Two Training Schools (linked to workshop with technology developers)

Tony Pridmore

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731013. This publication reflects only the view of the author, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

Document information

EU Project N°	731013	**Acronym**	EPPN2020
Full Title	European Plant Phenotyping Network 2020		
Project website	www.eppn2020.plant-phenotyping.eu		

Deliverable	**N°**	D5.4	**Title**	Two Training schools (linked to workshop with technology developers)
Work Package	**N°**	5	**Title**	Networking

Date of delivery	**Contractual**	31/10/2018 (Month 18)	**Actual**	14/11/2018 (Month 19)
Dissemination level	X	**PU Public, fully open, e.g. web**		
		CO Confidential, restricted under conditions set out in Model Grant Agreement		
		CI Classified, information as referred to in Commission Decision 2001/844/EC.		

Authors (Partner)	UNOTT			
Responsible Author	**Name**	Tony Pridmore	**Email**	**tony.pridmore@nottingham.ac.uk**

Version log			
Issue Date	**Revision N°**	**Author**	**Change**
14/11/2018	1	Uli Schurr	review by WP leader
15/11/2018	1	François Tardieu	Validated version
29/03/2019	2	Tony Pridmore	Addition of section 3 following the EC review of 1st reporting period

Executive Summary

This deliverable describes the first of two Training Schools operated as part of WP5 of EPPN2020. The first School, on Deep Machine Learning, was delivered at the School of Computer Science, University of Nottingham, on May 23rd 2018.

Objectives:

To provide an introduction to and understanding of the basic concepts of deep machine learning and particularly Convolutional Neural Nets (CNNs)

To provide practical experience of the use of these techniques

To foster discussion of the potential of CNNs for application in image-based plant phenotyping.

Rationale:

The target audience for the School was mixed, including both technology developers with experience of classical image analysis and machine learning methods and plant and crop scientists with limited knowledge of the computational sciences. It was therefore important that, while key concepts were introduced, the presentation should not be too technically detailed. The intention was that those with no prior knowledge of the area should leave with sufficient understanding to allow them to access the deep learning literature by themselves.

It was also important that these audience members gain some practical experience of the process of developing, training and applying a CNN. A significant proportion of the School was therefore given over to practical exercises, supported by staff from the Computer Vision Laboratory (CVL), University of Nottingham.

To allow participants with more advanced knowledge to develop further, CVL staff were available to answer more detailed technical questions and discuss participants' projects and problems.

Main Results:

A half-day Training School on Deep Machine Learning was held on May 23rd 2018 at the School of Computer Science, University of Nottingham. 14 participants attended.

Authors/Teams involved:

Prof. Tony Pridmore (UNOTT)
Dr Michael Pound (UNOTT)
Dr Reza Soltaninejad (UNOTT)
Dr John Atanbori (UNOTT)
Mr Aaron Jackson (UNOTT)

Table of contents

1. MOTIVATION

Deep machine learning approaches have revolutionized image analysis and computer vision in recent years, achieving performance on key image analysis tasks such as image segmentation and object recognition that significantly extend the state of the art.

The key difference between deep and previous “classic” machine learning methods lies in what is learnt. Classic methods require the user to decide which features of the images to be analysed are important. When attempting to decide e.g. whether a leaf is present in an image fragment, the user must first specify a set of features (leaf tip, veination, etc) upon which the decision should be made. Software tools must also be developed which can extract those features from the target images. The resulting set of ‘handcrafted’ image features are then input to a machine learning component which is trained to associate feature properties with e.g. the presence of absence of a leaf. Though classic methods can be powerful, there are significant drawbacks. The success of the system depends mainly on the features chosen and the performance of the feature extraction method. If inappropriate features are chosen it may simply be impossible for the learning element to learn a viable method.

Deep machine learning methods learn both how to achieve the task and what features are needed to make the task possible. Learning is “end-to-end”, the learning component is provided with input images and desired output and learns a complete solution. Deep methods are not a panacea: the problem is shifted from selecting image features to designing a deep method with an appropriate structure. In many cases, however, this is an easier task. Deep learning architectures designed for one domain also often transfer to similar tasks in other application areas.

Given the high proportion of phenotyping tasks and installations that include some element of digital image analysis, it is important that the European Plant Phenotyping community is both aware of the potential of deep learned approaches and able to make effective use of them.

2. A TRAINING SCHOOL ON DEEP MACHINE LEARNING

The EPPN2020 training school on deep machine learning took place on Wednesday 2rd May 2018 in the main teaching laboratory of the School of Computer Science Univerity of Nottingham. The School was advertised first within the EPPN2020 consortium, and then externally. 14 participants attended.

The School began with an introductory lecture (45 mins) by Prof Tony Pridmore. This introductory lecture presented the key terms and ideas underlying deep learning and convolutional neural nets - image processing by convolution, classic neural net structure, loss functions and learning – before showing how they are combined in CNNs. The structure of early CNNS, designed to perform classification tasks, provided illustrative examples of the practical application of deep methods.

This was followed by a practical session (45 mins) lead by Dr Michael Pound. The session was built around the MNIST text recognition dataset, chosen because the task is familiar to a wide audience and the data widely used as a benchmark in machine learning. The participants first worked through exercises demonstrating the classic approach of feature detection and classification by a standard machine learning tool: the support vector machine (SVM). SVMs are among the most powerful classic methods and it was felt that increased awareness of this would be useful to the participants. More importantly, the practical session provided practical experience of both the process followed and the tools used when developing deep learned solutions.

Figure 1: Prof Tony Pridmore, leading the EPPN2020 training session in Nottingham

Following a networking coffee break during which CVL staff were available to discuss participants' specific experience and problems (30 mins), a short lecture by Prof Pridmore (30 mins) introduced some of the more recent CNN archtitectures, the tasks for which they had been designed, and discussed their strengths and weaknesses. The aim here was to provide some appreciation of the range of CNNs available, and provide pointers to literature expected to be of value to the phenotyping community.

The final session was given over to short describing the use of a variety of CNN architectures to address current phenotyping challenges, specifically:

- *Segmenting colour images of aeroponically grown cassava roots*: Dr John Atanbori
- *Segmenting 3D (volumetric) images*: Dr Reza Soltaninejad/Mr Ezenwoko Benson
- *Plant shoot feature detection*: Dr Michael Pound.

3. Review and Lessons learnt from this Training school 'Imaging and deep learning'

The aims of the training course were first to present the key terms and ideas underlying deep learning and convolutional neural nets - image processing by convolution, classic neural net structure, loss functions and learning - before showing how they are combined in CNNs. This was done to provide delegates with the basic understanding of the field needed to be able to interpret the large amounts of literature and many different CNN architectures currently being produced.
Then, we used industry standard development software to train, run and evaluate the performance of classic machine learning and CNN solutions to selected image classification

problems. This gave trainees practical experience of both the process followed and the tools used when developing deep learned solutions. This was done to improve the ability of biological scientists to engage with the computer scientists and engineers who develop CNN solutions. EPPN2020 colleagues had reported in previous meetings that CNNs were seen within the community as a 'black box' solution; this needs to change if the potential of deep learning in phenotyping is to be realized.
Overall, trainees could see, in practical terms, how deep learning changes the task of developing image analysis solutions. Rather than write detailed code specifying how an image should be processed, developers now design a suitable network architecture and training scheme. Finally, trainees had a general discussion over short "real life" presentations on the use of CNNs to address current phenotyping challenges, namely segmenting colour images of roots, segmenting 3D volumetric images and detecting plant shoot features.
In these discussions, trainees were able to ask questions of the course team that focused on particular aspects of both CNN architectures and the development process; this would not have been possible before the training session. More detailed conversations around a specific application have continued since the training session between Dr Pound and the group from Louvain. This may or may not have happened without the training session, but certainly proceeded more smoothly, because of improved mutual understanding, after it.

Overall, the participants gave a very positive feedback of this training course: the fact that it was adapted to people with no background in imaging as well as the practical aspect of it (concrete case studies) were the two points most appreciated. The presentations and discussion about real life situations allowed trainees to see the advantages and limitations of the use of deep learning in image analysis, and several delegates requested instructions on how to obtain the tools used, and reported that they now felt able to install and experiment with them at their own institutions.

The following points were derived from the training discussions to benefit EPPN2020. (i) Adapt the current workflows for using CNNs, but also interact with JRA1 leaders to adapt current CNNs to specific problems, (ii) Develop ways for tracing the novel methods of imaging into organized workflows, able to be included and shared via the information system under development in the JRA3.

The whole process of workflows will be the occasion of a joint work of EPPN2020 and EMPHASIS-prep for developing a service to the phenotyping community, which will summarise examples of good practices in data collection and image analysis.

4. Participant List

Last Name	First Name	Your email	Status
Dhondt	Stijn	stdho@psb.vib-ugent.be	Post-Doc
Draye	Xavier	xavier.draye@uclouvain.be	PI
Georgii	Elisabeth	elisabeth.georgii@helmholtz-muenchen.de	Bioinformatician
Pavicic	Mirko	mirko.pavicic@helsinki.fi	PhD
Summerer	Stephan	stephan.summerer@alsia.it	PhD
GINER	Jean-Baptiste	jbginer@asa-sas.com	MSc
Rincent	Renaud	renaud.rincent@inra.fr	PhD
Passot	Sixtine	sixtine.passot@uclouvain.be	Post-Doc
Tardieu	Francois	francois.tardieu@inra.fr	Senior scientist
Zivcak	Marek	marek.zivcak@uniag.sk	Assistant Professor, Researcher

Brestic	Marian	marian.brestic@uniag.sk	Professor
Alexandersson	Erik	erik.alexandersson@slu.se	Associate Professor
Couvreur	Valentin	valentin.couvreur@uclouvain.be	Post-Doc
Millet	Emilie	emilie.millet@wur.nl	Post-Doc
ALTAZIN	Thomas	taltazin@asa-sas.com	PhD
Schurr	Ulrich	u.schurr@fz-juelich.de	Prof.

5. ANNEXES: AGENDA AND PRESENTATIONS

EPPN Deep Learning Training Session - Wednesday 23rd May 2018

Prof Tony Pridmore and Dr Michael Pound

Room A32, School of Computer Science, University of Nottingham
Jubilee Campus, Wollaton Road
Nottingham NG8 1BB

Programme

14.00 **Introduction to Deep Learning and Convolutional Neural Nets (CNNs)**

The aim of this introductory lecture is to present the key terms and ideas underlying deep learning and convolutional neural nets - image processing by convolution, classic neural net structure, loss functions and learning – before showing how they are combined in CNNs. The structure of early CNNS, designed to perform classification tasks, will provide illustrative examples.

14.45 **Practical session – train and deploy a CNN**

Here we will use industry standard development software to train, run and evaluate the performance of classic machine learning and CNN solutions to selected image classification problems. This will give some practical experience of both the process followed and the tools used when developing deep learned solutions.

15.30 Coffee/Tea

16.00 **Alternative CNN Architectures**

Deep learning fundamentally changes the task of developing image analysis solutions. Rather than write detailed code specifying how an image should be processed, developers must design a suitable network architecture and training scheme. This short lecture will introduce some of the more recent archtitectures and discuss their strengths and weaknesses.

16.30 **Plant Phenotyping Case Studies**

Short presentations describing the use of a variety of CNN architectures to address current phenotyping challenges:

Segmenting colour images of aeroponically grown cassava roots: Dr John Atanbori
Segmenting 3D (volumetric) images: Dr Reza Soltaninejad/Mr Ezenwoko Benson
Plant shoot feature detection: Dr Michael Pound

17.30 Close.

EPPN Deep Learning Training Session - Wednesday 23rd May 2018

Room A32, School of Computer Science, University of Nottingham

Tony Pridmore & Mike Pound

14.00 Welcome & Lecture: Introduction to Deep Learning and Convolutional Neural Nets (CNNs)

14.45 Practical: Train and deploy classic ML and a CNN

15.30 Coffee/Tea

16.00 Lecture: Alternative CNN Architectures

16.30 Plant Phenotyping Case Studies

Segmenting colour images of aeroponically grown cassava roots: John Atanbori

Segmenting 3D (volumetric) images: Reza Soltaninejad/Ezenwoko Benson

Plant shoot feature detection: Mike Pound

17.30 Close

European Plant Phenotyping Network

EPPN 2020

Deep Learning: An Introduction

Introduction to Deep Learning and Convolutional Neural Nets

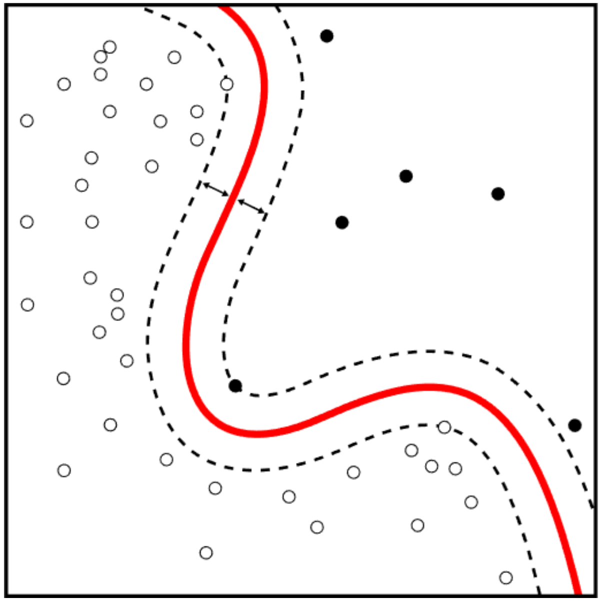
Tony Pridmore

Michael Pound

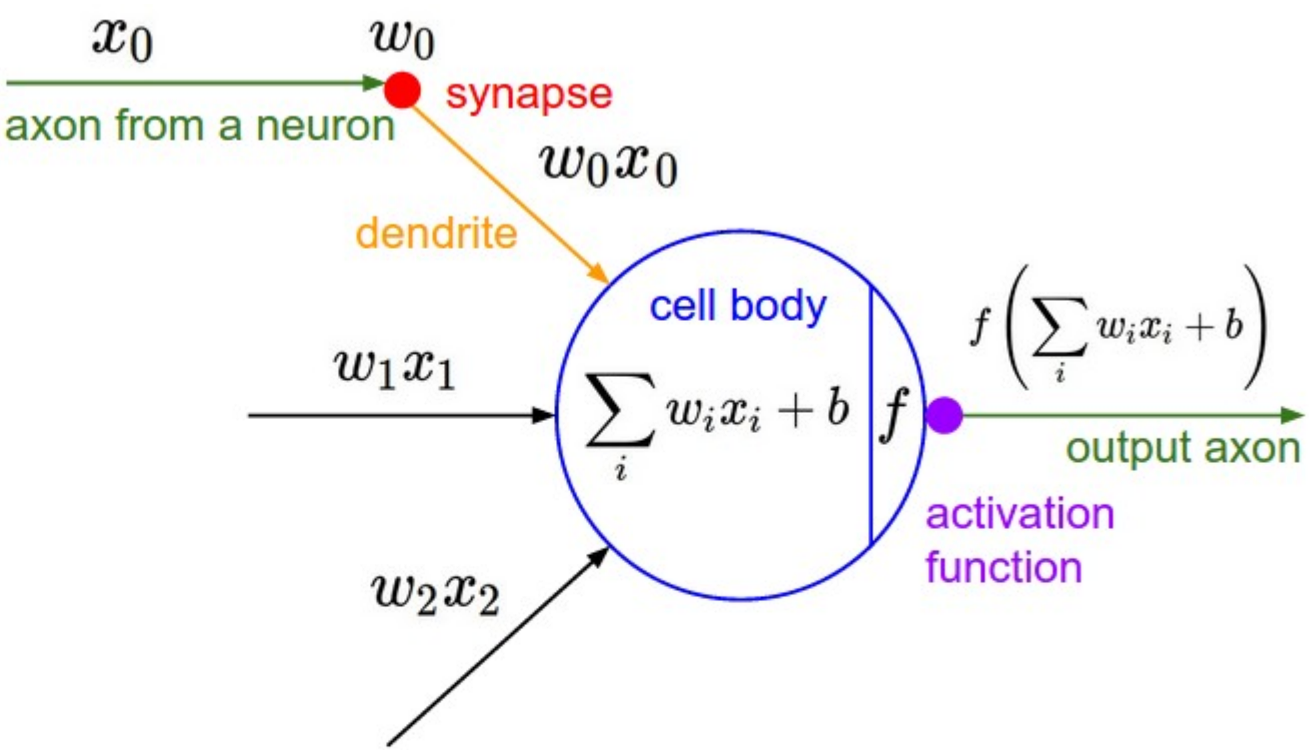
Machine Learning

"The construction of algorithms that can learn from and make predictions on data"

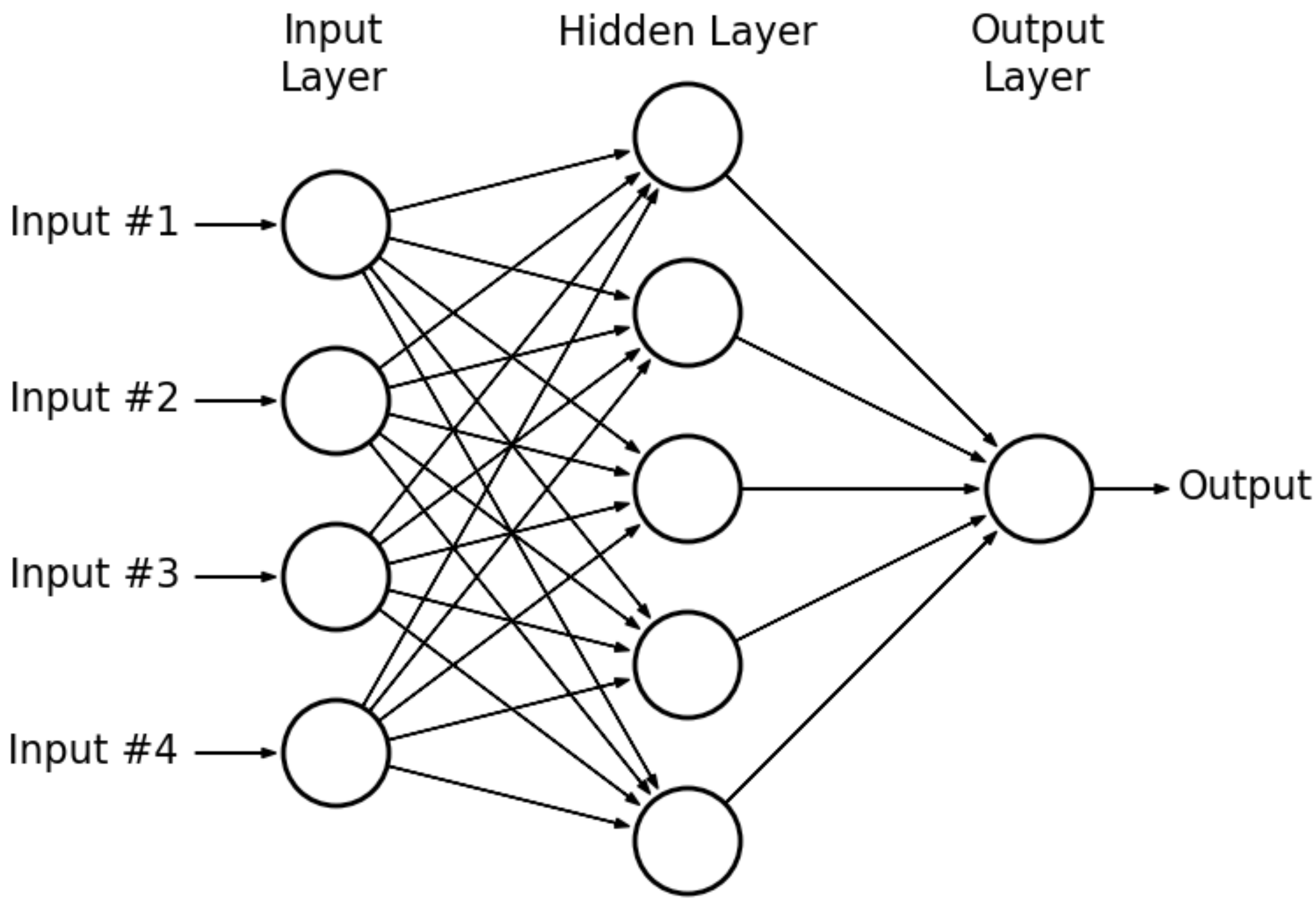
- When we talk about modern AI, we're usually referring to Machine Learning



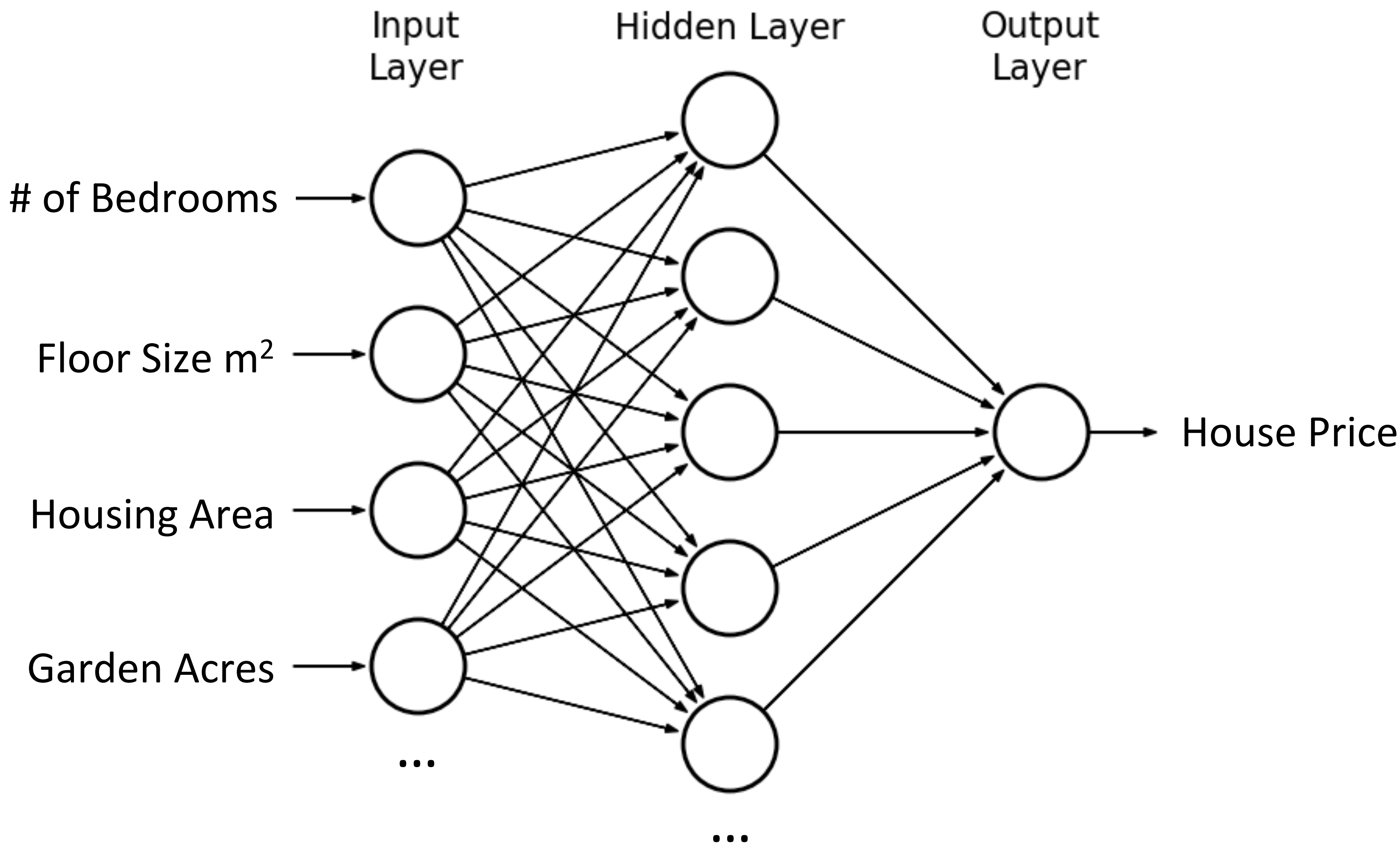
Neurons


x_0
w_0
synapse
axon from a neuron
$w_0 x_0$
dendrite
cell body
$w_1 x_1$
$\sum_i w_i x_i + b$
f
$f\left(\sum_i w_i x_i + b\right)$
output axon
activation
function
$w_2 x_2$

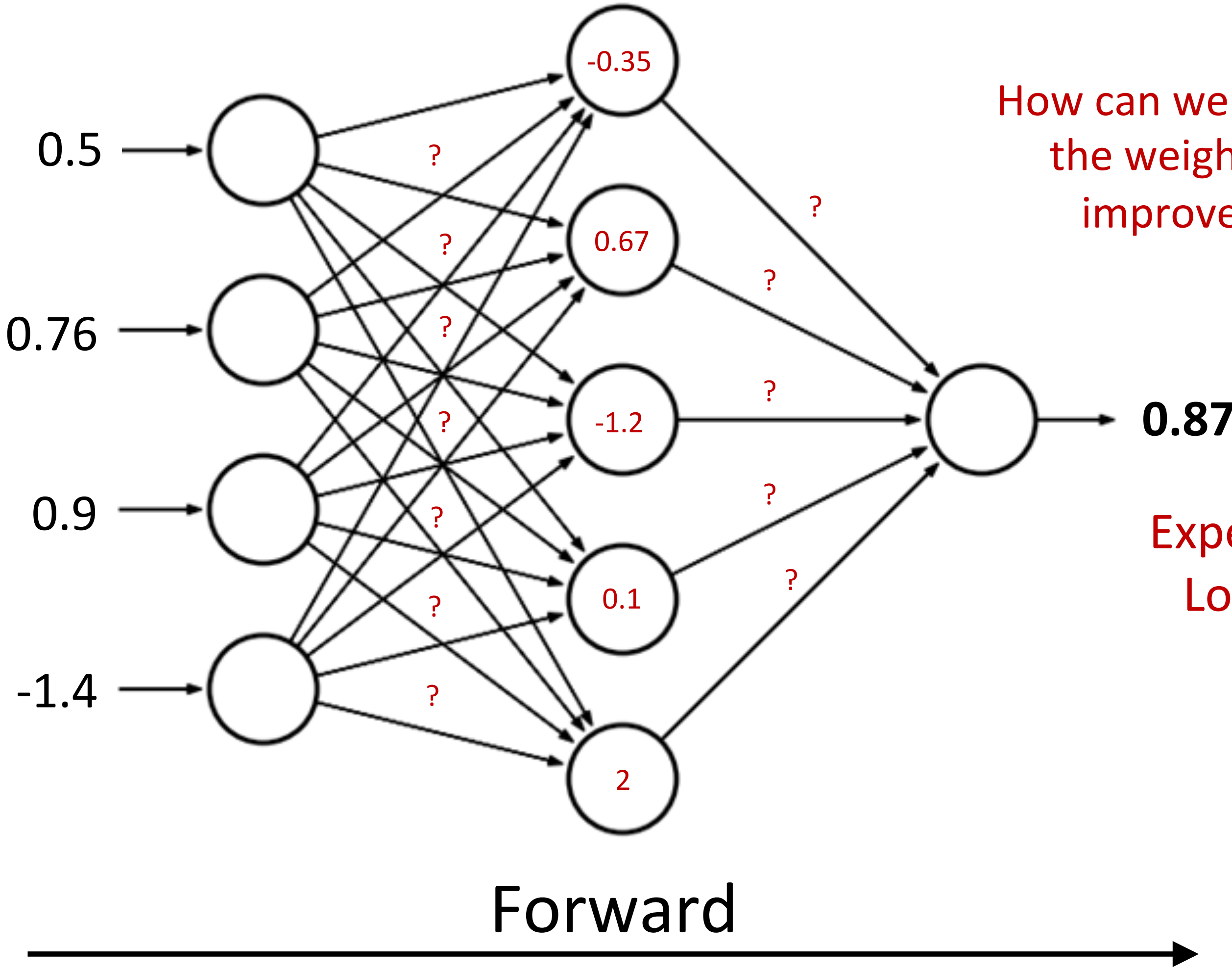
Neural Networks



House Prices



Training

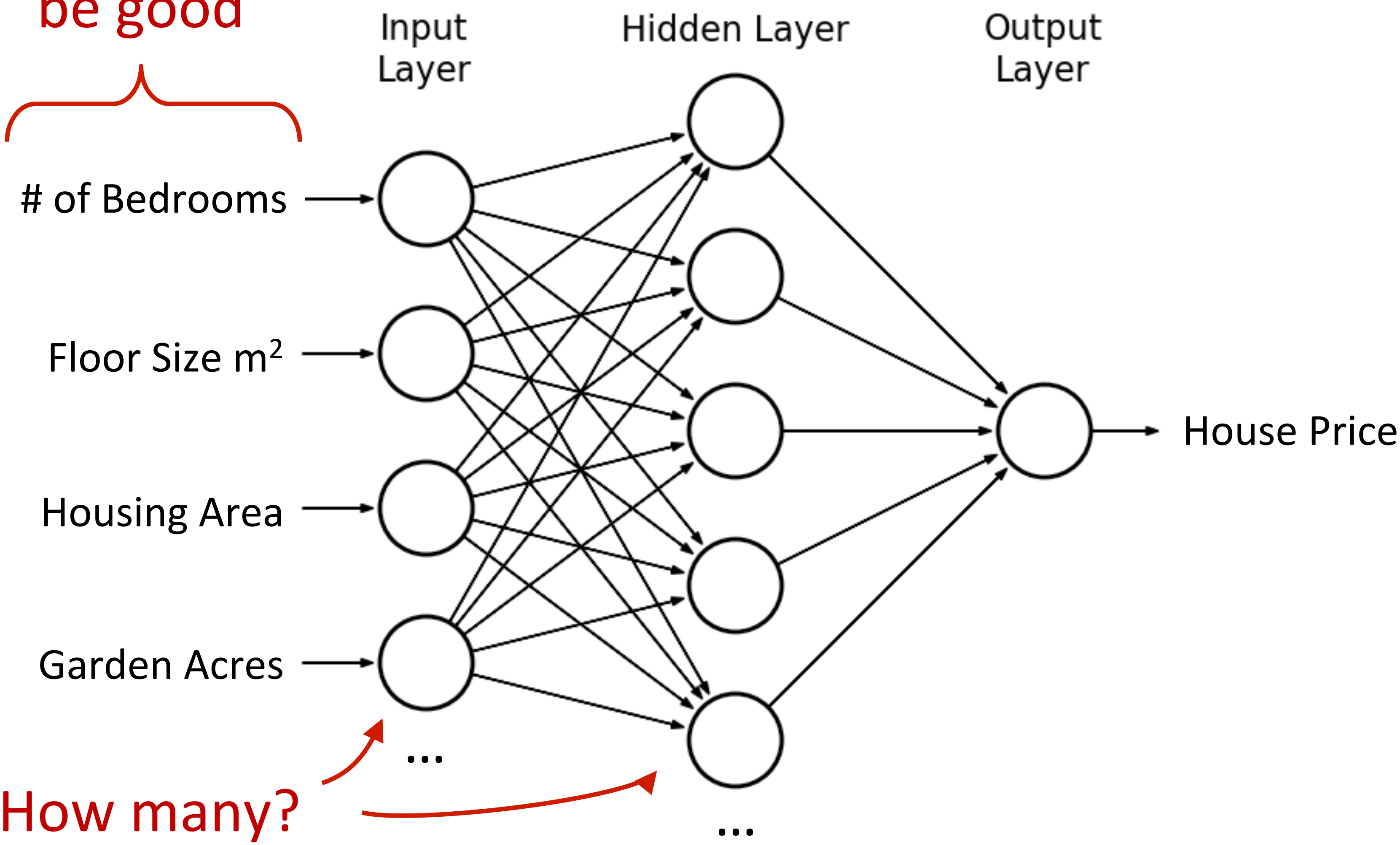


How can we change the weights to improve **L**?

Expected **0.74**
Loss **L** = **0.13**

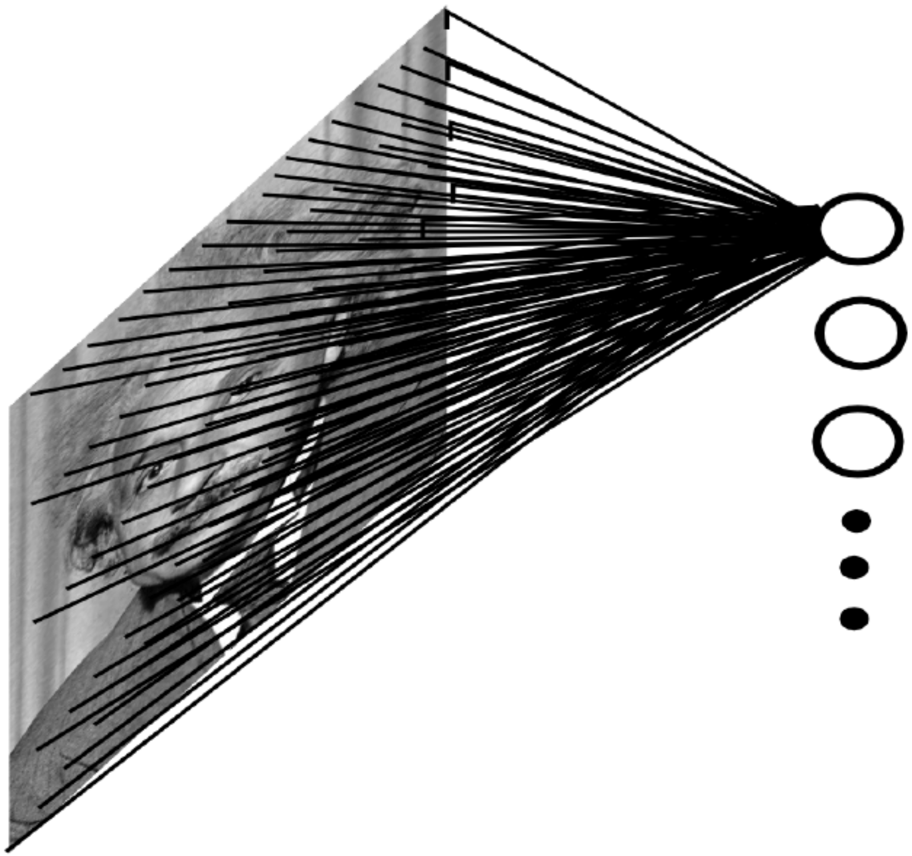
Some Problems

These need to be good



How many?

Working with images

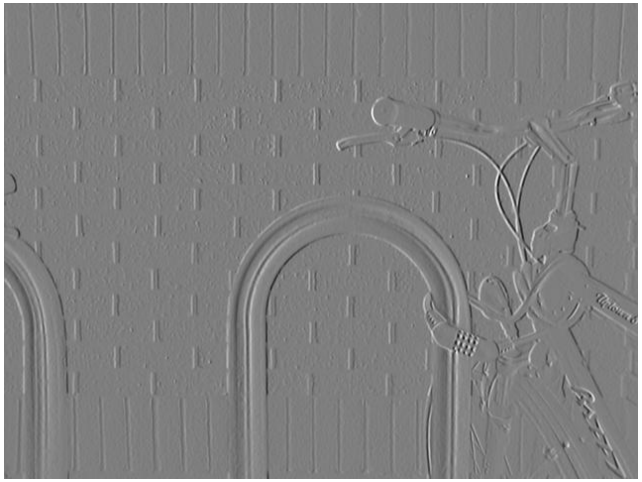


- Traditional neural nets are fully connected
- 200x200 image, 40K hidden units (1 per pixel) means ~2B weights to learn
- Would require an impractically large training set
- Need to extract features from images before learning

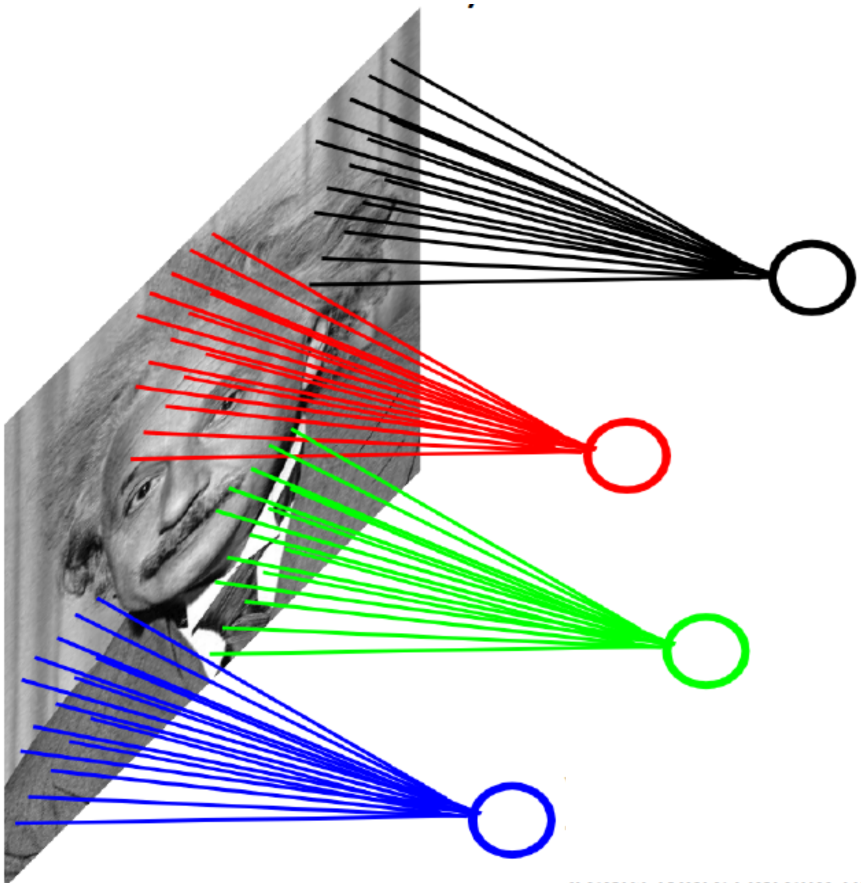
Neural Networks and Convolution

- Image processing and feature detection can be achieved via convolution
- Processed images contain weighted sums of small sub-areas of the original

-1	0	1
-2	0	2
-1	0	1

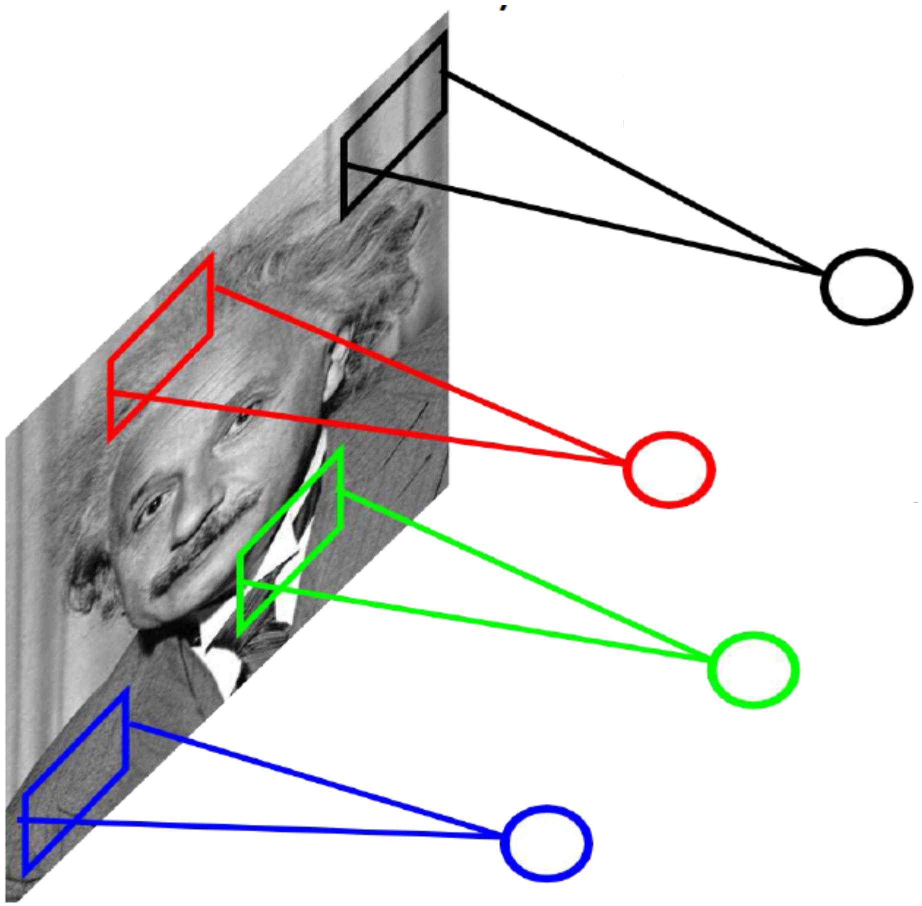


But neurons can do convolution



- Locally connected
- e.g. 200x200 image, 40K hidden units, 10 x 10 filter means 4M weights to learn

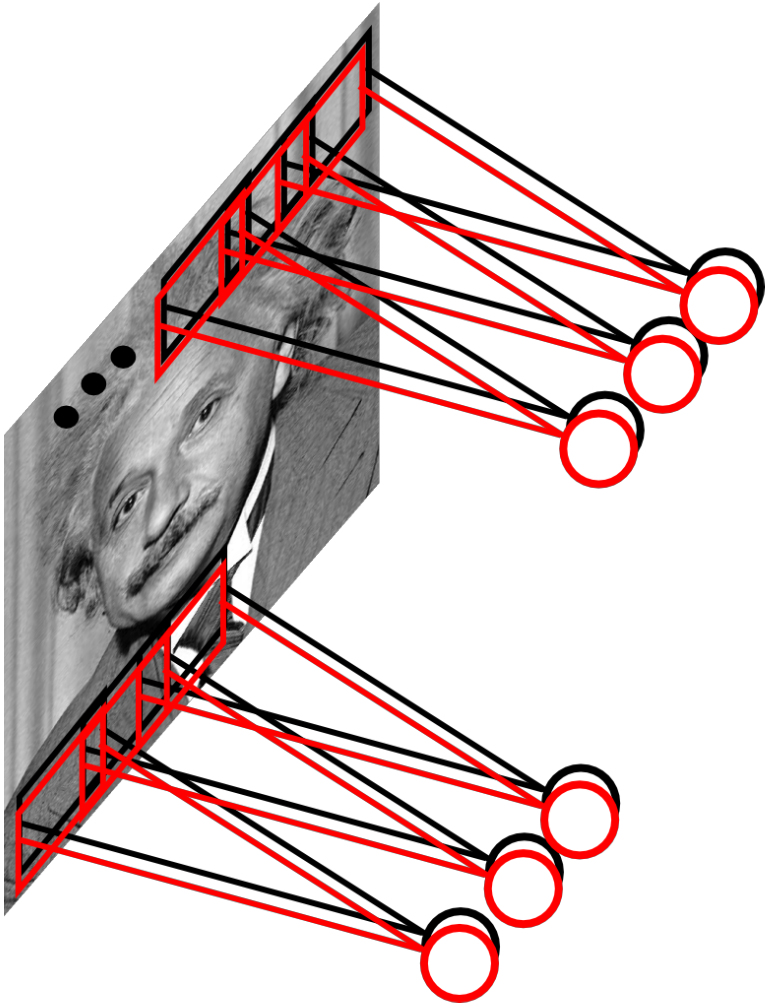
Convolution Layers



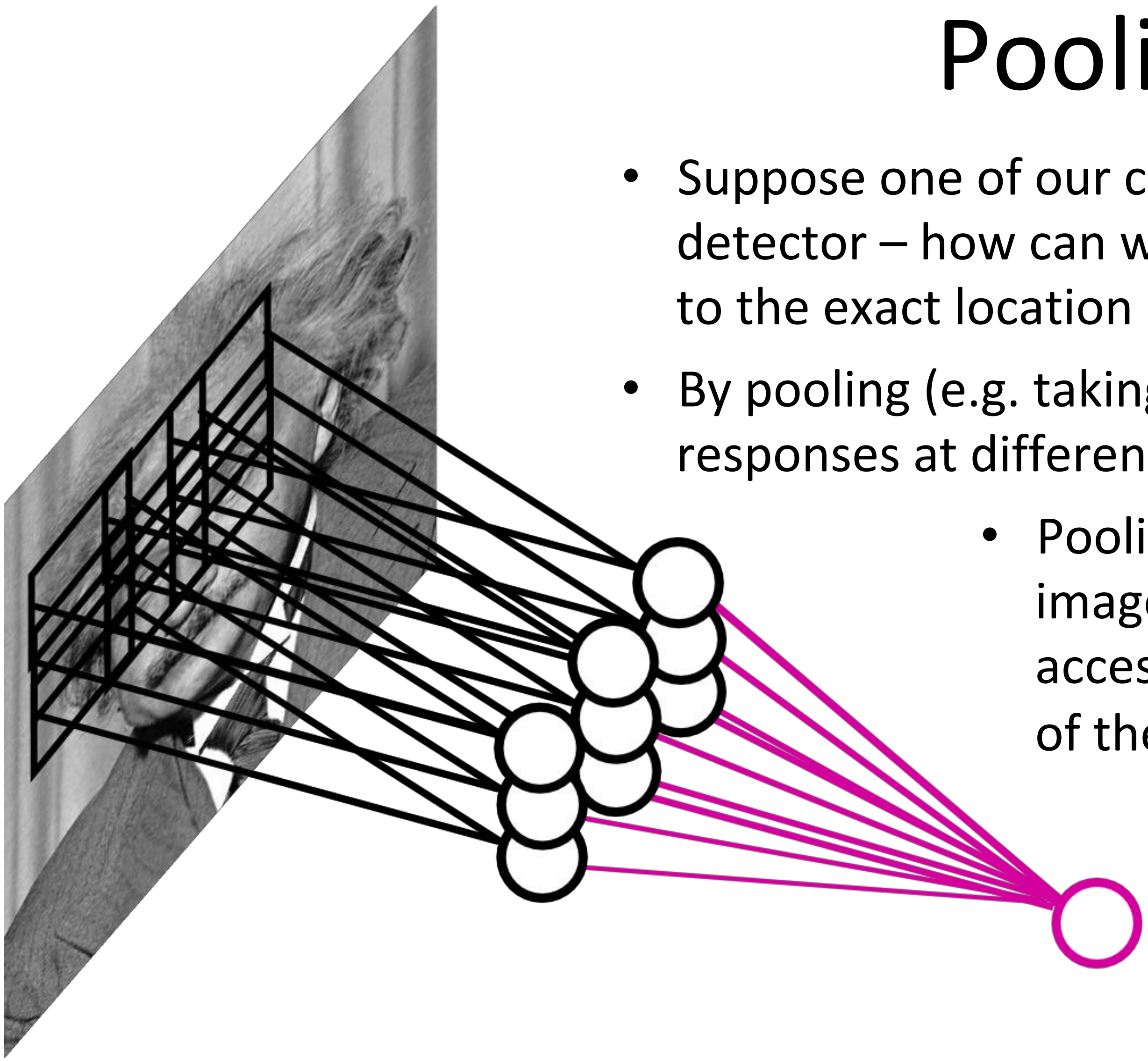
- In image processing/ vision we usually want to apply the same convolution mask at each location
- Each neuron in a layer has the same weights
- e.g. 200x200 image, 40K hidden units, 10 x 10 mask means only 100 weights to learn

Convolution Layers

- We can afford to learn multiple filters
  - e.g. 100 10x10 masks is only 10K parameters
- Convolutional layers are filter banks performing convolutions with learned kernels (masks)
- Could be applied to all pixels, or have a small 'stride'



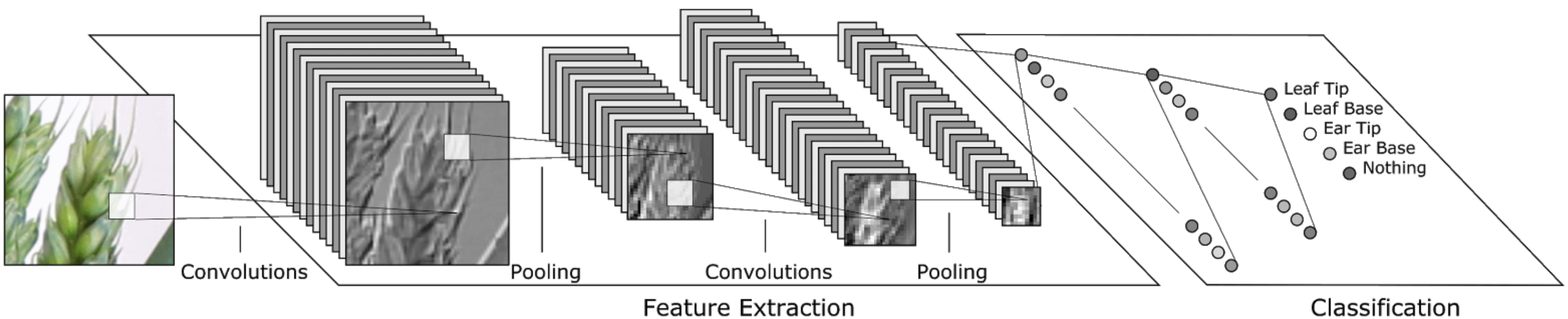
Pooling Layers



- Suppose one of our convolutions is an eye detector – how can we make the net robust to the exact location of the eye?
- By pooling (e.g. taking the max) filter responses at different locations
  - Pooling also shrinks the image, so later filters access larger subsections of the data

Convolutional Neural Networks

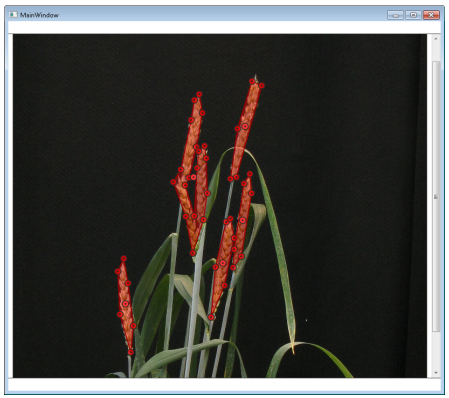
- Early layers perform convolutions
- Final layers perform classification
- All weights & convolution masks are learned



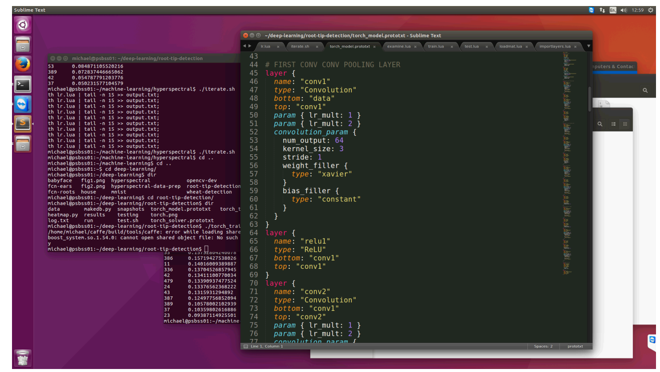
Convolutional Neural Network

How this works in practice

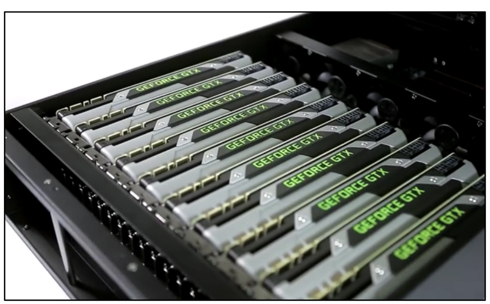
1. Capture and annotate dataset



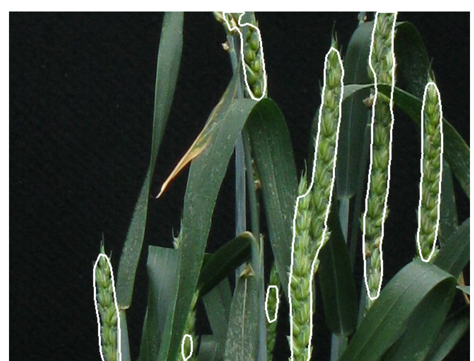
2. Design network architecture



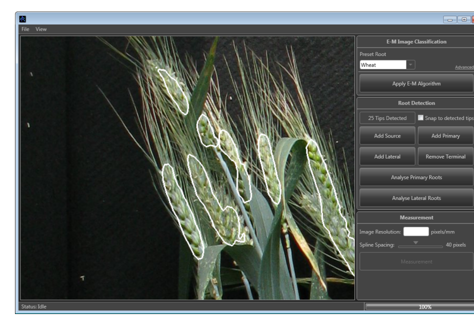
3. Train Network



4. Test Network

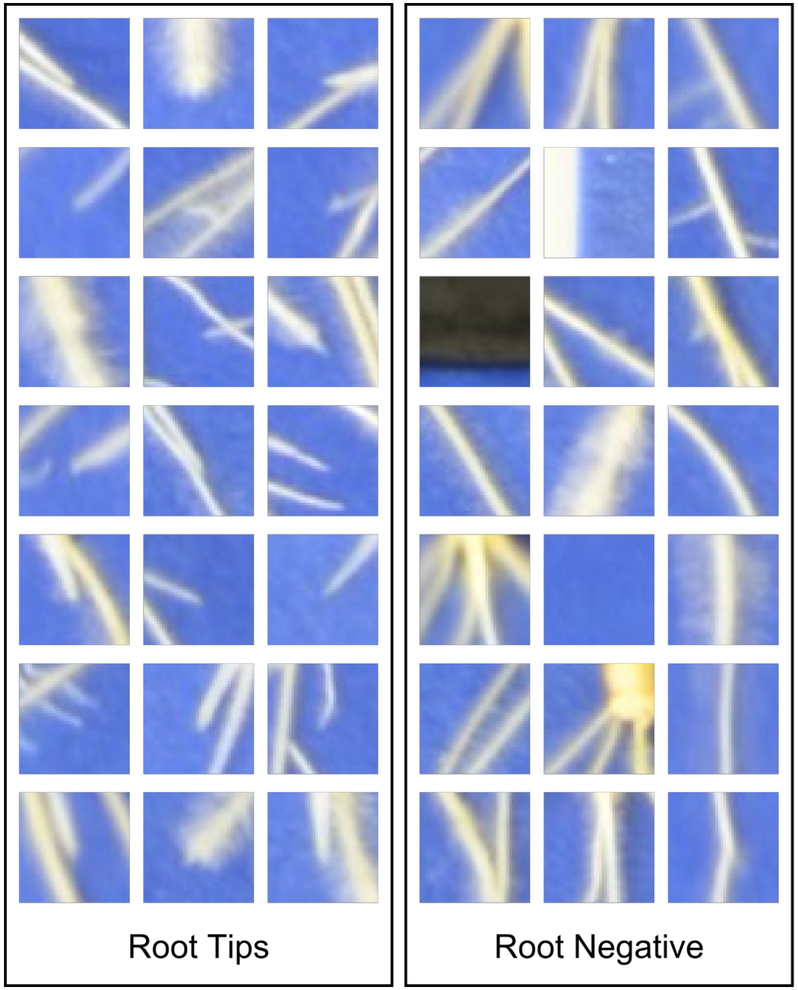


5. Deploy



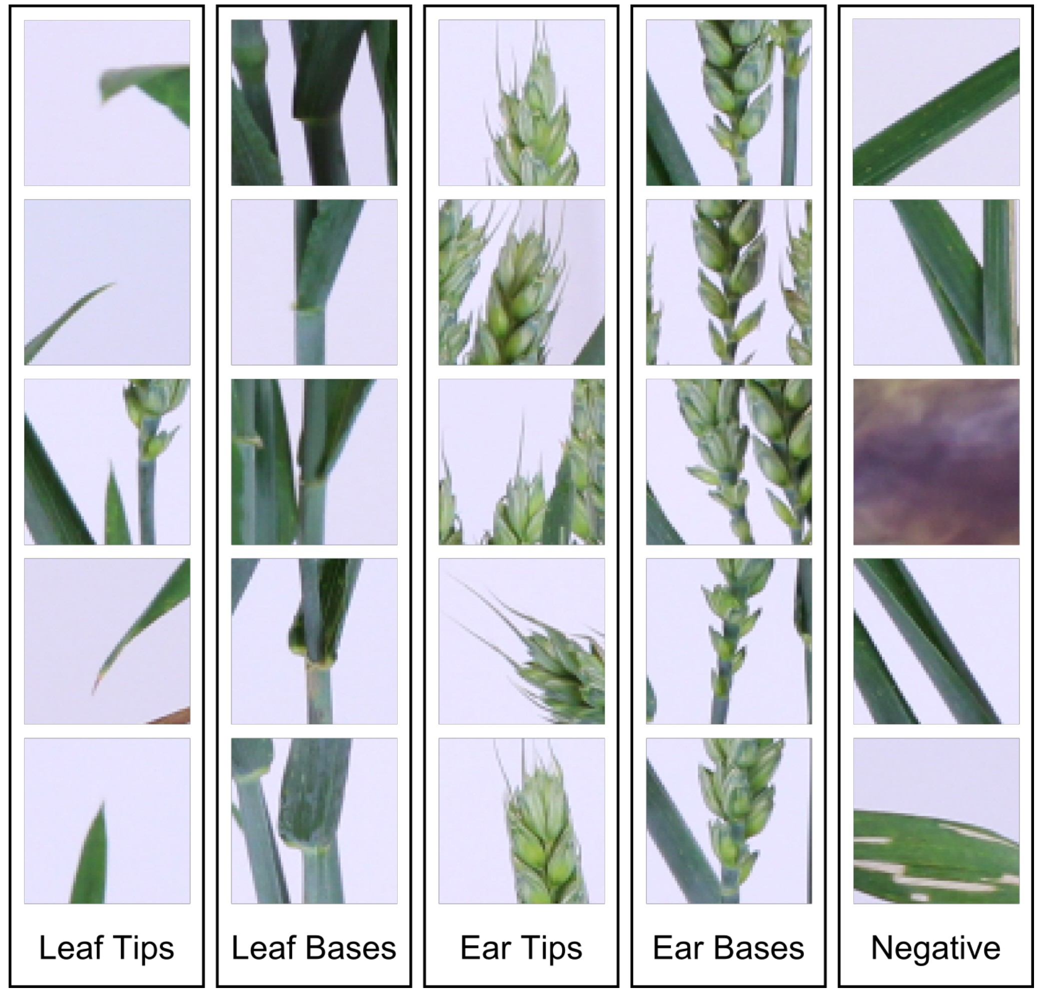
Root Feature Detection

- Images of 32x32 pixels were used
- Positive examples from each annotated root tip
- Negative both at random, and specifically on non-tip root
- ~44k images total (35k training, 9k validation)



Shoot Feature Detection

- Images of 64x64 pixels were used
- Positive examples from annotations
- Negatives generated at Harris feature locations
- ~62k images total (50k training, 12k validation)



Classification Results

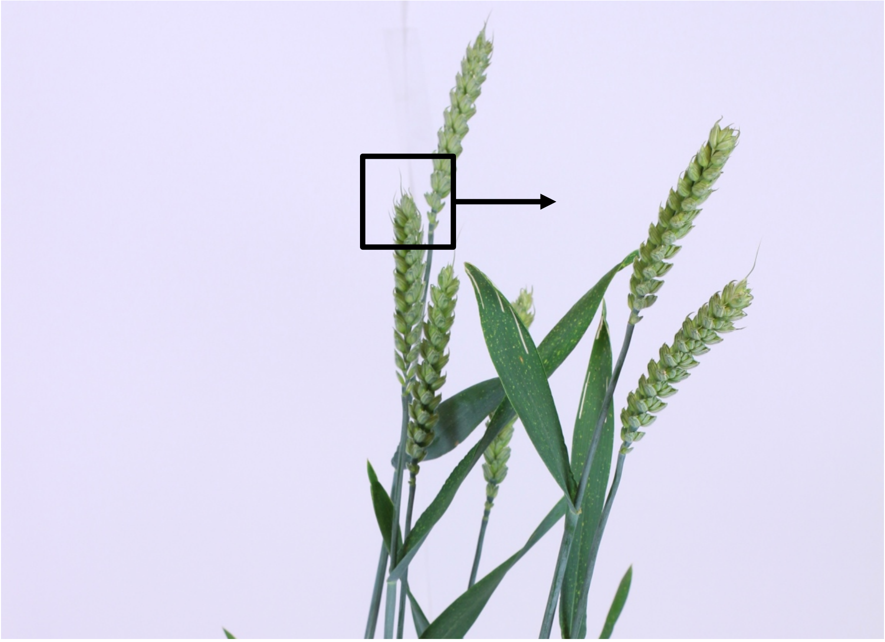
- Accuracy was measured on an unseen validation set
- Average accuracies of 98.4% and 97.3%

Feature	Accuracy (%)
Root Tip	97.5
Root Negative	98.9
Total/Average	**98.4**

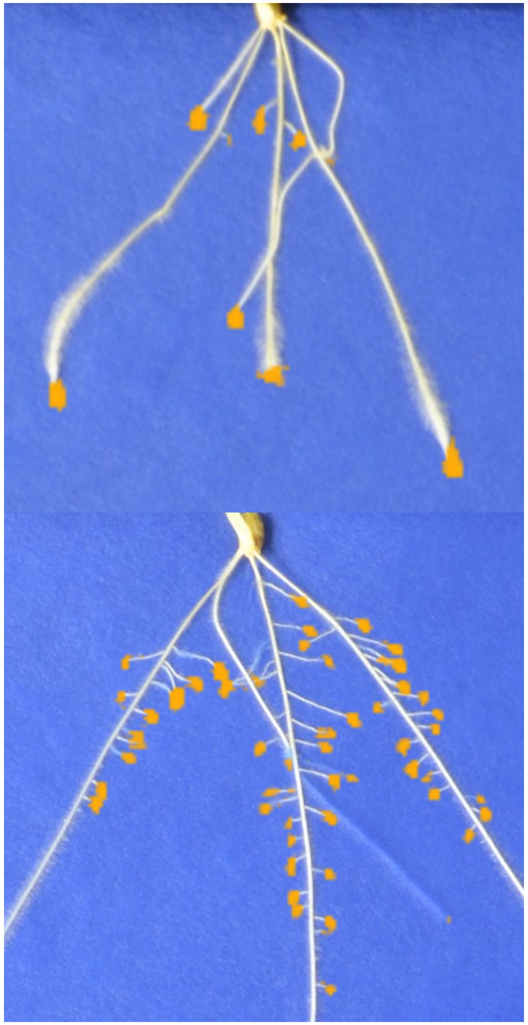
Feature	Accuracy (%)
Leaf Tip	95.2
Leaf Base	97.8
Ear Tip	97.9
Ear Base	97.1
Shoot Negative	97.8
Total/Average	**97.3**

Localisation

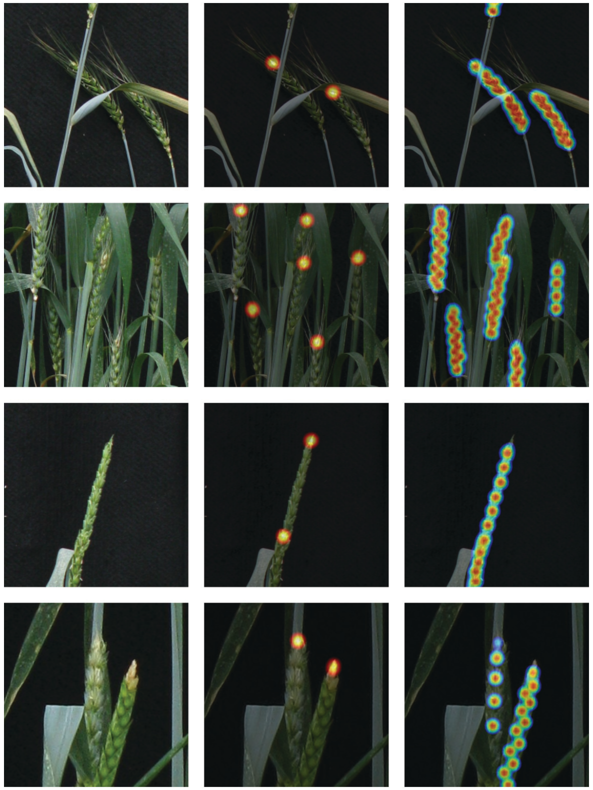
- We can scan an image, classifying at regular intervals to perform localisation



Localisation Results



Conclusions



- CNNs are a supervised learning method
- Success requires well annotated datasets, and careful design of CNN architectures
  - Convolution layers
  - Pooling layers
  - Decision layers

- Pound, M., et al, Deep Machine Learning provides state-of-the-art performance in image-based plant phenotyping, GigaScience. 2017

Acknowledgments

- Jonathan Atkinson
- Adrian Bulat
- Marcus Griffiths
- Aaron Jackson
- Erik Murchie
- Alexandra Townsend
- Georgios Tzimiropoulos
- Darren Wells
- Michael Wilson

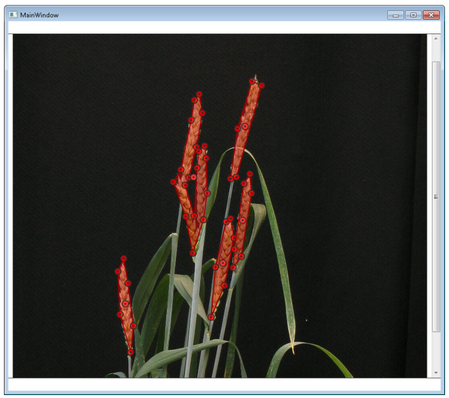
BBSRC
bioscience for the future

Alternative CNN Architectures

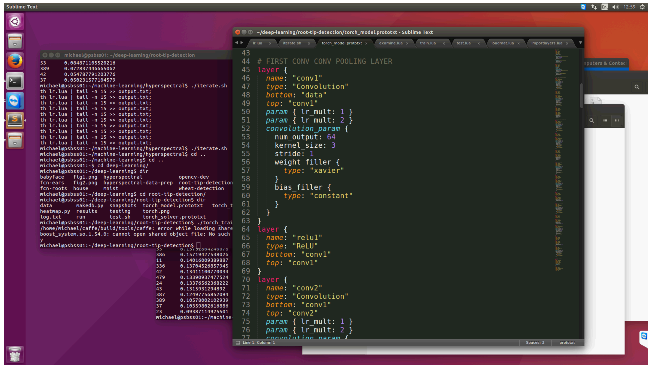
Tony Pridmore
Michael Pound

How this works in practice

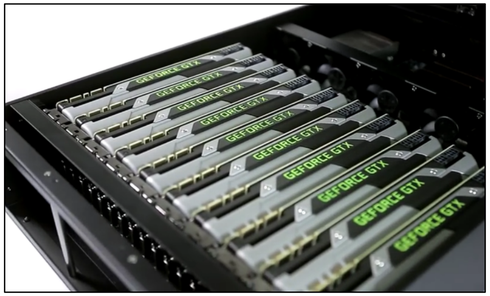
1. Capture and annotate dataset



2. Design network architecture

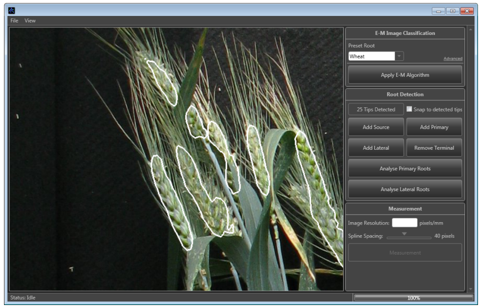


3. Train Network



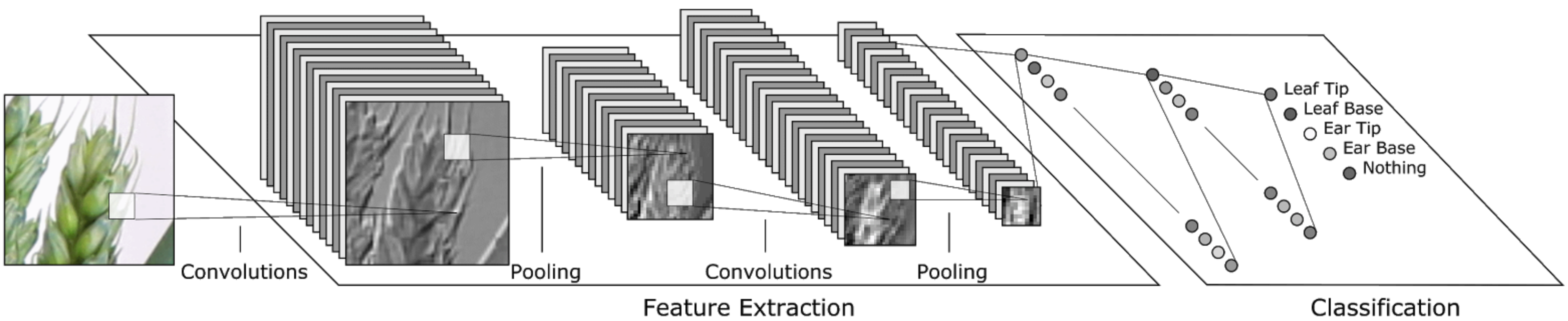
4. Test Network

5. Deploy

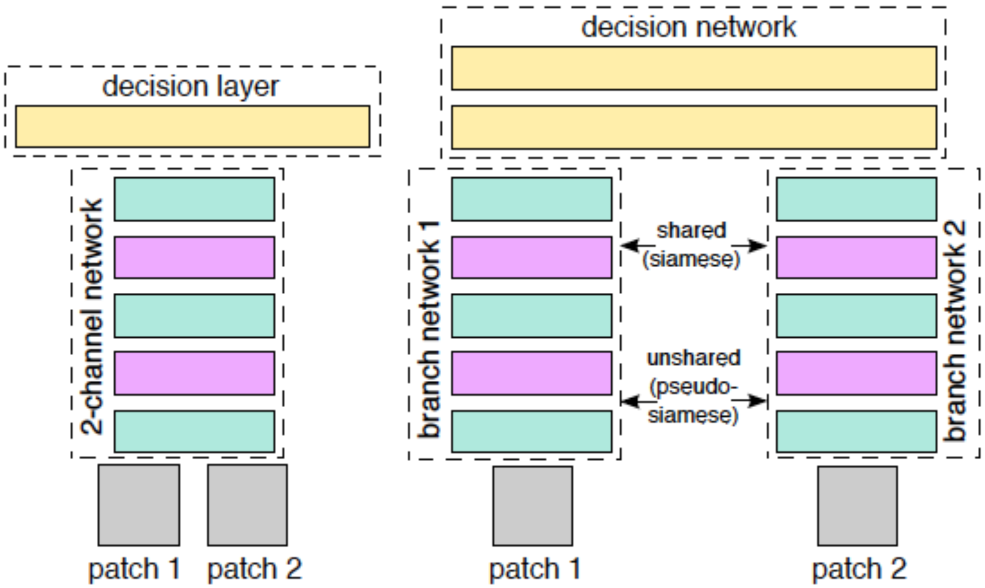


Classic CNNs

- Early layers perform convolutions
- Final layers perform classification
- All weights & convolution masks are learned

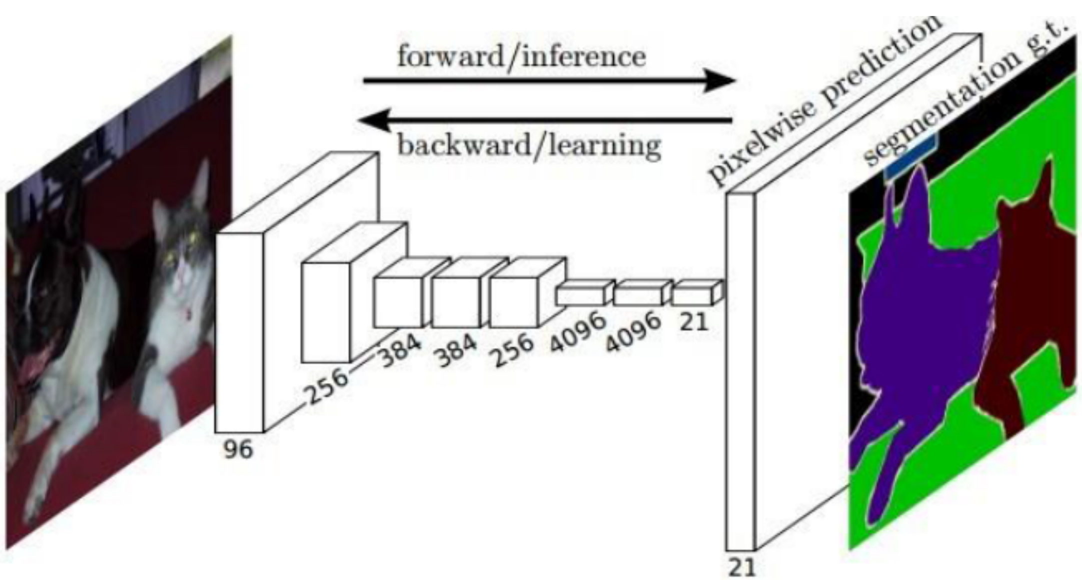


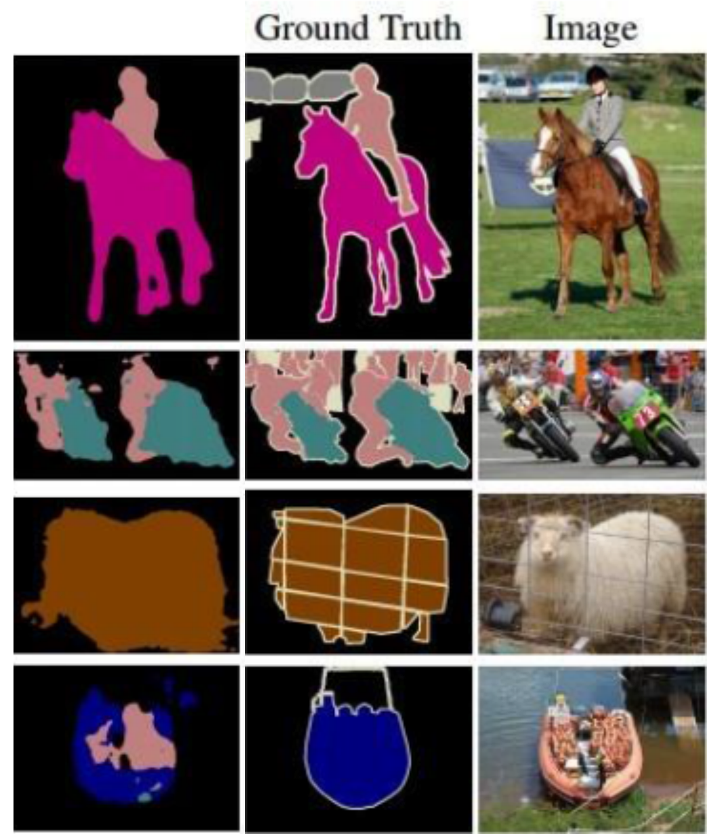
Multiple input images



- CNN approach has been applied to classic binocular stereo, wide baseline stereo and general image matching
- 2-channel approach has been more effective so far

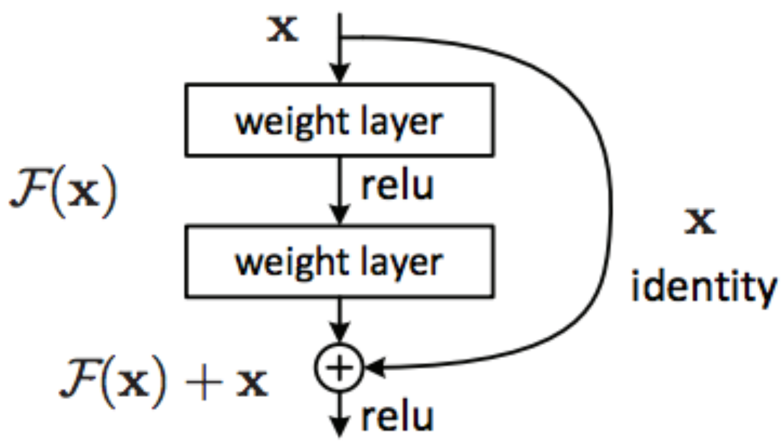
Fully Convolutional Architectures

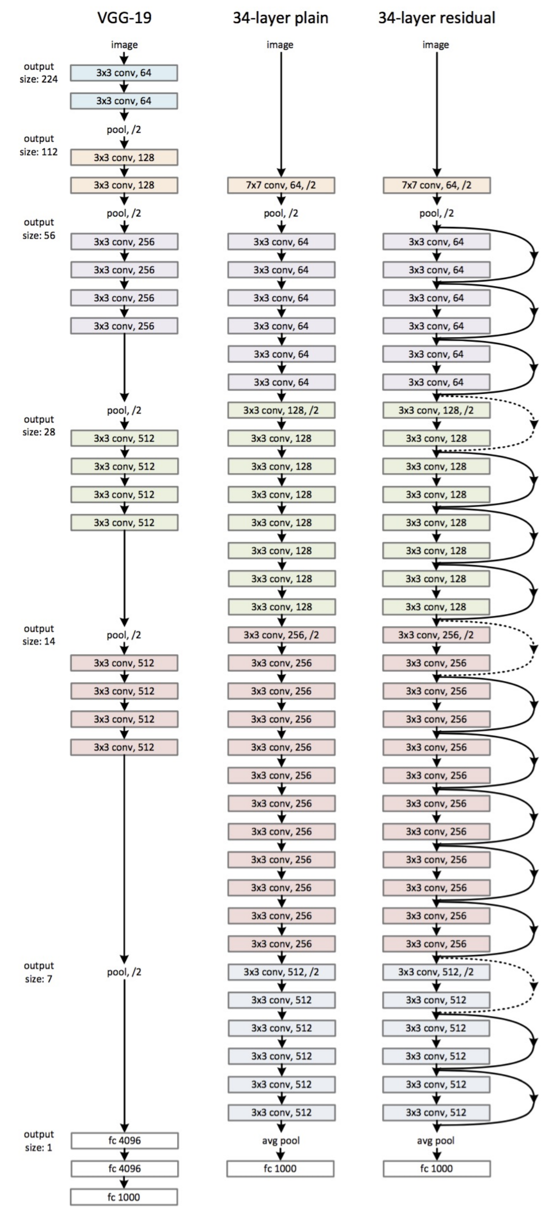




Fully Convolutional Networks for Semantic Segmentation
Long, Shelhamer, Darrell

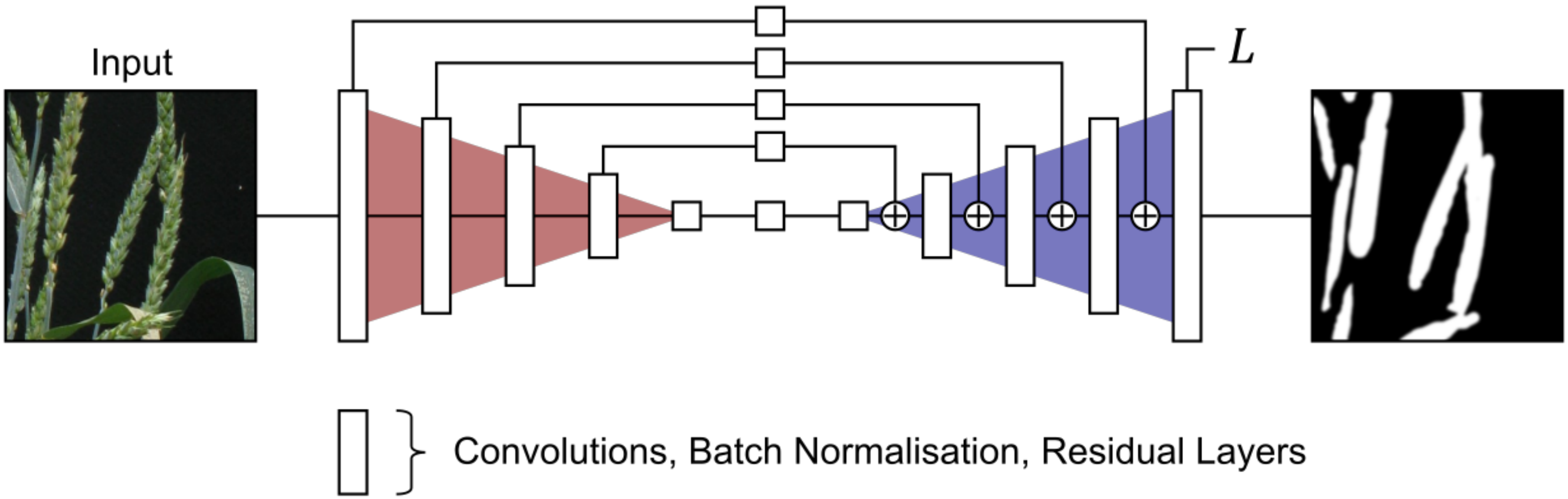
Residual Networks





- Skip layers allow the original features to pass through unaltered
- Aids in the training of very deep networks

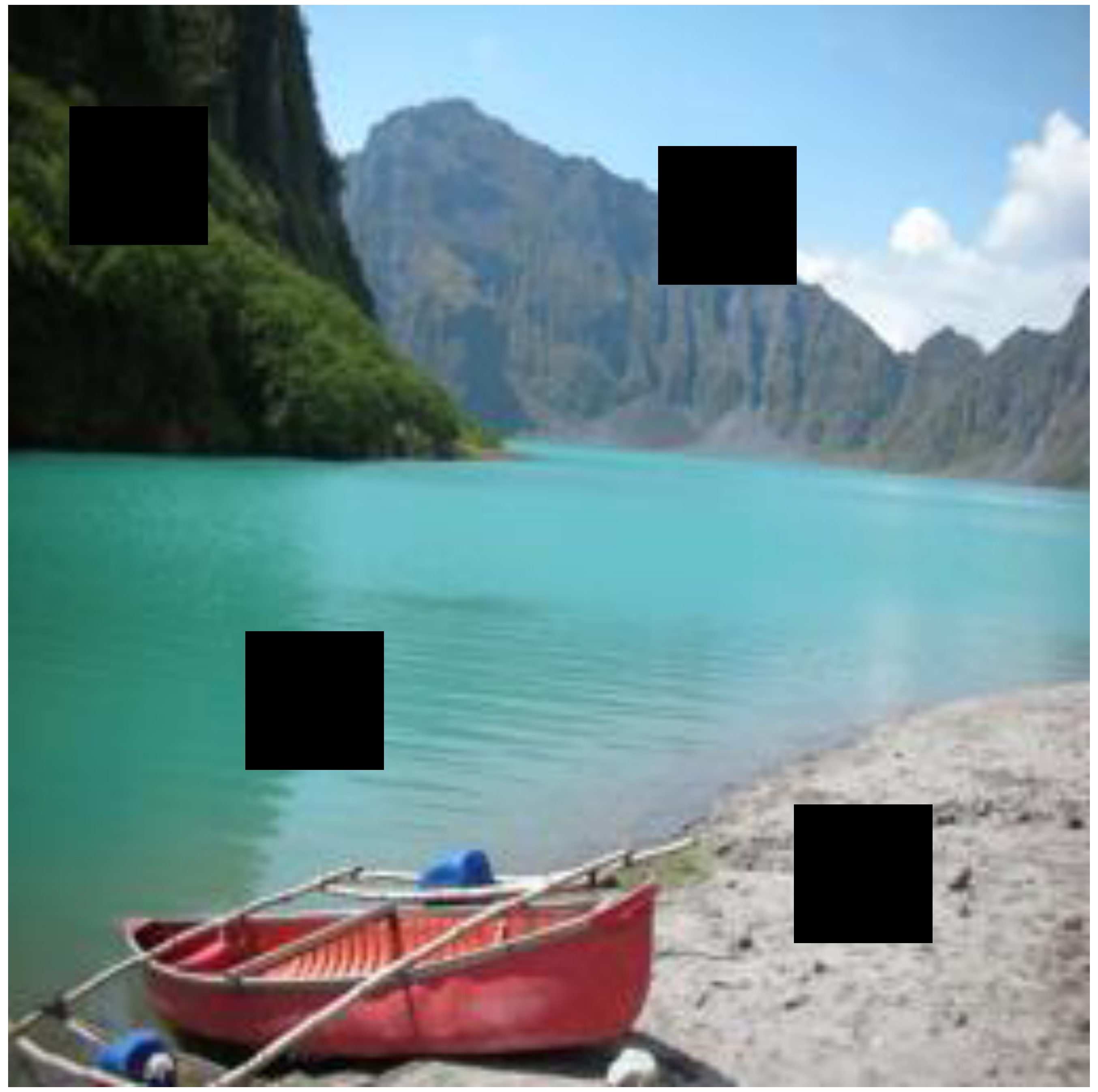
Encoder-decoder Networks

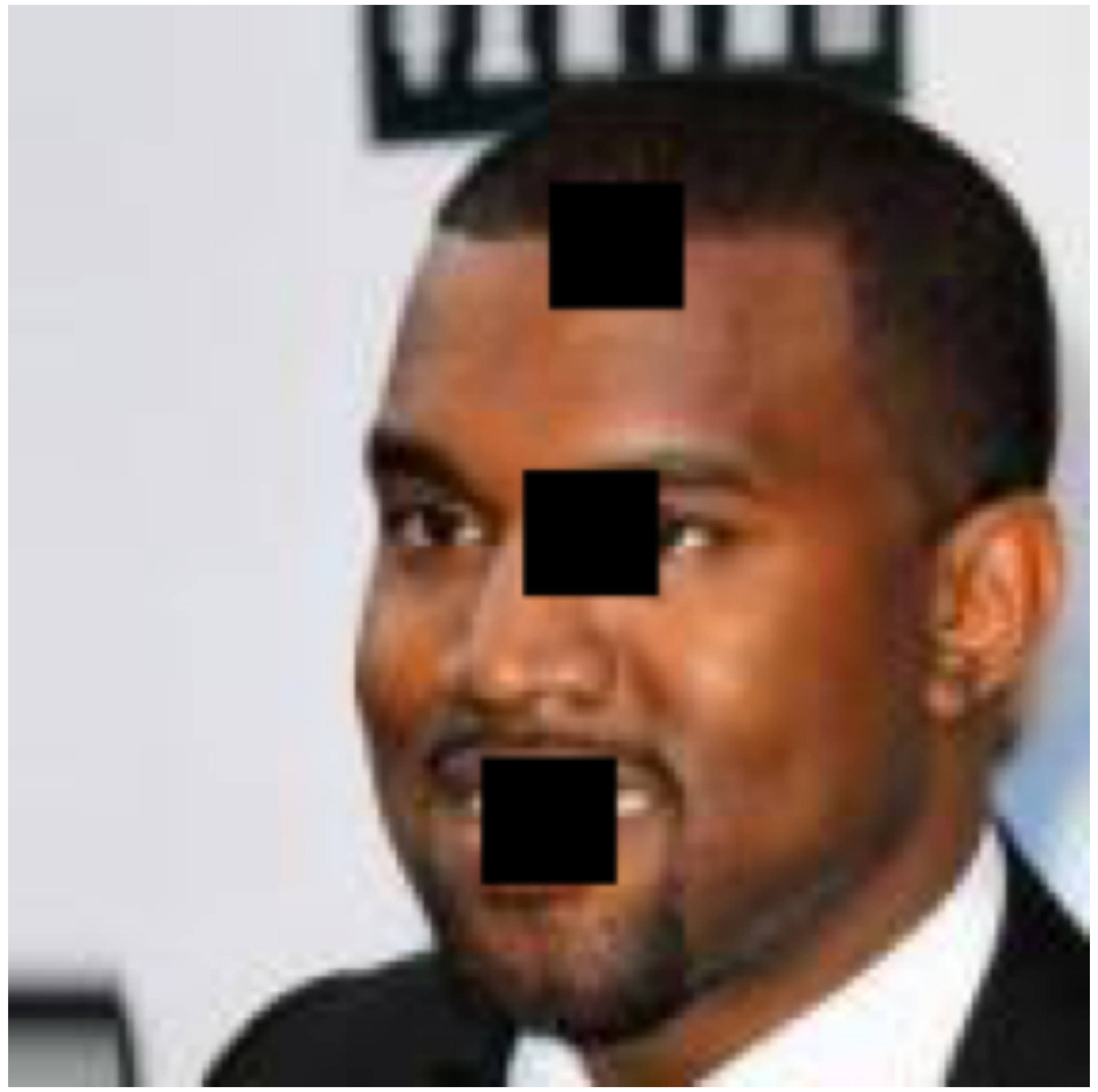


- Derives features from an input images, and then upsamples this back into an output image
- Extra connections (skip layers) help in producing accurage segmentation output

Object Completion Using Generative Deep Learning

By Oscar Mason
Supervised by Tony Pridmore

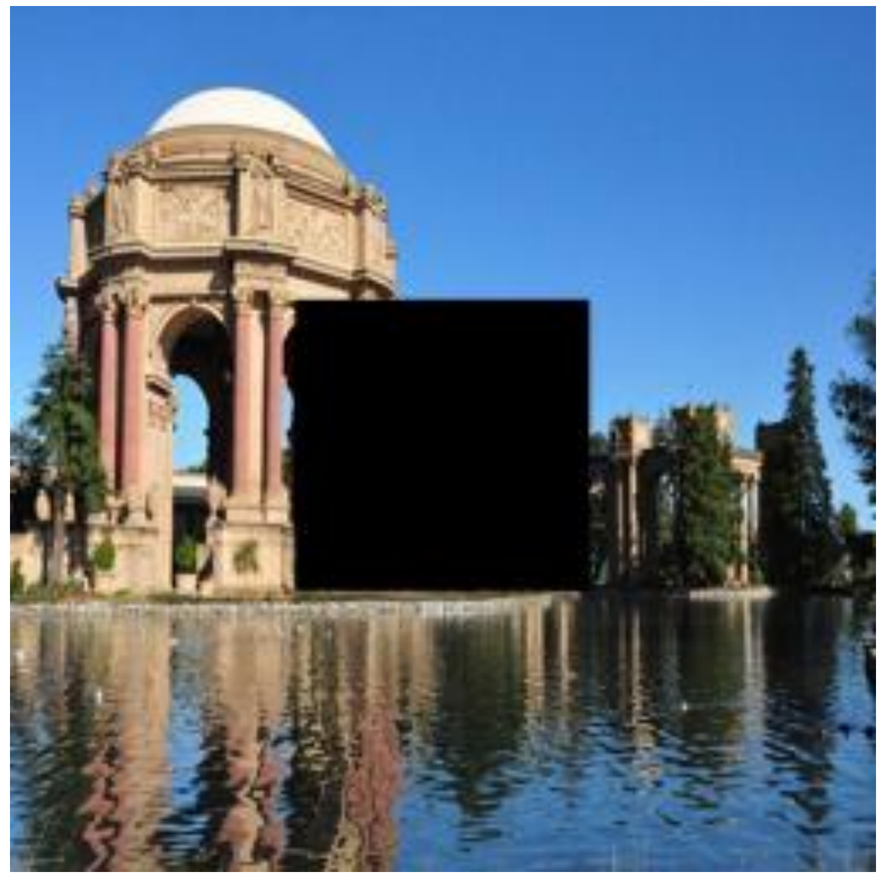




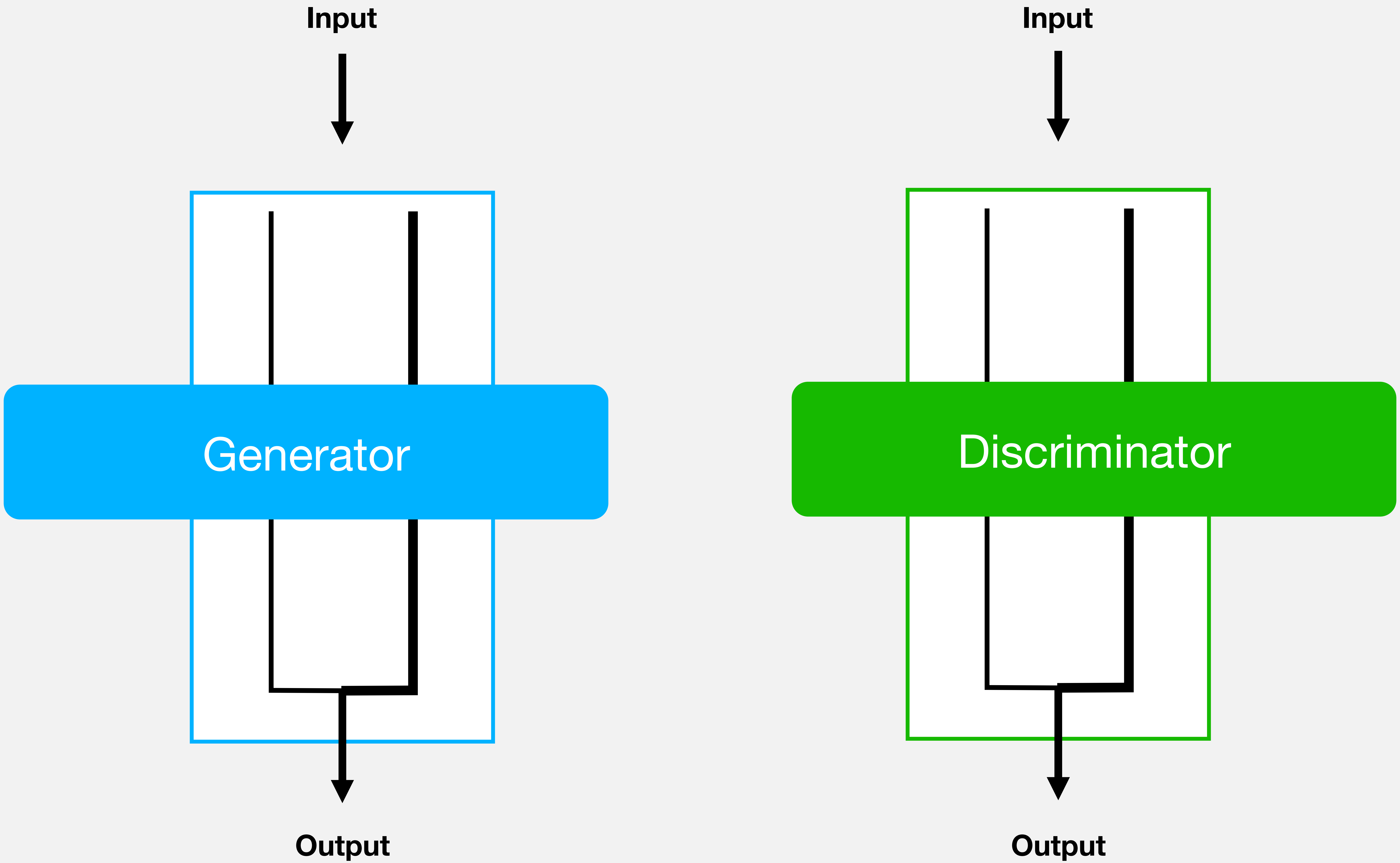
Generative Adversarial Networks

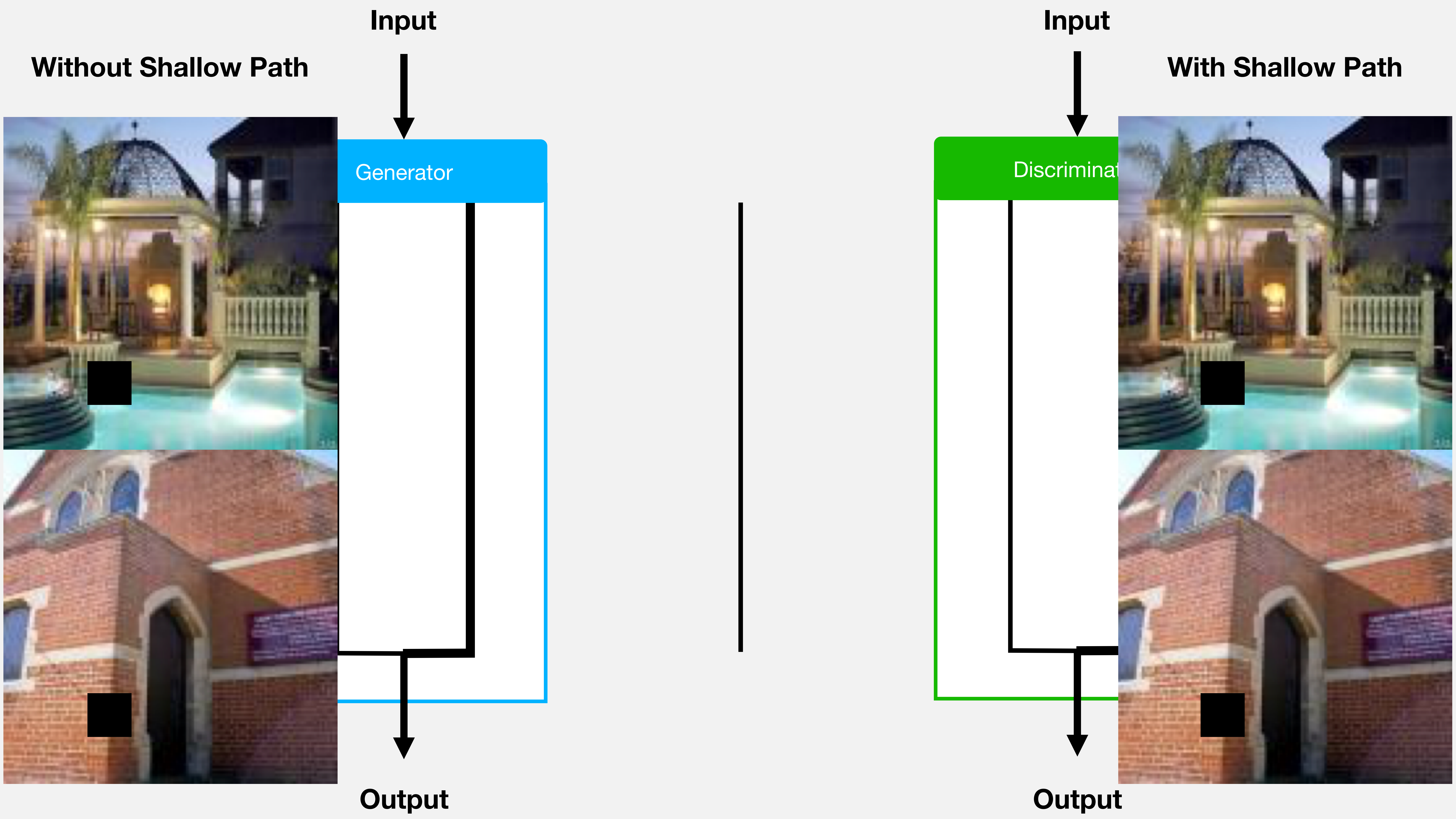
Discriminator

Generative Adversarial Networks



Generator


Input
Generator
Output
Input
Discriminator
Output



Input
Without Shallow Path
Generator
Output
Input
With Shallow Path
Discriminat
Output

Network

A novel architecture

- Small and Large Kernel Paths
- Shallow Path in Generator

Feature Selection

18.8 Million Trainable Parameters

38 Layers