Analytics

The Delta Mission

Increase access and usefulness of data to **empower changemakers** working to better their communities

We accomplish this through our two core programs

Data Service Grant

Bridging the nonprofit skill gap by enabling nonprofits to **accelerate their impact** through leveraging **skilled data professionals** on project opportunities.

Machine Learning for Good

Empower anyone, anywhere to leverage data for good in their communities by **building technical capacity** around the world.

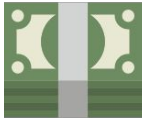
Since 2013...

Completed **~50 projects** with nonprofits and social impact organizations

Open-sourced and taught our machine learning curriculum to students globally

Recruited **200+ professionals** to work on projects

Donated over **20,000 hours**

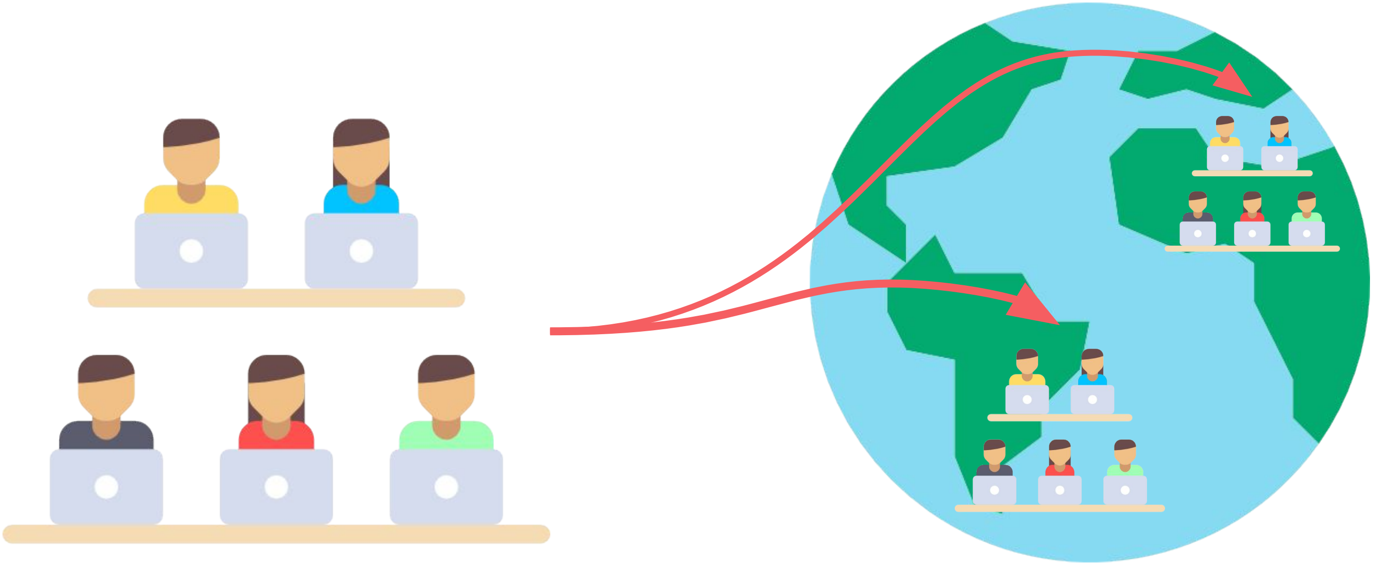


$0 charged for services

Analytics

Teaching Fellows

Delta has been working since 2013 to help non-profits around the world with their data. In 2017, we started the Teaching Fellows program.



Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

2019

2020

Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

Beta tested, revised, and open-sourced our curriculum
Taught subsequent classes in Morocco + the SF Bay Area

2019

2020

Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

Beta tested, revised, and open-sourced our curriculum
Taught subsequent classes in Morocco + the SF Bay Area

2019

Explored supplementing our curriculum with active learning techniques

2020

Now, we are going global.

Why Global Teaching Fellows?

Alicia Tsai Follow
Feb 26 · 5 min read

We just opened up our Global Teaching Fellows program. This is a big change from requiring that all our teaching fellows are local in the Bay Area. We wanted to explain why we made this change and what makes us excited about the 2020 teaching fellow program.

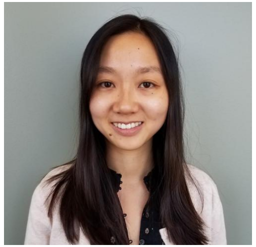
Delta Analytics is a 501(c)(3) nonprofit organization that was founded 6 years ago with a focus on helping nonprofits and social impact organizations leverage their data for good.

Three years ago, Delta Analytics started a Teaching Fellows Initiative. The goal of the program is to build technical capacity around the world. In 2017, we developed our pilot course, which focused on explaining the fundamentals of machine learning.

Our goal: support educators all around the world.

Meet our 2020 Bay Area teaching fellows!

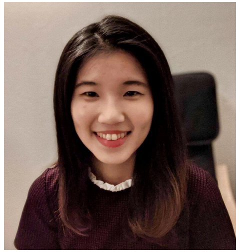
Sara Hooker
Co-Director



Amanda Su
Co-Director

Brian Spiering
Code Lead

Melissa Fabros
Teaching Lead



Alicia Tsai
Teaching Fellow

Raul Maldonado
Teaching Fellow

Allie Wang
Teaching Fellow

Meet our 2020 global teaching fellows!

Analytics

Today's lecture …

Convolutional Neural Networks

October 2, 2020

A little about me...

Dina Machuve

Lecturer, Nelson Mandela African Institution of S&T

Global Teaching Fellow 2020, Delta Analytics

Agenda

1. Tanzania applications for Convolutional Neural Networks
    a. Chicken production background and challenges
    b. Motivation for CNNs in developing countries

2. How we can use CNNs for chicken health monitoring?
    a. Introduction to CNNs
    b. Architecture of CNNs
    c. Application in the field

3. Coding example on CNNs for chicken health

Challenge

Which chicken is sick?

A

B

C

Challenge

Chicken health matters

TANZANIA

- Low poultry productivity from diseases
- Lack of trusted poultry data
- Chickens: 36 mil
- Households with chickens: 4.6 mil
- Population: 56.32 mil (2018)
- GDP: US$ 55.5 bil (2019)

Tanzania Livestock Modernization Initiative, TLMI, (2015)

Challenge

How do we improve chicken health?

- Farmers seldom keep records for managing their chicken health
- Rely on word of mouth
- Chicken diseases are devastating:
    - High mortality rates
    - Low production rates
    - Some affect human
    - Farmers have limited access to livestock field officers

Challenges at the coalface: Tanzania's Livestock Field Officers (2018)

Challenge

Chicken diseases monitoring now

Salmonella, Newcastle and Coccidiosis chicken diseases:

- Diagnosed by lab procedures using droppings samples
- It takes 3 -4 days to get results
- Clinical signs
- Access to the services by farmers is expensive and limited
- Extension officers lack tools for rapid diagnostics

Challenge

Chicken monitoring opportunity

With better tools:

- Farmers can diagnose chickens themselves
- Diagnosis time 4 days to 3 minutes!
- Isolate sick chickens faster to prevent disease spread
- 100% of chicken losses from Salmonella will be reduced due to early diagnosis
- Avoid economic loss by acting in a timely manner

Motivation

Image data is better than lab data in developing countries

- Ubiquity of mobile phones in developing countries:
  - 48.86 mil mobile phone subscriptions in Tanzania
  - 444 mil in Africa and 250 mil have smartphones
  - Mobile phones as a sensor

- Low levels of literacy and multiple languages
  - Images (geo-coded) form a universal data format
- Socially, people are obsessed in images!

TCRA, Quartely Communications Statistics, March 2020
Mobile is Sub-Saharan Africa: Can world's fastest-growing mobile region keep it up? (2018)

Convolutional Neural Networks (CNNs)

1. Introduction to CNNs
2. Architecture of CNNs
3. Application in the field

Intro
CNNS

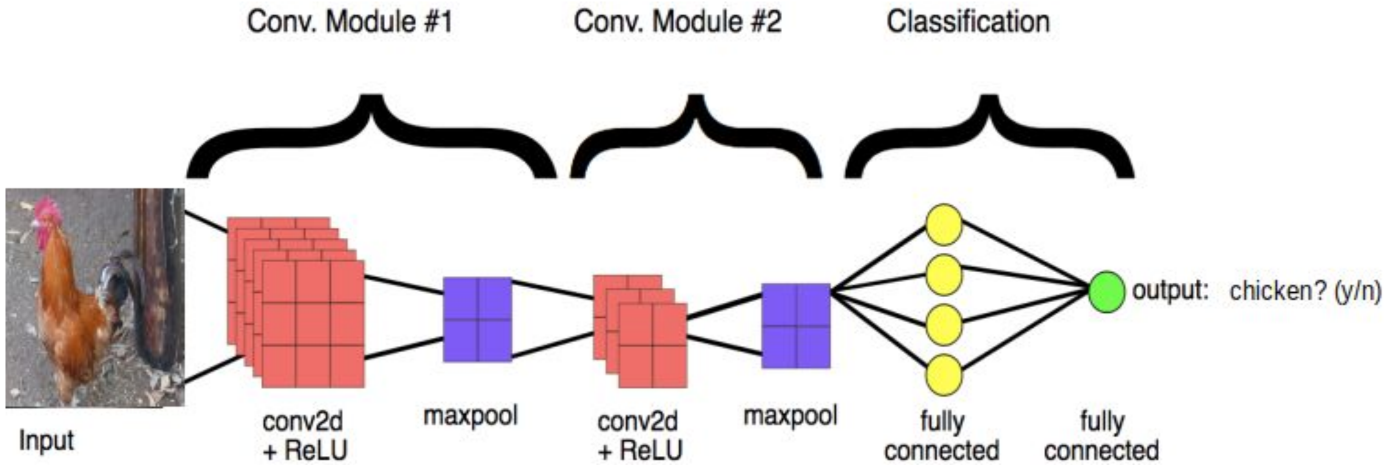
CNNs are specialized kind of neural network

- Process grid-shaped data
  - Time-series data
  - Image data
- Commonly applied to analyzing visual imagery
- Successful in practical applications:
  - Win many contests, ImageNet challenge (ILSVRC)
  - OCR and handwriting recognition systems
  - Diseases diagnostics

Intro CNNS

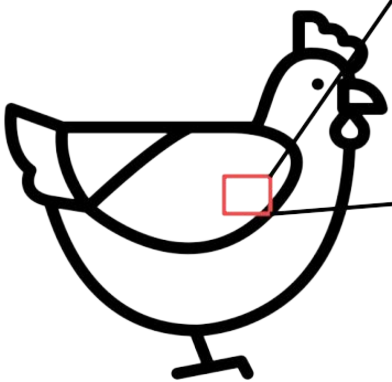
Types of layers in a CNN

- **Input layer:** the raw image data
- **Convolutional layer:** convolutions between neurons and a patch in the input
- **Rectified linear unit layer**
- **Pooling layer**
- **Fully connected layer:** computes the output scores in the last layer



Convolution

CNN considers structure



```
[[169 164 152 152 152 151 140 149 144]
 [163 168 156 153 154 146 144 146 144]
 [152 157 149 147 146 132 129 140 139]
 [143 144 142 136 133 130 133 143 139]
 [144 136 138 134 130 131 132 146 142]
 [133 129 134 131 128 130 136 142 144]
 [127 123 131 132 132 129 133 145 151]
 [127 133 130 129 125 126 140 152 157]
 [117 118 130 132 128 133 149 157 159]]
```

- Image is a high-dimensional vector
- Visual abstractions of input image for learning
- The numbers are pixel intensities (0-255)

What is Convolution?

The operation that applies a filter (w) to an input feature map (x) to compute new features

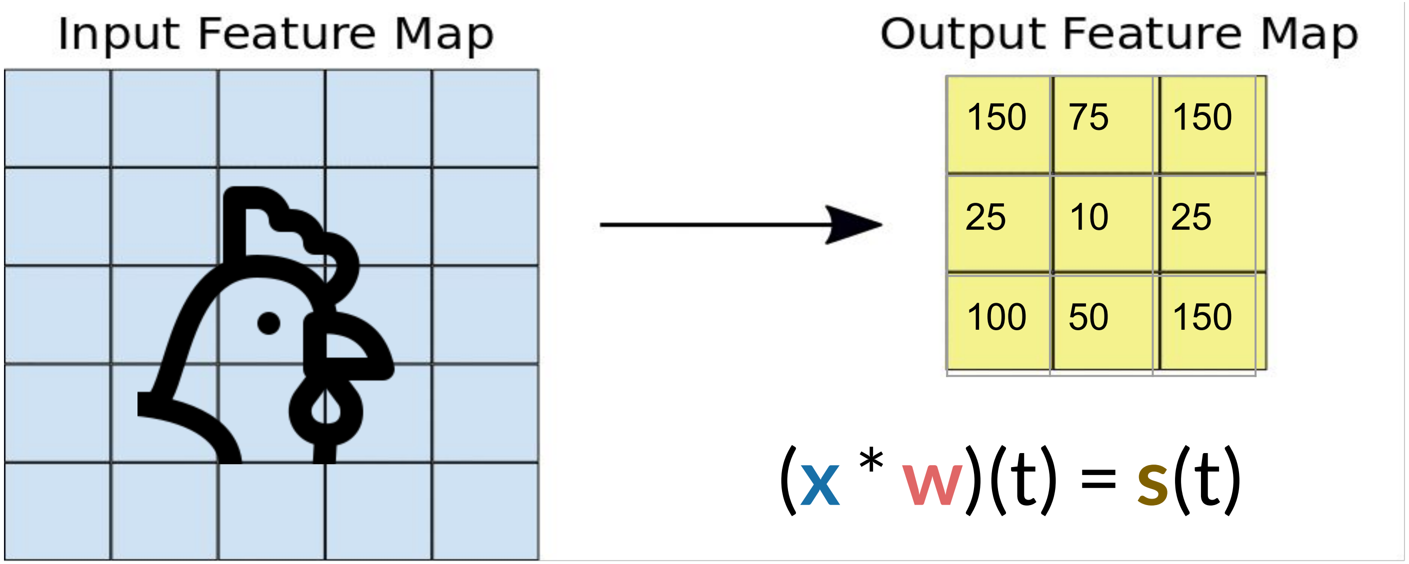


Image source: Machine Learning Practica, Google

Convolution

Applies a filter to a patch of an image to produce an output feature map.

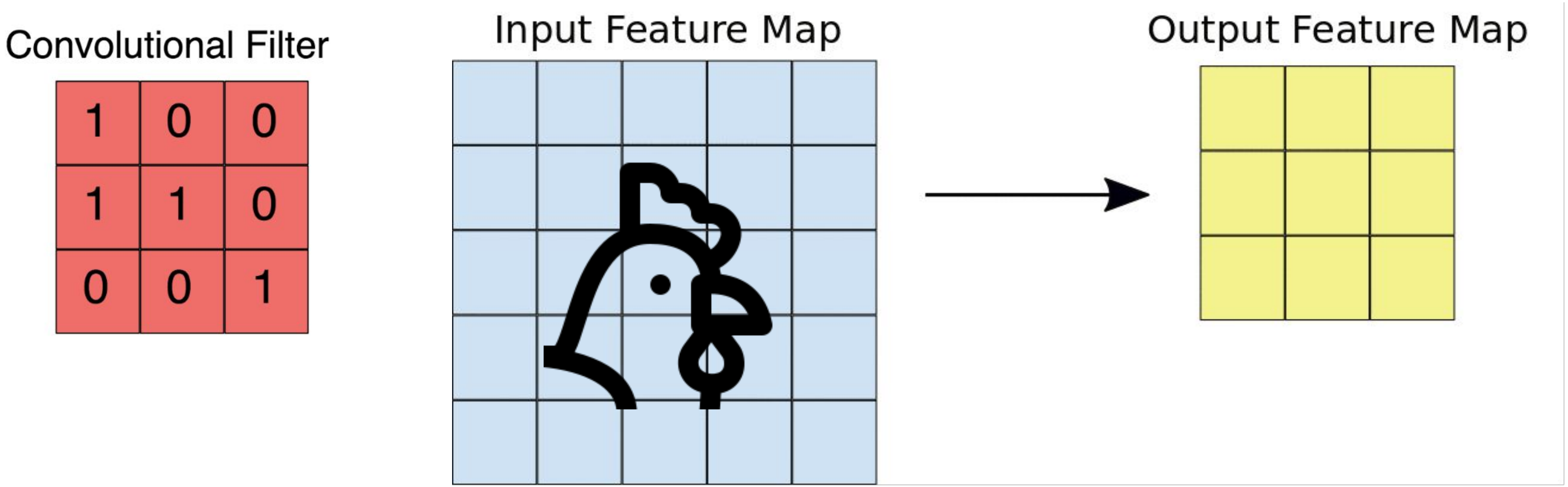


Image source: Machine Learning Practica, Google

Convolution

Useful definitions

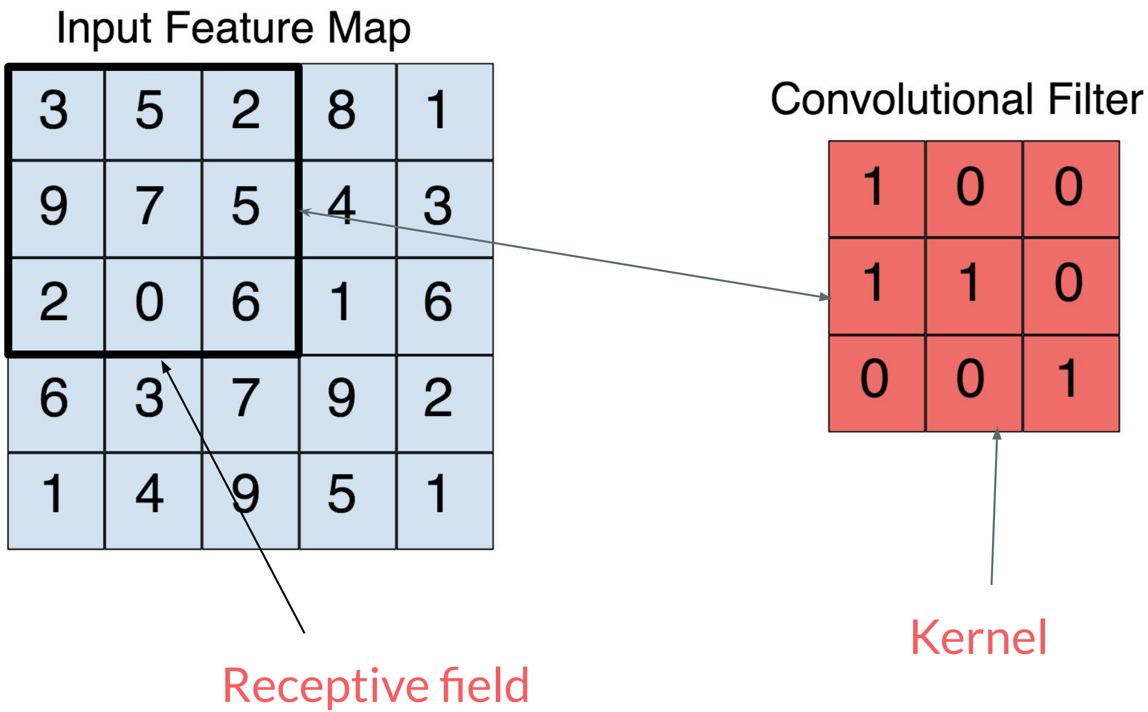
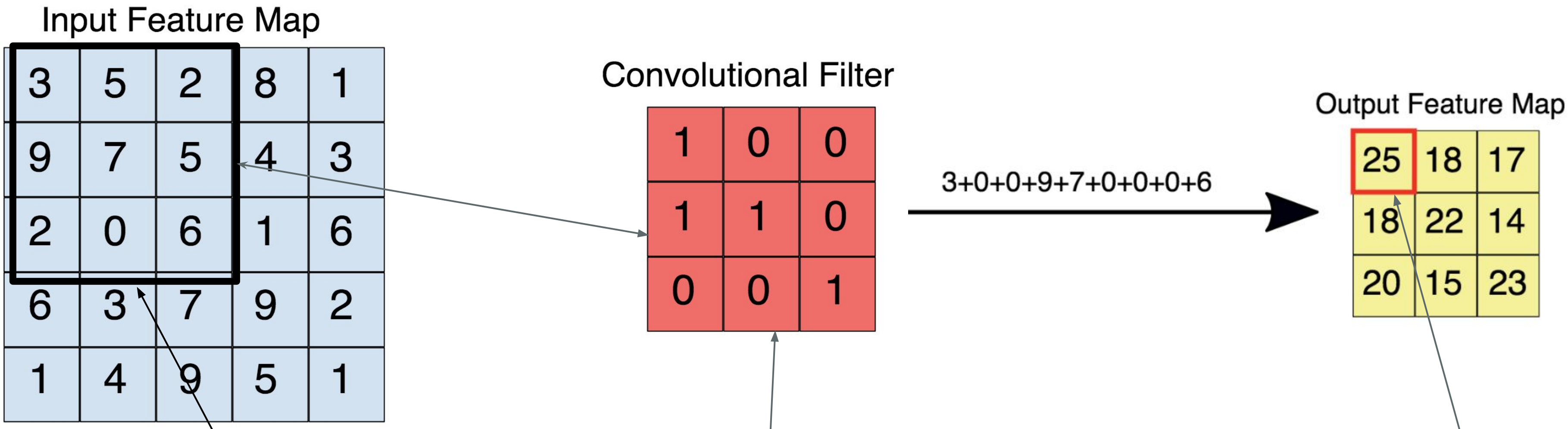


Image Credit: Sarah Hooker (2018)

Convolution

Useful definitions



Receptive field: region of the input space the kernel pays attention to; the inputs that affect the value .

kernel/filter: we slide the kernel over the input. At each location the element wise product of input and kernel is computed, which is then summed up to obtain the output activation.

Neuron: the sum of the elementwise product of filter and receptive field.

Check for Understanding

What is the element wise multiplication of the convolutional filter and the highlighted receptive field?

Input Feature Map

3	5	2	8	1
9	7	5	4	3
2	0	6	1	6
6	3	7	9	2
1	4	9	5	1

Convolutional Filter

1	0	0
1	1	0
0	0	1

→ ?

Output feature map:

A.

4	0	1
6	0	1
2	0	0

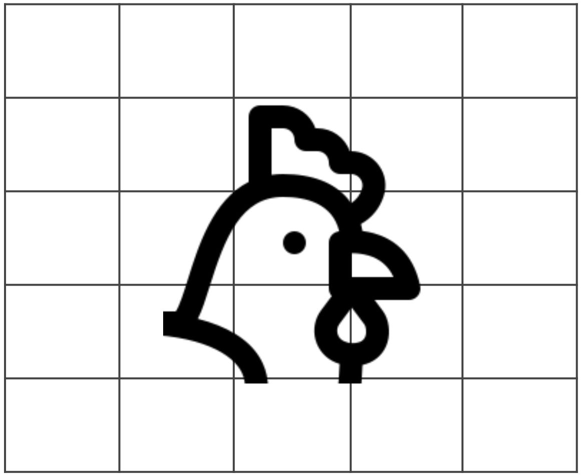
B.

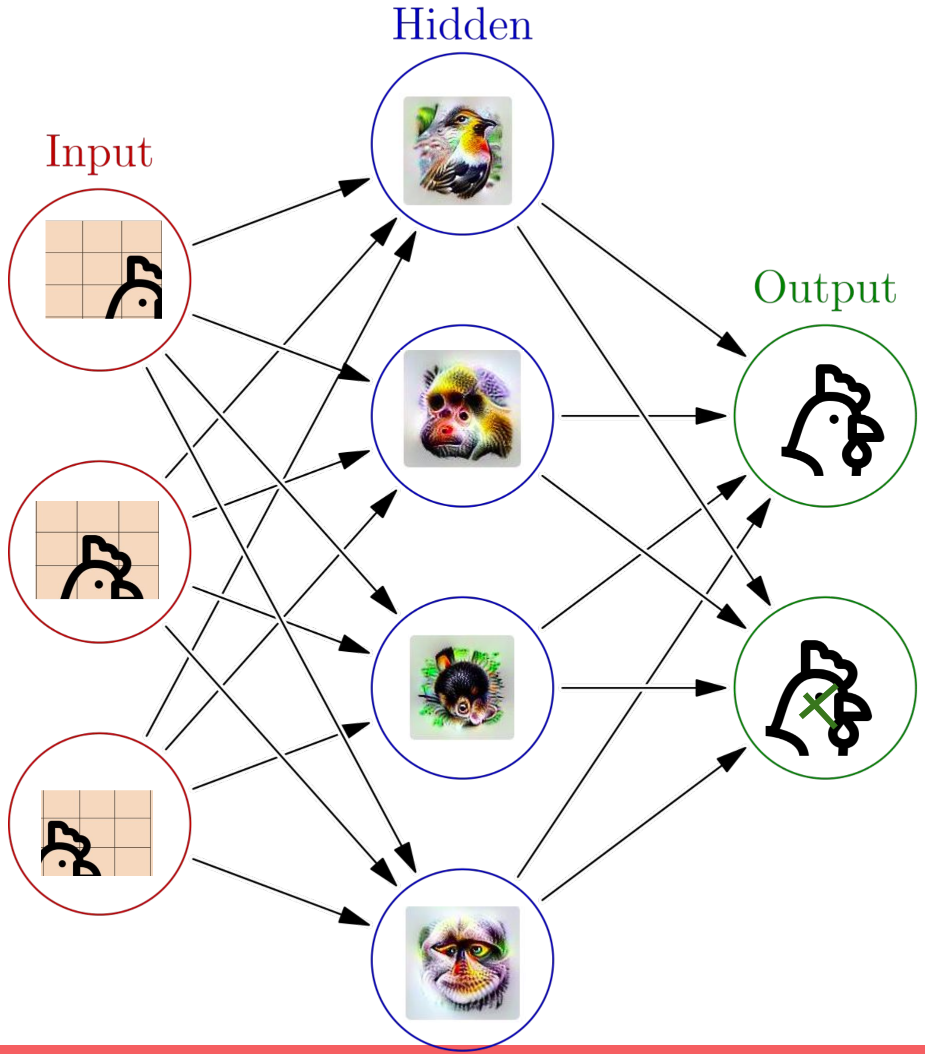
5	0	0
6	1	0
0	0	2

What is a neural net?

- You can run a convolution over an image
- But the neural net architecture does the decision making
  - Tuning the weights and biases on neurons
  - Many layers of computation “neurons” deciding a classification.
- The power of neural net lies in their *hidden layers*

What is a neural net?





Neural net

What the hidden layer "sees"

Making sense of these activations is hard because we usually work with them as abstract vectors:

$a_{3,6}$ = [0, 0, 0, 0, 90.0, 0, 0, 0, 61.4, 0, 0, 27.3, ...]

With feature visualization, however, we can transform this abstract vector into a more meaningful "semantic dictionary".

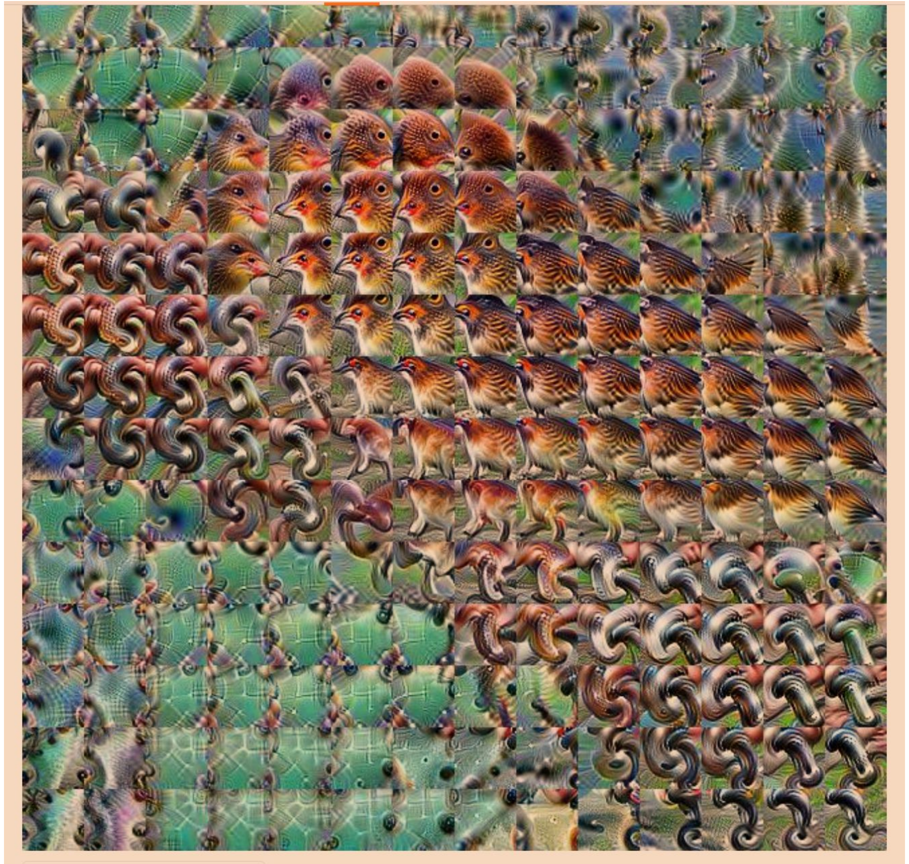
https://distill.pub/2018/building-blocks/

Neural net

What the network can "see"

Convolutional neural nets are good solution for the challenges around improving chicken health:

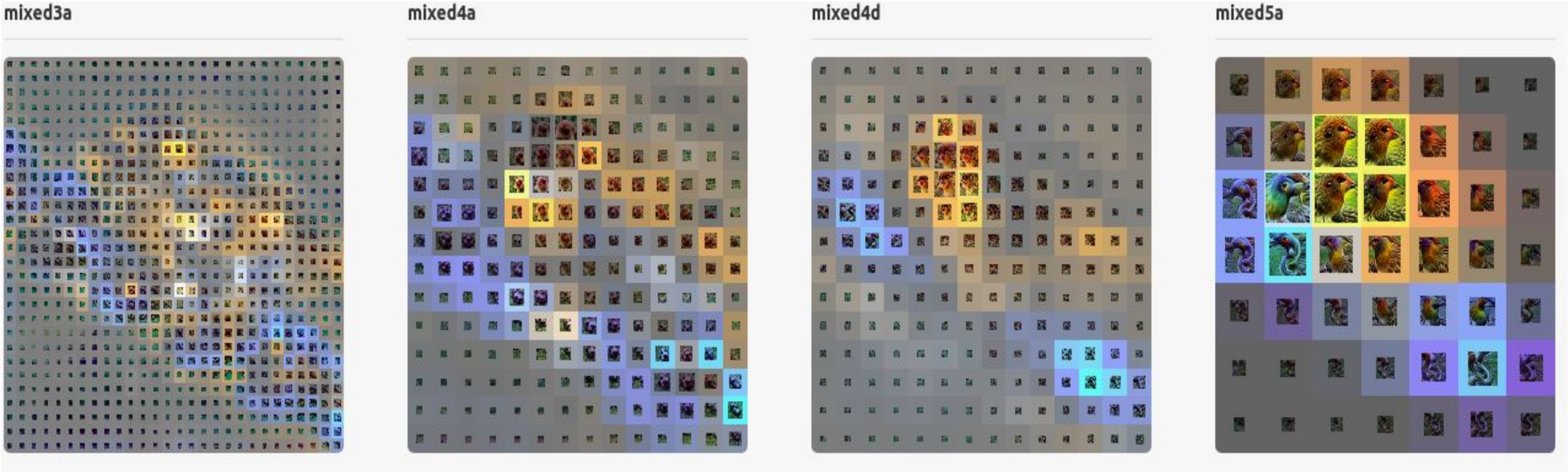
- The network can detect patterns on the layer level and the aggregate level.
- Can detect patterns that the human eye might miss
- Once trained can detect patterns quickly.



https://distill.pub/2018/building-blocks/

Neural net

What the network "assembles"



https://distill.pub/2018/building-blocks/

Application in the field

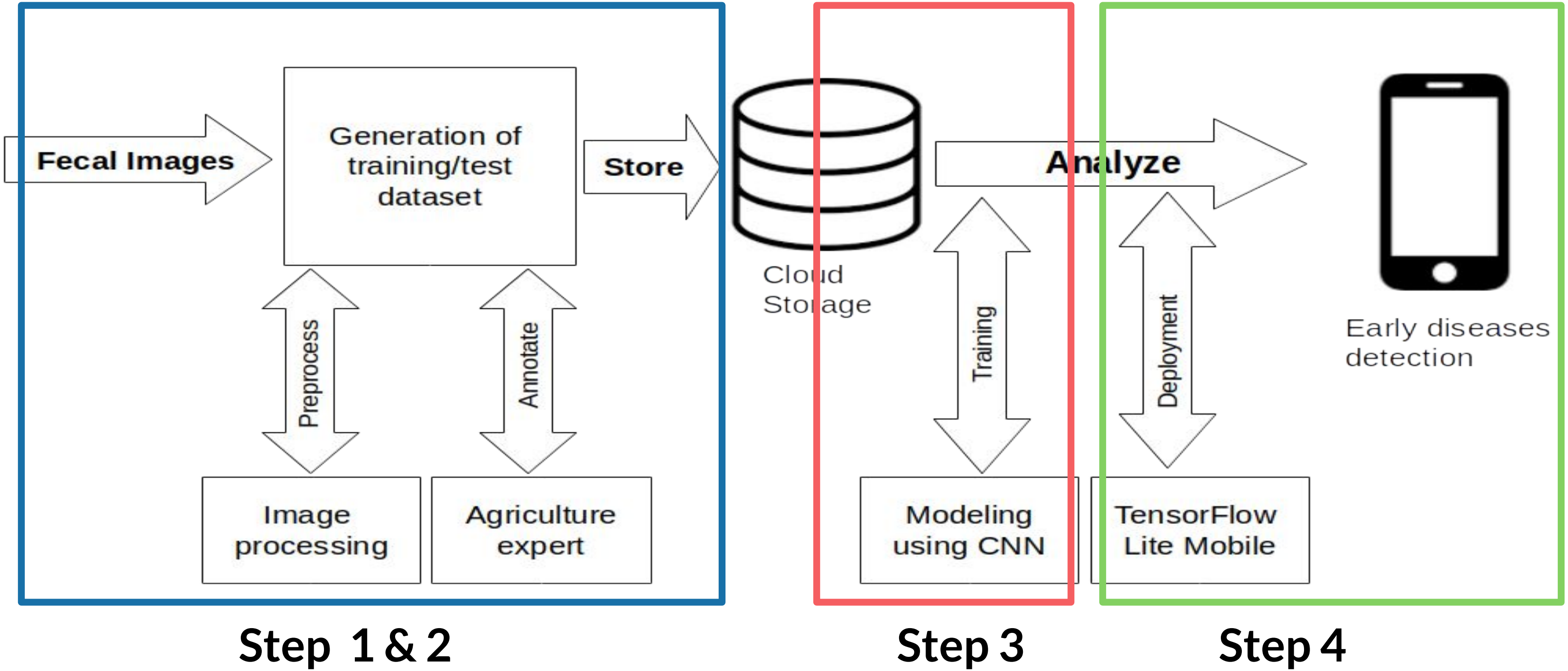
Task

Can we train a model that can correctly classify a set of images for chicken diseases?

After training, our model should return the correct label for each image.

Task

Workflow



Step 1: Building dataset is labor intensive

- Physically go to farms & take pictures

Step 1: Build dataset

- Find supportive chicken farmers
  - building trust with farmers
- Field work at farms
  - Accompanied by field livestock officers or veterinarian
  - Travel to farms takes 2-3 hours
  - It takes ≈4 hours to collect 150 images
- Why chicken droppings?
  - Clinical signs are similar
  - Color and shape of droppings supplement clinical signs

Dataset: Chicken Droppings

Class	# Images
Healthy	777
Coccidiosis	891
Salmonella	1133
TOTAL	**2801**

- Target **8000** images

Healthy

Coccidiosis

Salmonella

Step 2: EDA on images

- Features of the dataset
  - Color and shape of droppings
  - the number of chickens
  - type of feed
  - geocodes
- Limitations in the data
  - limited/missing clinical history data of the chickens
- Record keeping is fundamental for improving productivity of chicken farmers

Step 3: Train Model

- Baseline model training
  - Binary Image classification
  - Sequential model on Keras
- Transfer learning for selection of appropriate model for deployment
  - On mobile for farmers and field officers
- Metrics for the model
  - accuracy

Modeling

Computer vision solutions are good for noticing patterns

Image Classification

Classify the main
object
in an image

Object localization

Predict the region of the image where the main object is present.

Object recognition

Predict the region where the objects are present and classify all objects in the image.

Task

After training, our model should return the correct label for each image

"healthy"

"coccidiosis"

"salmonella"

Step 4: Implement trained model in a mobile app / Productionize model

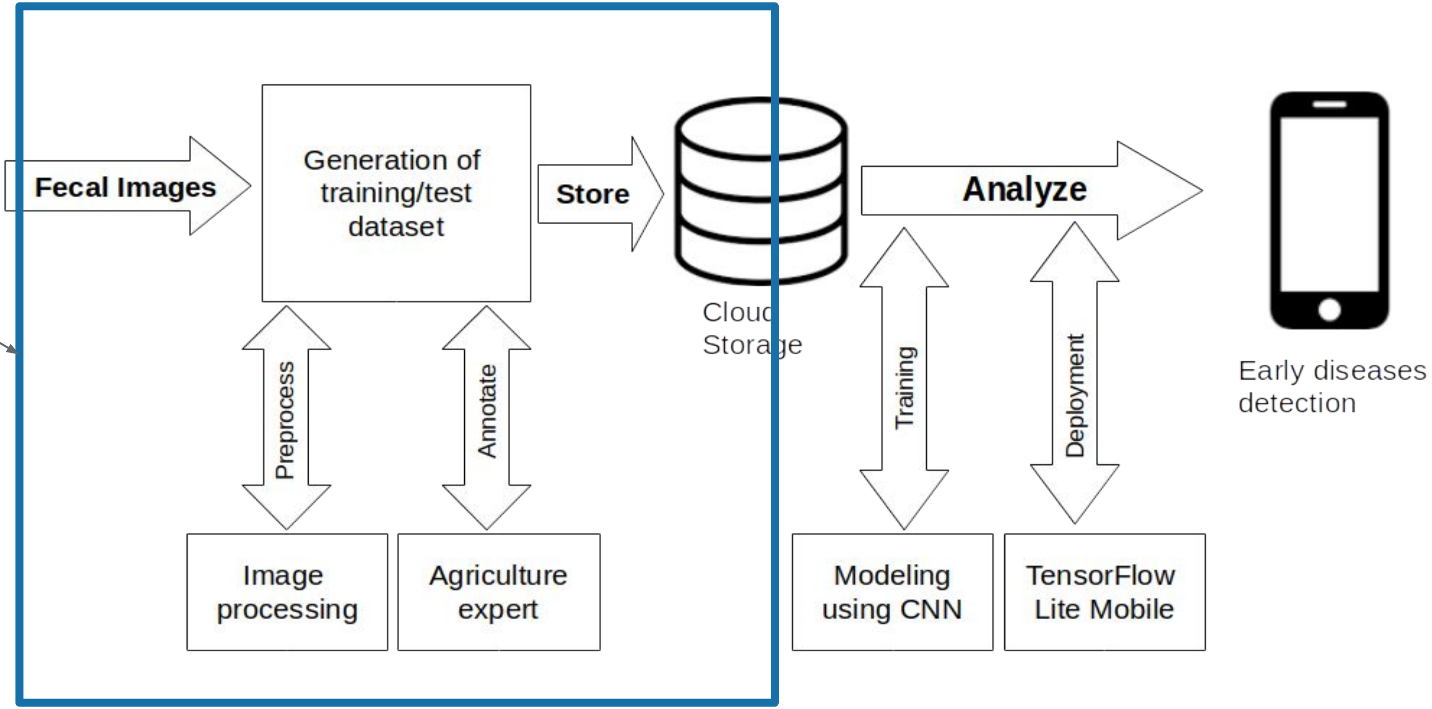
- The model will be deployed on Android mobile application
  - Access will be Free
- Train farmers and livestock officers to use it
- Model will be served on the mobile app
  - To allow usage offline
  - Model updates will be online

Our Vision

- A mobile application for use by farmers and livestock officers:
    - For early detection of Coccidiosis, Salmonella and Newcastle
- A machine learning dataset for poultry diseases diagnostics
- Reduced outbreaks of diseases for farmers using the app

Our Progress

We are here!



Our Progress

- Funding:
  - OWSD Early Career Fellowship from Nov. 2019 to Oct. 2021
  - Won IndabaX AI4D innovation grant to fund collection of more data progress on this work in August 2020
- Collaboration Partners
  - Delta Analytics Teaching Fellowship
  - MD4SG Development Group
- Collected initial dataset of 2801 images, target 8000
  - Dataset will be shared on open access

Our current Challenges

- Team is small and stretched because have to do other jobs in addition to this research
- Need more farmer buy-in
- Training is online on Google Colab
  - Challenge on unstable internet

Further exploration

- Slides and Colab notebook will be made available on the Project website
- Goodfellow, I., Courville, A., & Bengio, Y. (2016). *Deep Learning* (Vol. 1). Cambridge: MIT press
- The Building Blocks of Interpretability. https://distill.pub/2018/building-blocks/

Coding Example on CNNs

Goal(s) for Coding walkthrough

- Binary Image Classification
  - Image pre-processing
  - Training using simple ConvNet model on Keras
- Visualizing the results
- Improving performance
  - data augmentation
  - Transfer learning
- Link to collab notebook

Thank you.......

Melissa
Mentor

Yad Conrad

Analytics

Alicia

Raul

Brian

Amanda

Allie

Sara

- 2020 Global Teaching Fellows

Contact

@dmachuve

https://twiga2.github.io/apdd/

Thank you. Questions?