Enhancing Semantic Segmentation of Aerial Images with Inhibitory Neurons

Ihsan Ullah, Sean Reilly, Michael G Madden

Machine Learning & Data Mining Research Group
School of Computer Science, National University of Ireland Galway

25th International Conference on Pattern Recognition,
Milan/Online, 10-15 January 2021

OÉ Gaillimh
NUI Galway

ICPR 2020

ROCSAFE
Remote Forensics

Outline of Talk

Our Contributions in Brief

Task: Semantic Segmentation

Context: ROCSAFE Project

Objective: Better Semantic Segmentation
while Re-using Existing Models

Solution: Inhibitory Convolutional Block

Results

Conclusions

Our Contributions in Brief

- Inhibitory Neurons first proposed in 2003
  - Inspired by neuro-biological research
  - Well suited to detecting contours and texture changes
- Propose new **Inhibitory Convolutional Block (ICB)**
  - Updated form of inhibitory neuron
  - Works in modern deep learning frameworks
- ICB Enables lateral inhibition to be added to **existing DNN models**
  - No need to retrain from scratch, can just fine-tune
  - Fast and efficient for model re-use
- Applications in object detection and semantic segmentation
  - State-of-art results on Aeroscape dataset for semantic segmentation:
  - 13.4% better than a similar model without ICB

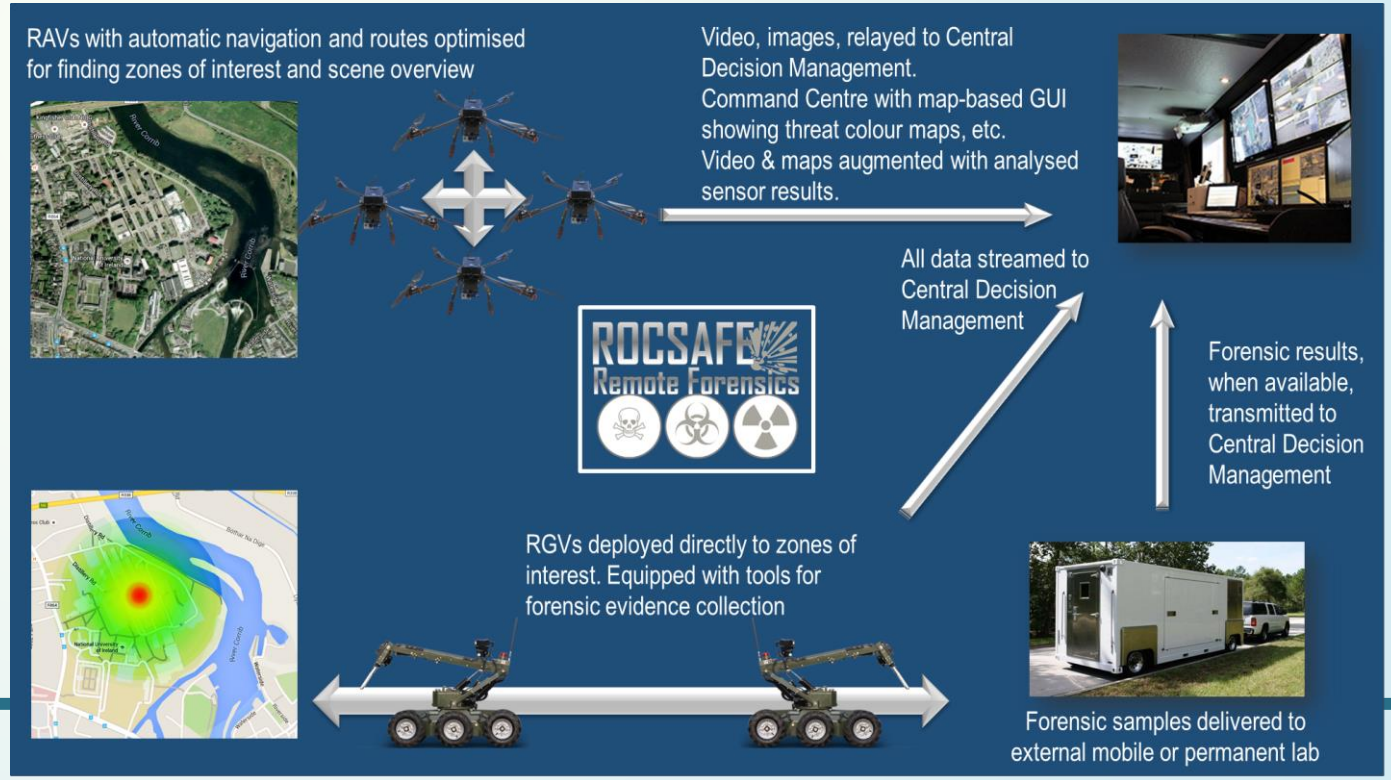
Context: ROCSAFE Project

Remotely Operated CBRNe Scene Assessment and Forensic Examination

- Horizon 2020 Research & Innovation Project
- 13 partners across Europe, led by NUI Galway

Innovations in Robotics, Sensors and AI to keep humans safe in potentially dangerous situations

- Assess the scene to quantify threats
- Rather than sending crime scene investigators into hazardous situations, send in drones and other robotic systems with appropriate sensors and AI



Task: Pixel-Level Semantic Segmentation

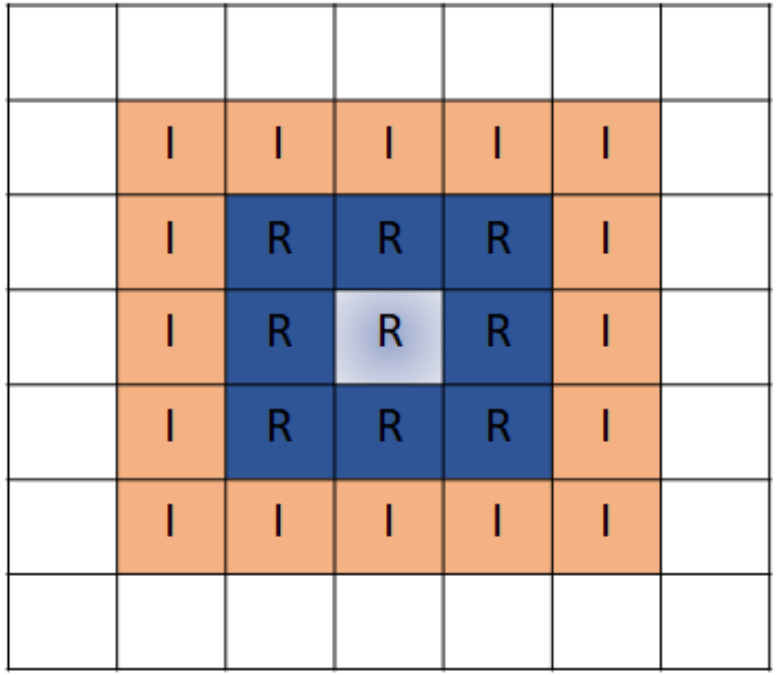
Goal: assign a colour to every pixel in an image, based on what that pixel represents

Challenge and Objectives

- The challenge:
  - Deep models like VGG and ResNet have millions of parameters
  - Semantic Segmentation needs large amount of labelled data to train: **for hazardous crime scene situations, such data is hard to obtain**
- Fine-tuning and transfer learning can help
  - Make use of models that were trained on less specialized datasets
- Therefore, our objectives are:
  - Develop an approach to semantic segmentation that works better than current state of the art
  - Make use of existing deep models
  - Improve speed

Background: Lateral Inhibition

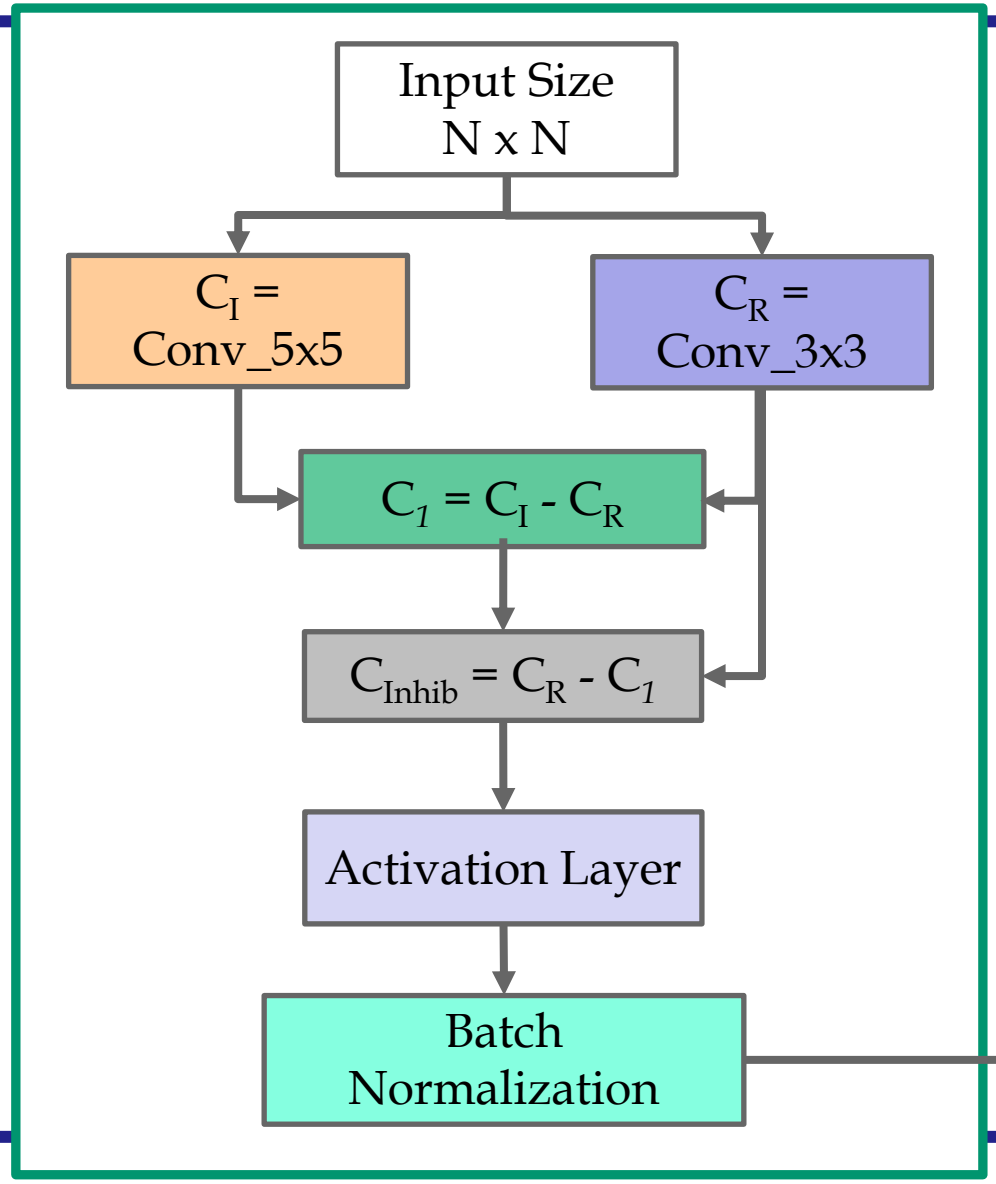
- Lateral Inhibition:
  - Capacity of an excited neuron to reduce the activity of its neighbours [1-3]
  - Excited neurons are called inhibitory neurons
  - Located immediately outside the receptive field of a neuron
- Result: an inhibitory field surrounding the receptive field [4]
  - Can improve texture analysis by ignoring the inner texture contours

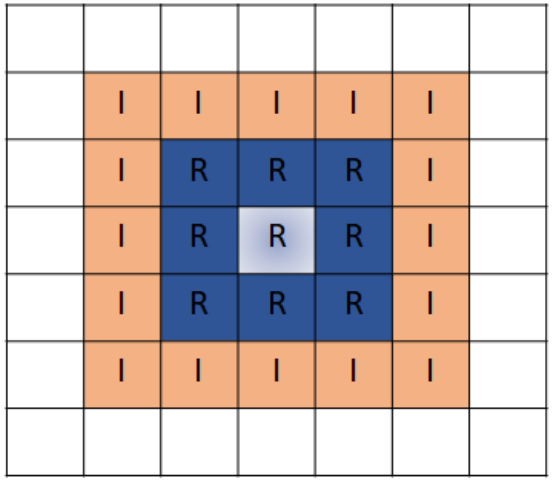


Related Work

- LIPNET [4-5]
  - lateral inhibition used in a neural network for image classification and segmentation
  - Bottom-up approach that removes the neighbouring neuron's impact by **subtracting** it
- Others
  - Contour detection [1] - 2003
  - Image segmentation [4] - 2008
  - Face detection [5] - 2013
  - Recently: deep CNN for saliency detection [6] - 2018

Proposed Inhibitory Convolutional Block (ICB)

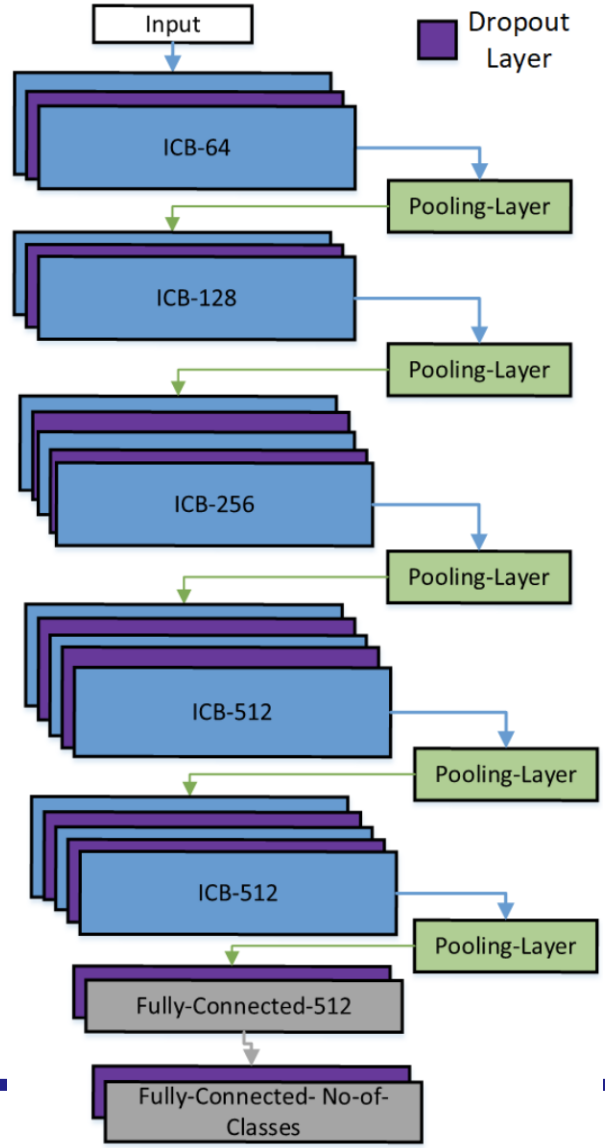




Inhibitory Convolutional Block, Cont'd

Enhanced inhibitory-VGG16 model (ICB-VGG16)

Applicable on all models where we have a Convolutional Layer



Experiments: Benchmark Datasets

Object Classification

- Models enhanced: VGG16 [7] and ResNet [8]
- Datasets: CIFAR-10 & CIFAR-100 datasets
  - 50,000 training and 10,000 testing images
  - Augmented with horizontal flipping, rotation ±15°, vertical/horizontal shifts 10%.

Semantic Segmentation

- Models enhanced: VGG16 and ResNet
- Dataset: Aeroscape of aerial view images [9]
  - 3269 images, 12 categories
  - Taken from 5 - 50 meters above ground level
  - Augmentation: horizontally flipping

Results : CIFAR-10 Object Classification

Name of Model	Layers	Accuracy
BL-VGG16*	16	93.43
ICB-VGG16	**16**	**93.62**
BL-ResNet-BS32*	20	90.31
BL-ResNet-BS200*	20	90.60
ICB-ResNet-BS200*	**20**	**91.63**
Inhi-ResNet-DO-0.3	20	92.21
ICB-ResNet-DO-0.3-Frac	**20**	**92.91**
ResNet [8]	110	93.57

BL, ICB and DO represent Baseline, Inhibitory Conv Block, and Dropout Layer, respectively.
* Indicates baseline model trained by ourselves.

Results: CIFAR-100 Object Classification

Name of Model	Layers	Accuracy
BL-VGG16*	16	69.88
ICB-VGG16	**16**	**71.28**
BL-ResNet-Aver*	20	66.22
ResNet-DO-0.3-Aver	20	66.85
ICB-ResNet-DO-0.3-Aver	20	66.96
ICB-ResNet-DO-0.3-Frac	**20**	**69.48**

ICB, Aver, and Frac represent Inhibitory Conv Block model, Average, and Fractional, respectively

* Indicates a baseline model trained by ourselves

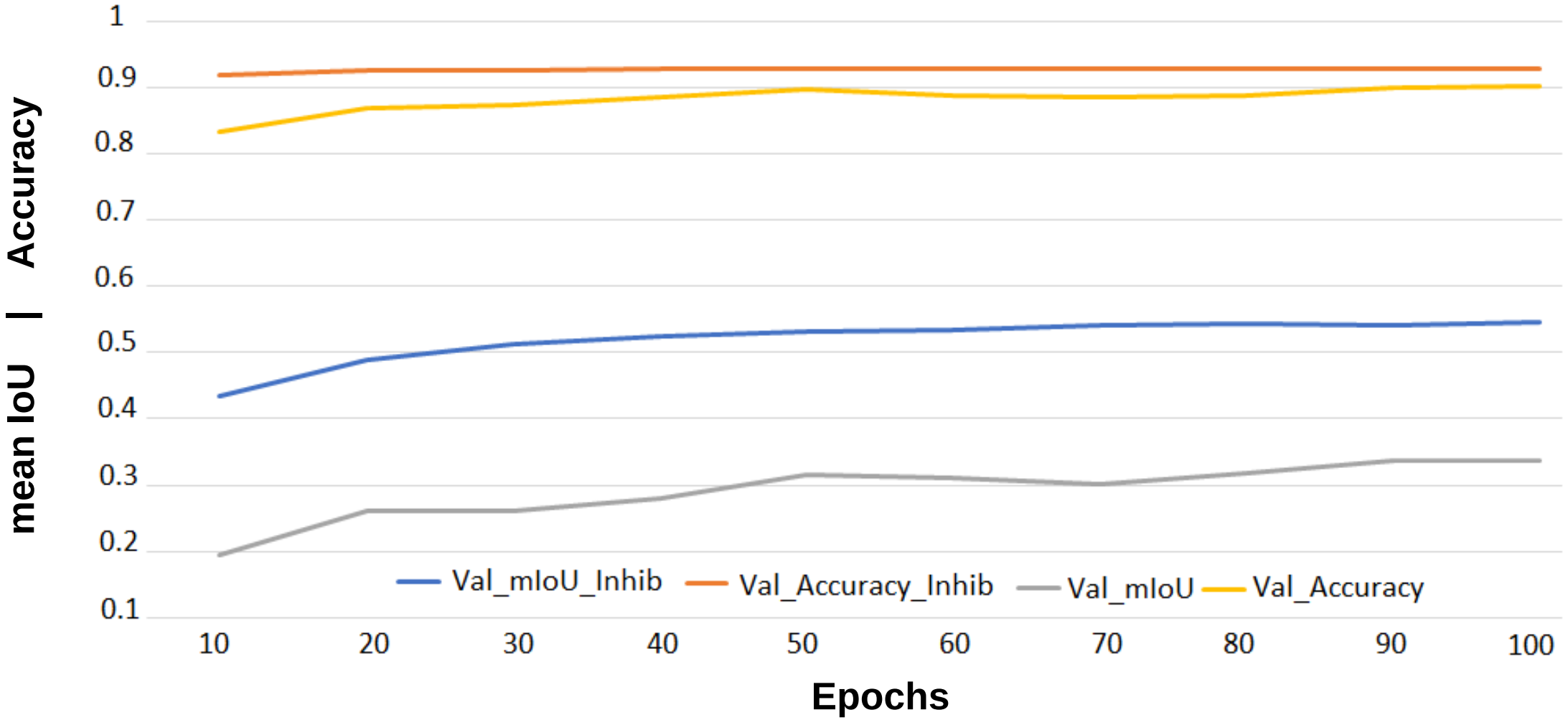
Results: Aeorscape Semantic Segmentation

Comparison of our inhibitory model vs published state-of-the-art

Model	Single Model/Ensemble	mIoU
FCN-Imagenet-8s [9]	Ensemble	47.56
SS-VGG16-FFT	Single	51.00
FCN-Ensemble-SingleSource [9]	Ensemble	53.08
FCN-Ensemble-MultiSource [9]	Ensemble	57.08
SS-IHB-VGG16-FFT	**Single**	**64.43**

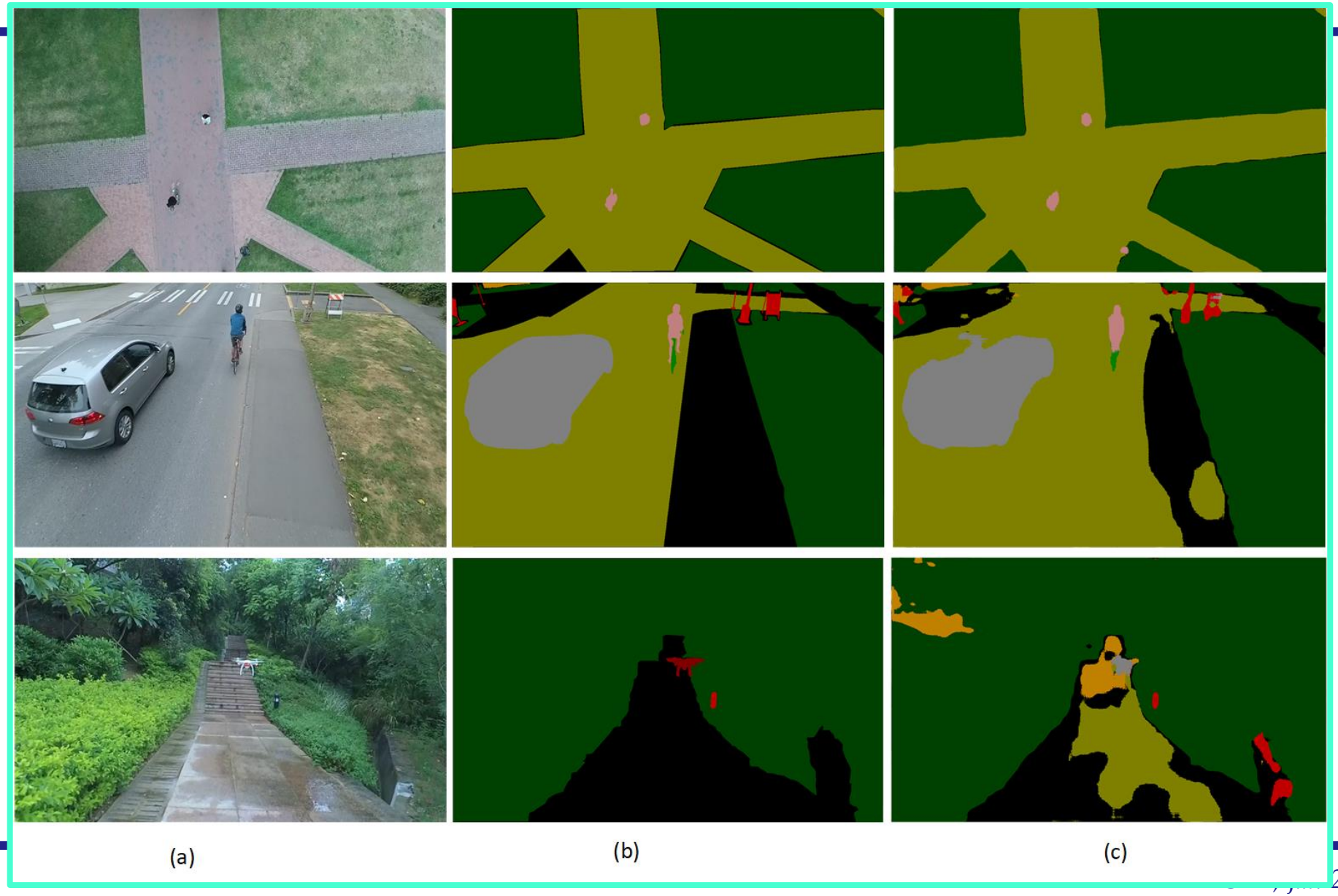
SS, IBH, FT1, and FFT stands for Semantic Segmentation, Inhibitory Conv Block, Fine Tuned on 1 layer), and Fully Fine Tuned (all layers)

Results: Aeorscape Semantic Segmentation



Results – Semantic Segmentation

a) Original Images
b) Ground Truth
c) Predicted



Conclusions

- We have proposed the Inhibitory Convolutional Block:
  a new approach for reducing the time to fine-tune deep learning models
  - Draws on prior work on receptive fields and inhibitory fields
  - General mechanism that can be applied to modify existing CNN architectures in order to introduce inhibitory neurons.
- Evaluation:
  - Strong results on object recognition datasets
  - State-of-the-art results for semantic segmentation of aerial images
- Key benefits:
  - Facilitates re-use of models that are implemented in Keras and Tensorflow
  - Fine-tuning for new applications is much faster than existing approaches.
  - Model re-use and fast learning both lead to lower GPU utilization, reducing energy consumption for the model training process, with consequent environmental benefits

Thank You

Prof. Michael Madden
Machine Learning & Data Mining Group
National University of Ireland Galway
michael.madden@nuigalway.ie
http://datamining.it.nuigalway.ie

We acknowledge funding from **European Union's Horizon 2020** Research and Innovation Programme, under Grant Agreement No. 700264.

References

[1] C. Grigorescu, N. Petkov, and M. A. Westen berg, “Contour detection based on nonclassical receptive field inhibition,” IEEE Transactions on Image Processing, vol. 12, no. 7, pp. 729–739, 2003

[2] T .A. Tjøstheim and C. Balkenius, “Cumulative inhibition in neural networks,” Cognitive Processing, vol. 20, no. 1, pp. 87–102, 2019

[3] C. Cao, Y. Huang, Z. Wang, L. Wang, N. Xu, and T. Tan, “Lateral Inhibition-Inspired Convolutional Neural Network for Visual Attention and Saliency Detection,” pp. 6690–6697, 2018

[4] B.J.T. Fernandes, G. D. C. Cavalcanti, and T. I. Ren, “Classification and segmentation of visual patterns based on receptive and inhibitory fields,” in2008 Eighth International Conference on Hybrid Intelligent Systems, 2008, pp. 126–131

[5] B.J.T. Fernandes, G. D. Cavalcanti, and T. I. Ren, “Lateral inhibition pyramidal neural network for image classification,” IEEE Transaction on Cybernetics, vol. 43, no. 6, pp. 2082–2092, 2013

[6] C. Cao, Y. Huang, Z. Wang, L. Wang, N. Xu, and T. Tan, “Lateral Inhibition-Inspired Convolutional Neural Network for Visual Attention and Saliency Detection,” pp. 6690–6697, 2018

[7] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” pp. 1–14, 2014.

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” vol. 9908 LNCS, pp. 630–645, 2016.

[9] I. Nigam, C. Huang, and D. Ramanan, “Ensemble Knowledge Transfer for Semantic Segmentation,” in IEEE Winter Conference on Applications of Computer Vision, WACV 2018, vol. 2018-Janua, 2018, pp. 1499–1508