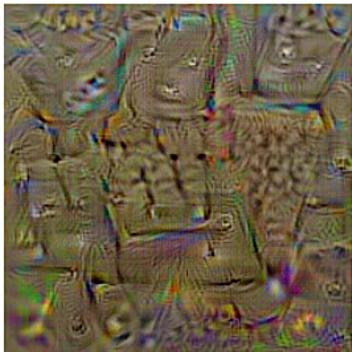
lemon



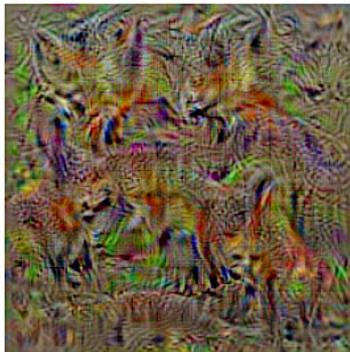
husky



washing machine



computer keyboard



kit fox

- Numerically computed images, illustrating the class appearance models, learnt by a ConvNet, trained on ILSVRC-2013
- Paper: “Deep inside convolutional networks: visualising image classification models and saliency maps” by Karen, Vedaldi, Zisserman

Network in Reverse

- paper: “Visualizing and Understanding Convolutional Networks”, 2013 by Zeiler and Fergus

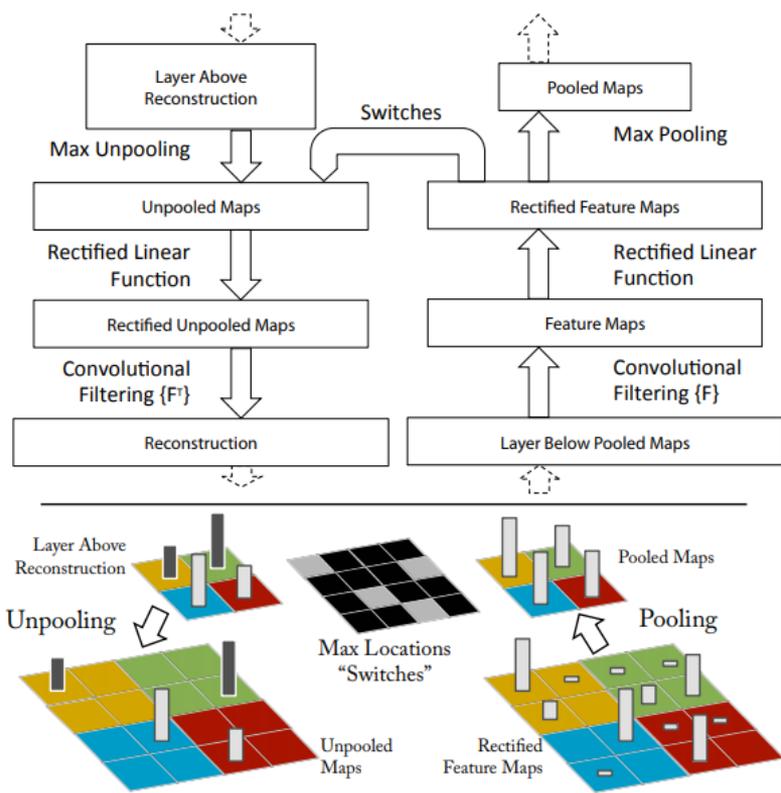


Figure 1. Top: A deconvnet layer (left) attached to a convnet layer (right). The deconvnet will reconstruct an approximate version of the convnet features from the layer beneath. Bottom: An illustration of the unpooling operation in the deconvnet, using *switches* which record the location of the local max in each pooling region (colored zones) during pooling in the convnet.

Network in Reverse



Network in Reverse



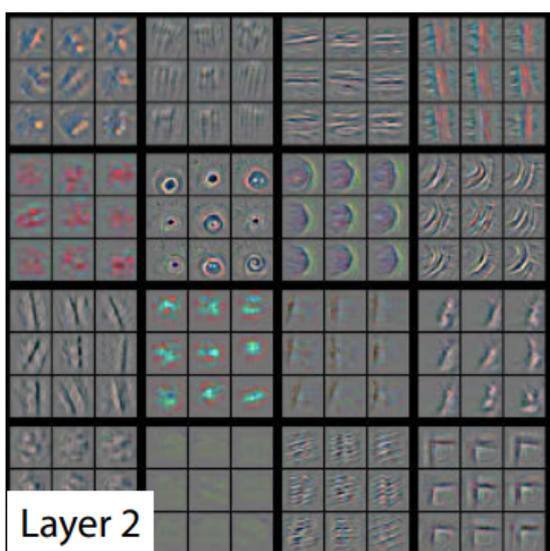


“We must go deeper”

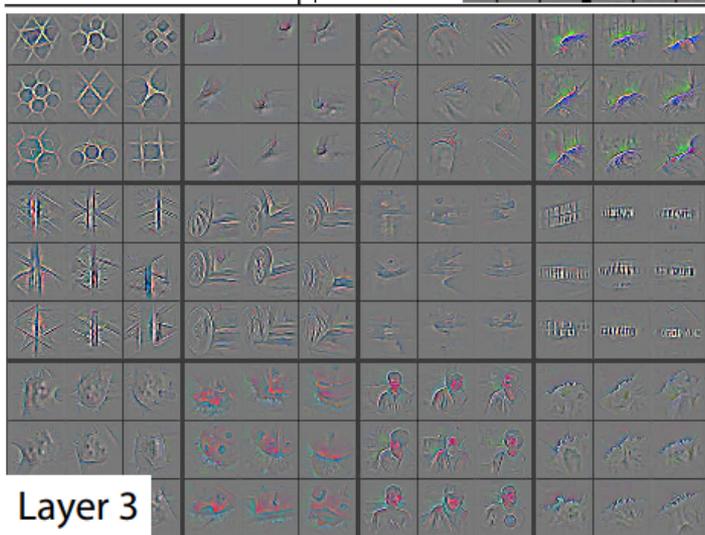
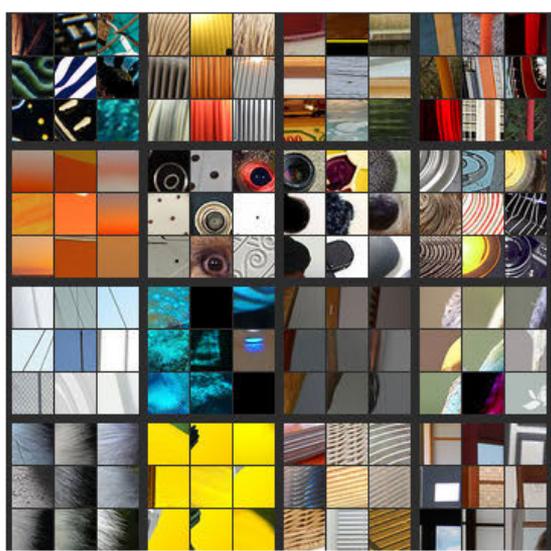
Tool:



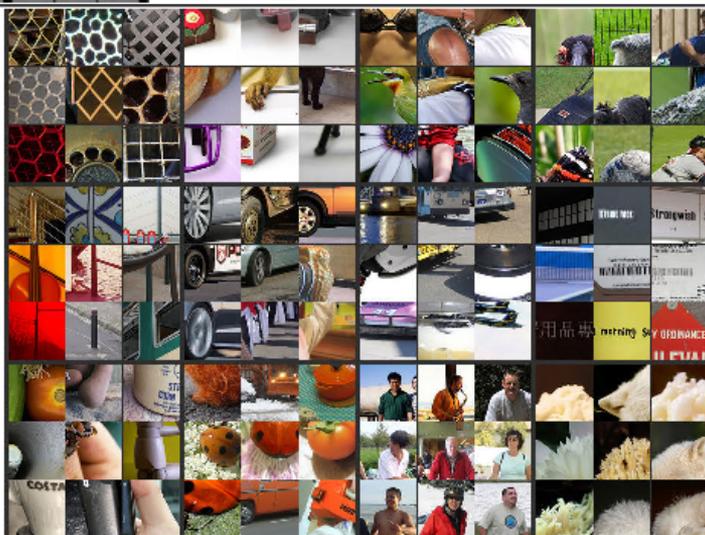
Layer 1



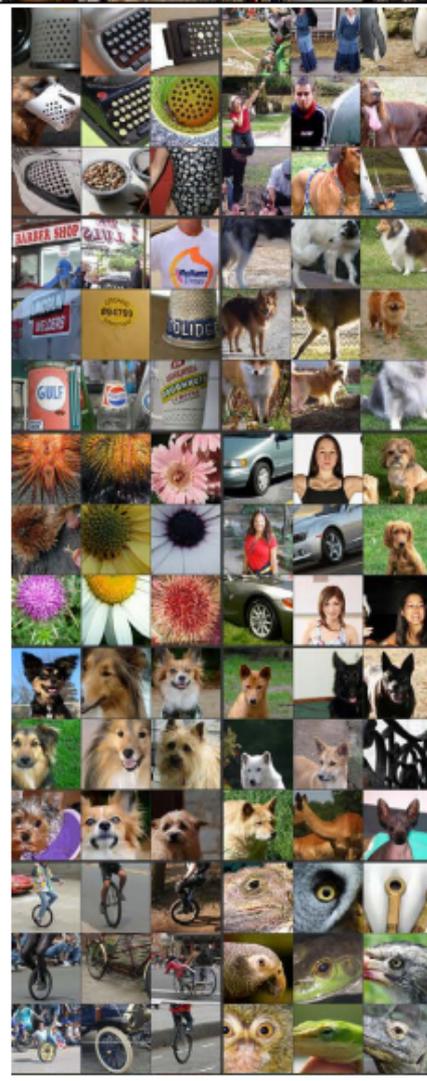
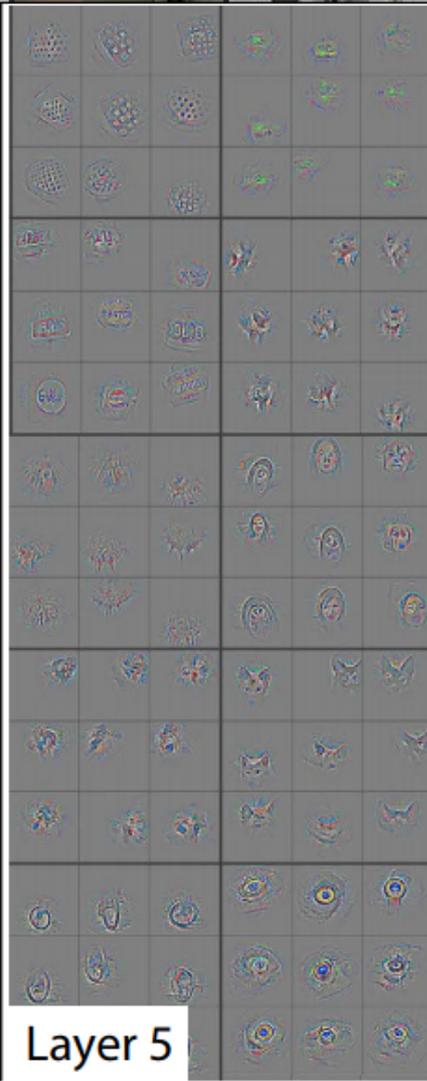
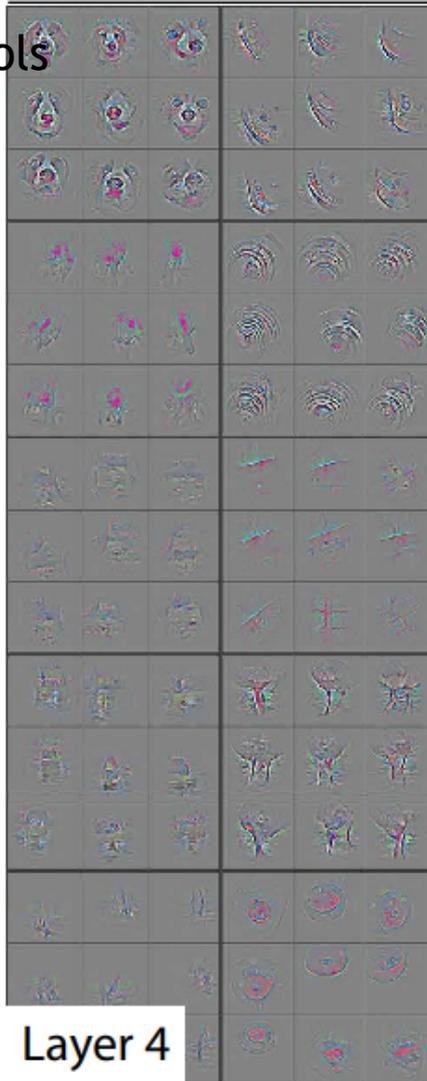
Layer 2



Layer 3



Tools

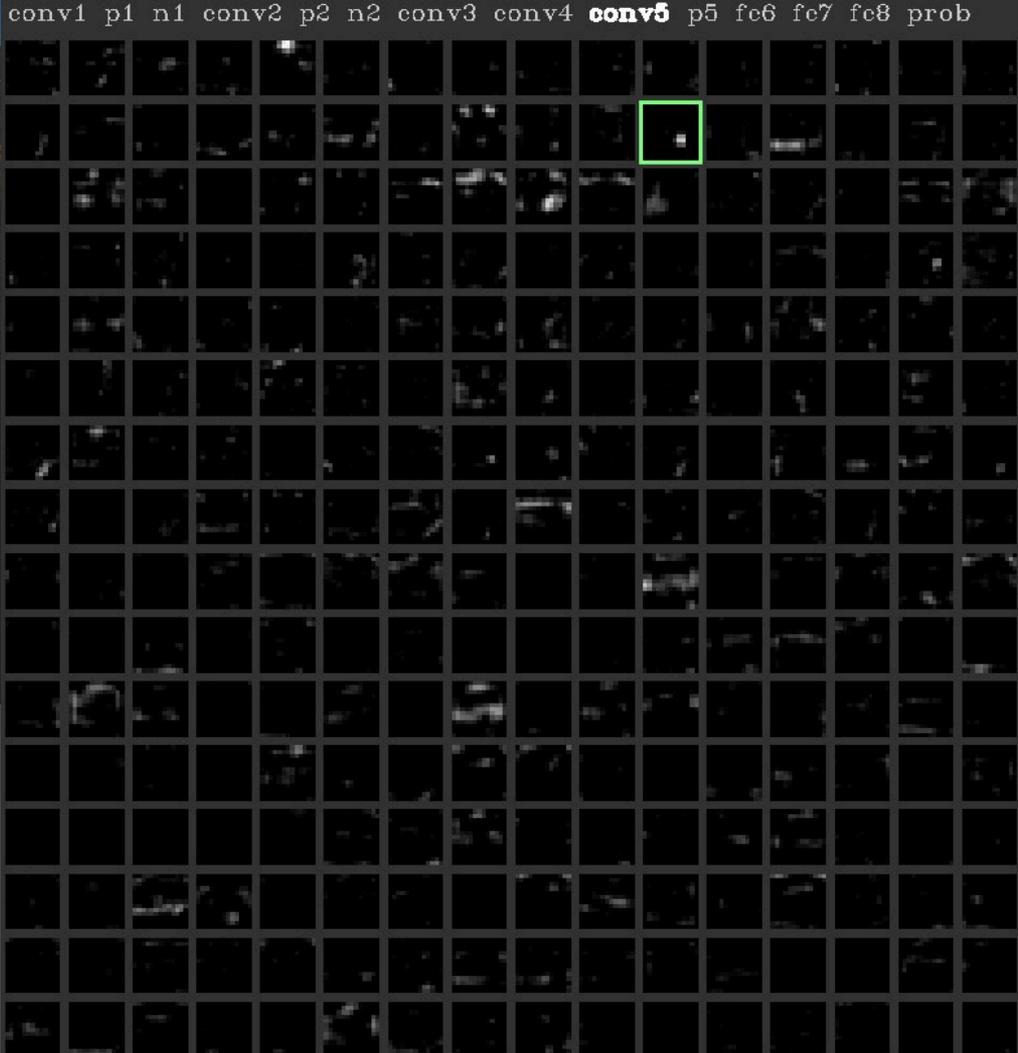
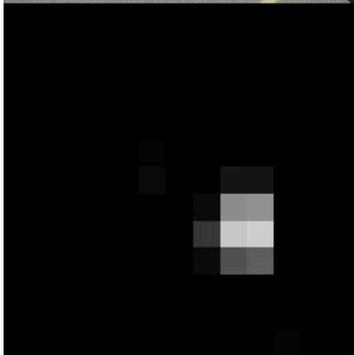




Deep Visualization Toolbox

- paper: “Understanding neural networks through deep visualizaton” by Yosinski et al.
- github: [yosinski/deep-visualization-toolbox](https://github.com/yosinski/deep-visualization-toolbox)

Too



fwd conv5:26 | Back: deconv (from conv5_26, disp raw) | Boost: 0/1 | FPS: 0.8

Take-Aways

- user confidence is key
- safety within the framework of defined concept of operations
- just follow standards
- open discussions in the industry, acceptance takes time
- build trust through understanding
- interpretability
- tools are available, e.g. LIME or Deep Vis Toolkit
- open the black box



HEAD OFFICE

60 Camelot Drive
Ottawa, Ontario K2G 5W6

PHONE 613 686 3988
TOLL FREE 1 866 799 1555
EMAIL info@searidgetech.com

