Object Recognition for Fun and Profit

Anil Thomas

SV Deep Learning Meetup
November 17th, 2015

Outline

- Neon examples
- Intro to convnets
- Convolutional autoencoder
- Whale recognition challenge

NEON

Neon

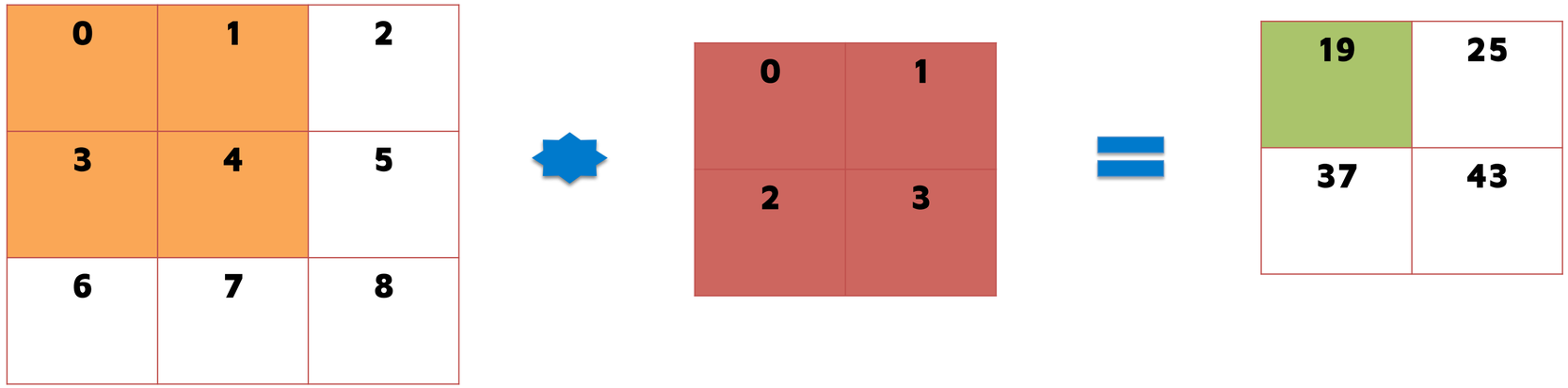
- Modular components
- Extensible, OO design
- Documentation
- neon.nervanasys.com

Backends	NervanaCPU, NervanaGPU NervanaEngine (internal)
Datasets	Images: ImageNet, CIFAR-10, MNIST Captions: flickr8k, flickr30k, COCO; Text: Penn Treebank, hutter-prize, IMDB, Amazon
Initializers	Constant, Uniform, Gaussian, Glorot Uniform
Learning rules	Gradient Descent with Momentum RMSProp, AdaDelta, Adam, Adagrad
Activations	Rectified Linear, Softmax, Tanh, Logistic
Layers	Linear, Convolution, Pooling, Deconvolution, Dropout Recurrent, Long Short-Term Memory, Gated Recurrent Unit, Recurrent Sum, LookupTable
Costs	Binary Cross Entropy, Multiclass Cross Entropy, Sum of Squares Error
Metrics	Misclassification, TopKMisclassification, Accuracy

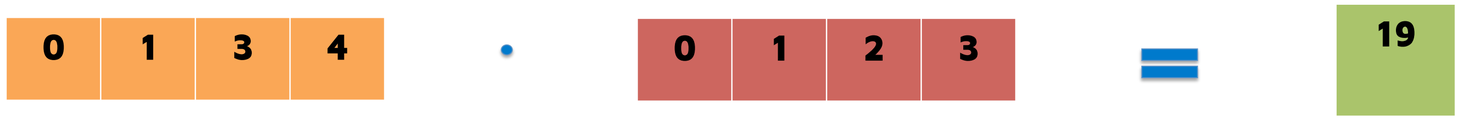
HANDS ON EXERCISE

INTRO TO CONVNETS

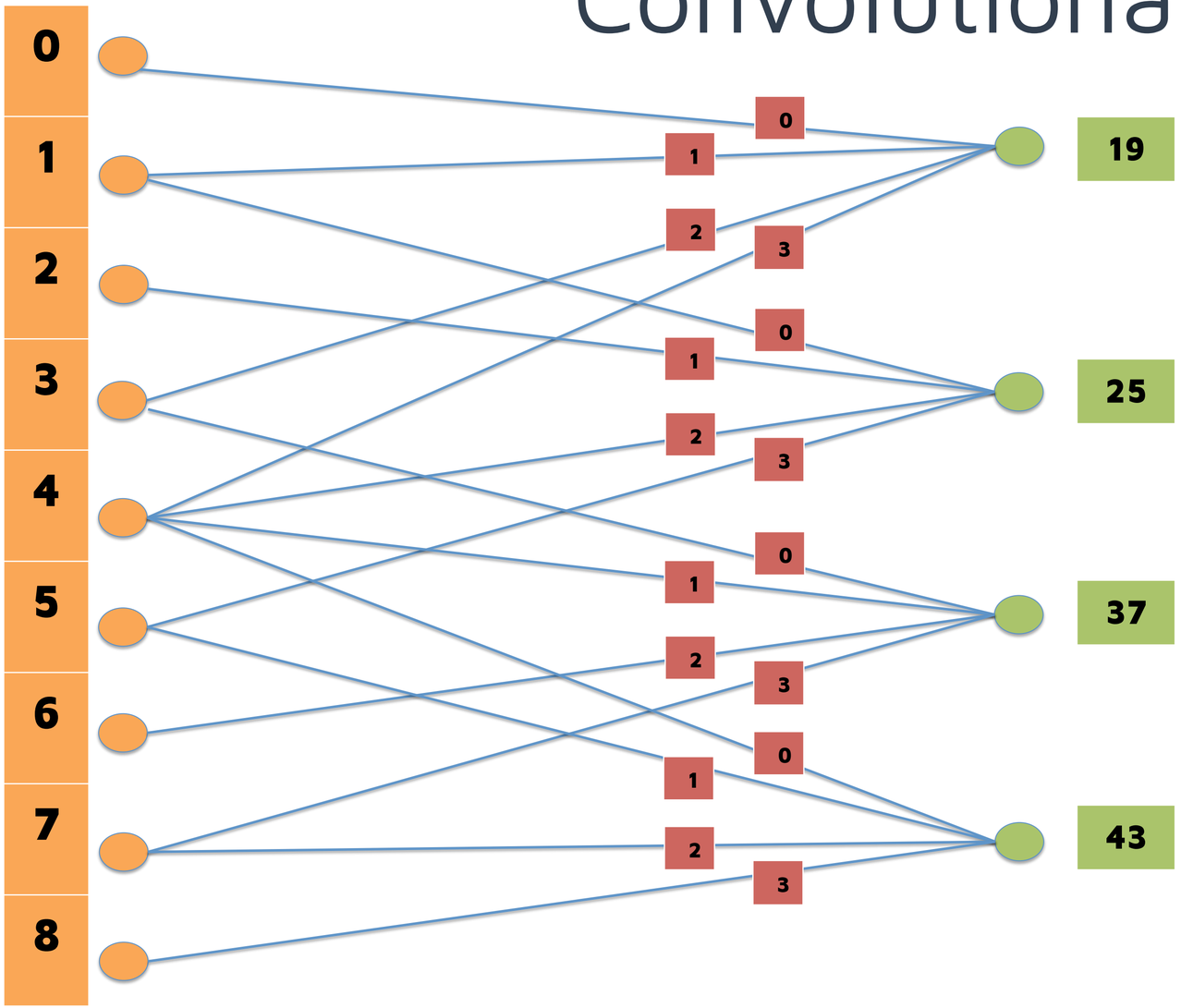
Convolution



- Each element in the output is the result of a dot product between two vectors



Convolutional layer



0	1	2
3	4	5
6	7	8

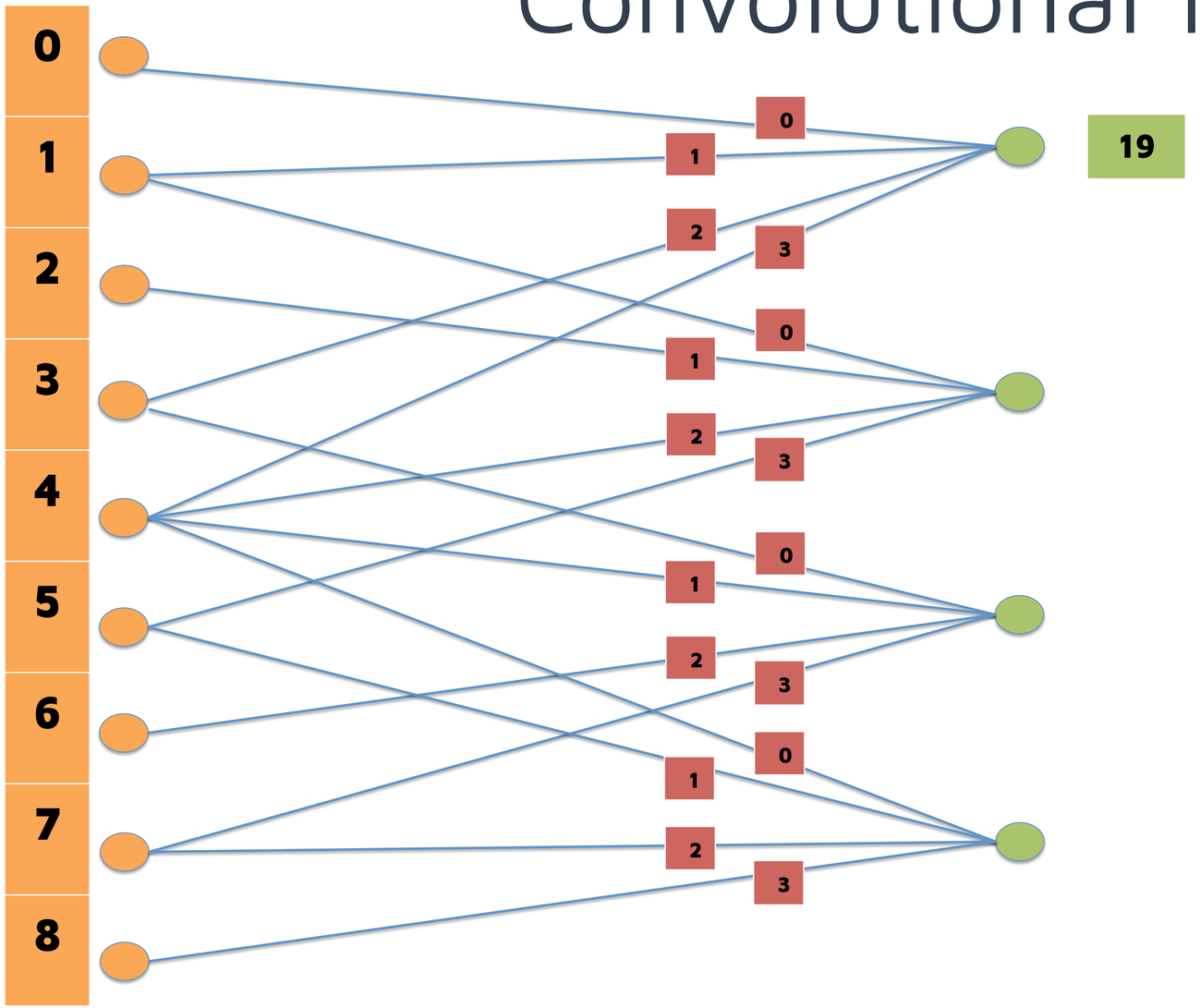
0	1
2	3

=

19	25
37	43

nervana

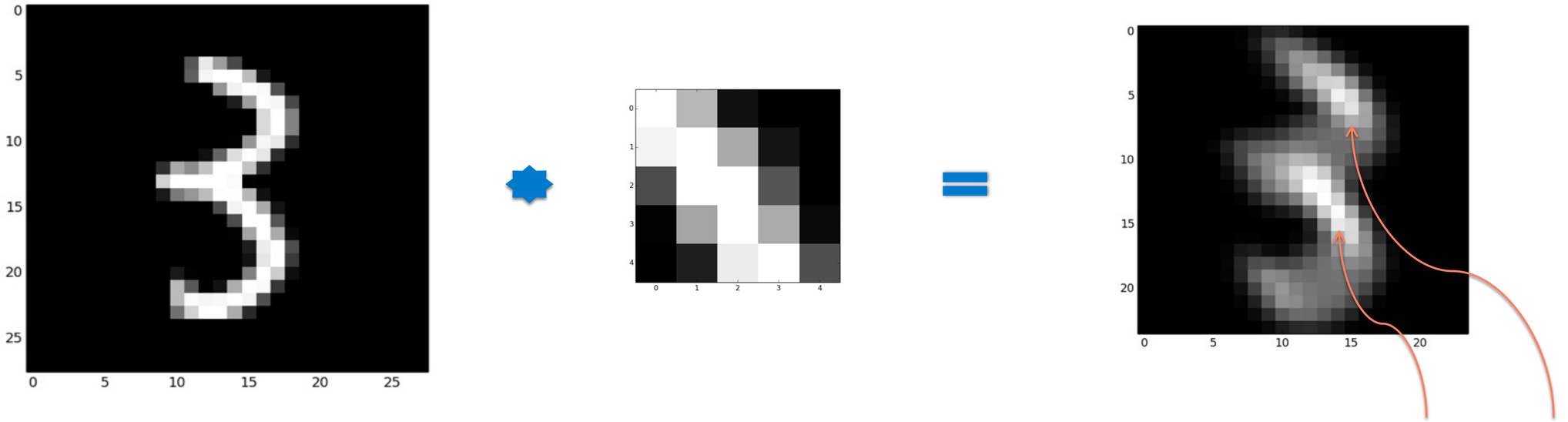
Convolutional layer



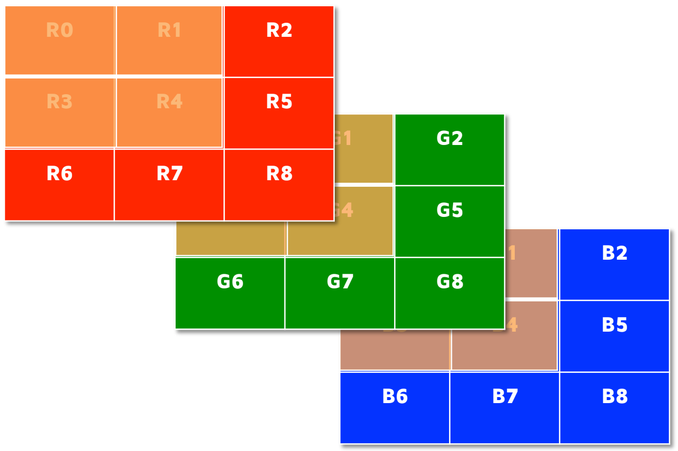
0 x 0 +
1 x 1 +
2 x 3 +
3 x 4 = 19

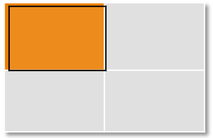
The weights are shared among the ● units.

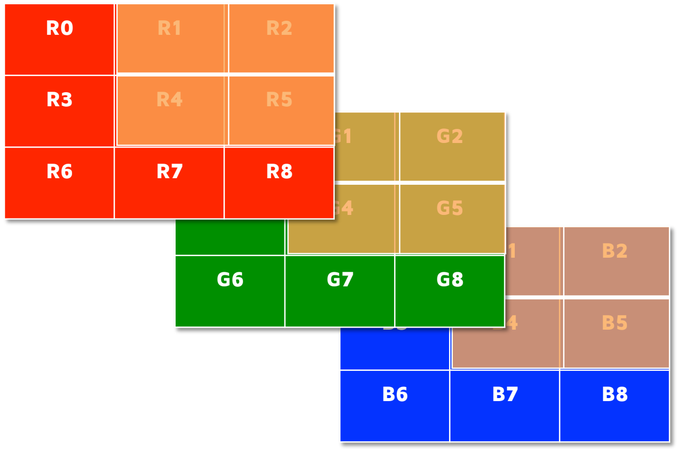
Recognizing patterns

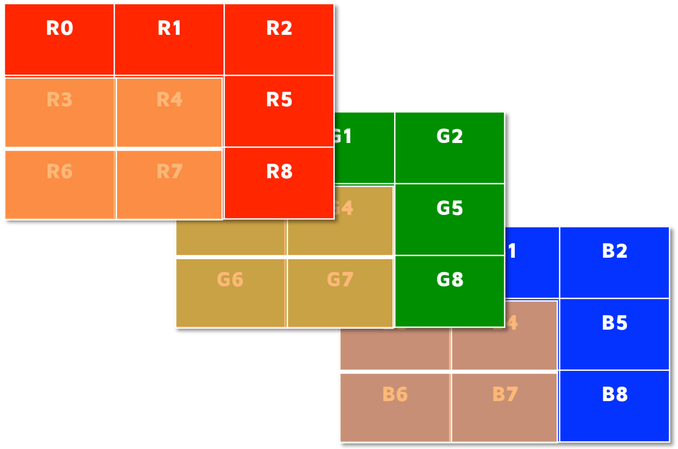


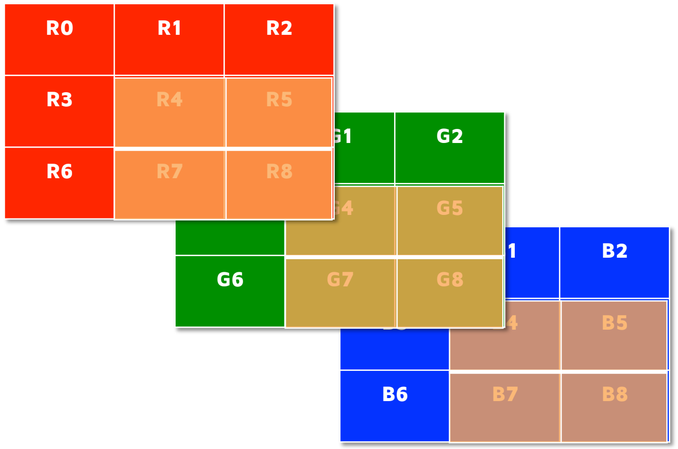
Detected the pattern!

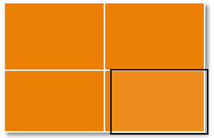

R0
R1
R2
R3
R4
R5
R6
R7
R8
G2
G5
G6
G7
G8
B2
B5
B6
B7
B8




R0
R1
R2
R3
R4
R5
R6
R7
R8
G2
G5
G6
G7
G8
B2
B5
B6
B7
B8


R0
R1
R2
R3
R4
R5
R6
R7
R8
G2
G5
G6
G7
G8
B2
B5
B6
B7
B8


R0
R1
R2
R3
R4
R5
R6
R7
R8
G1
G2
G4
G5
G6
G7
G8
1
B2
4
B5
B6
B7
B8



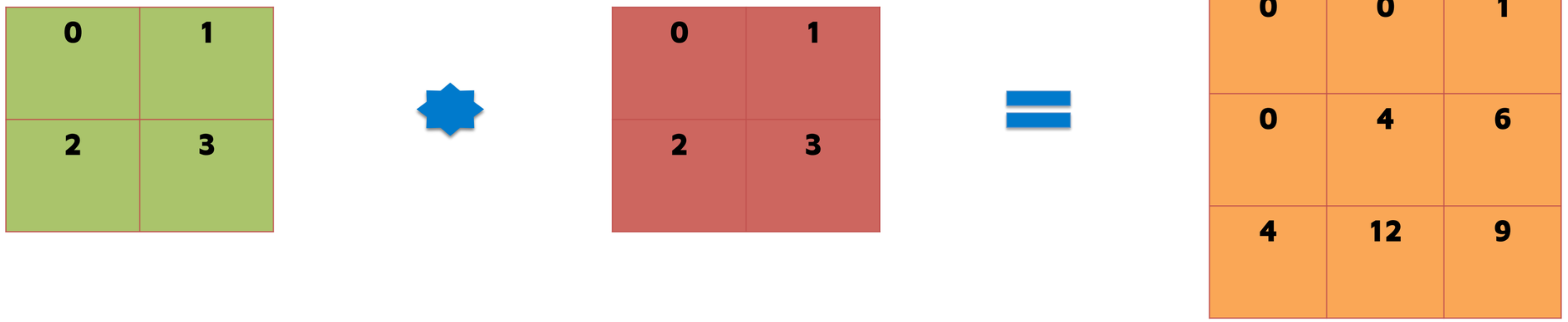
Max pooling

- Each element in the output is the maximum value within the pooling window

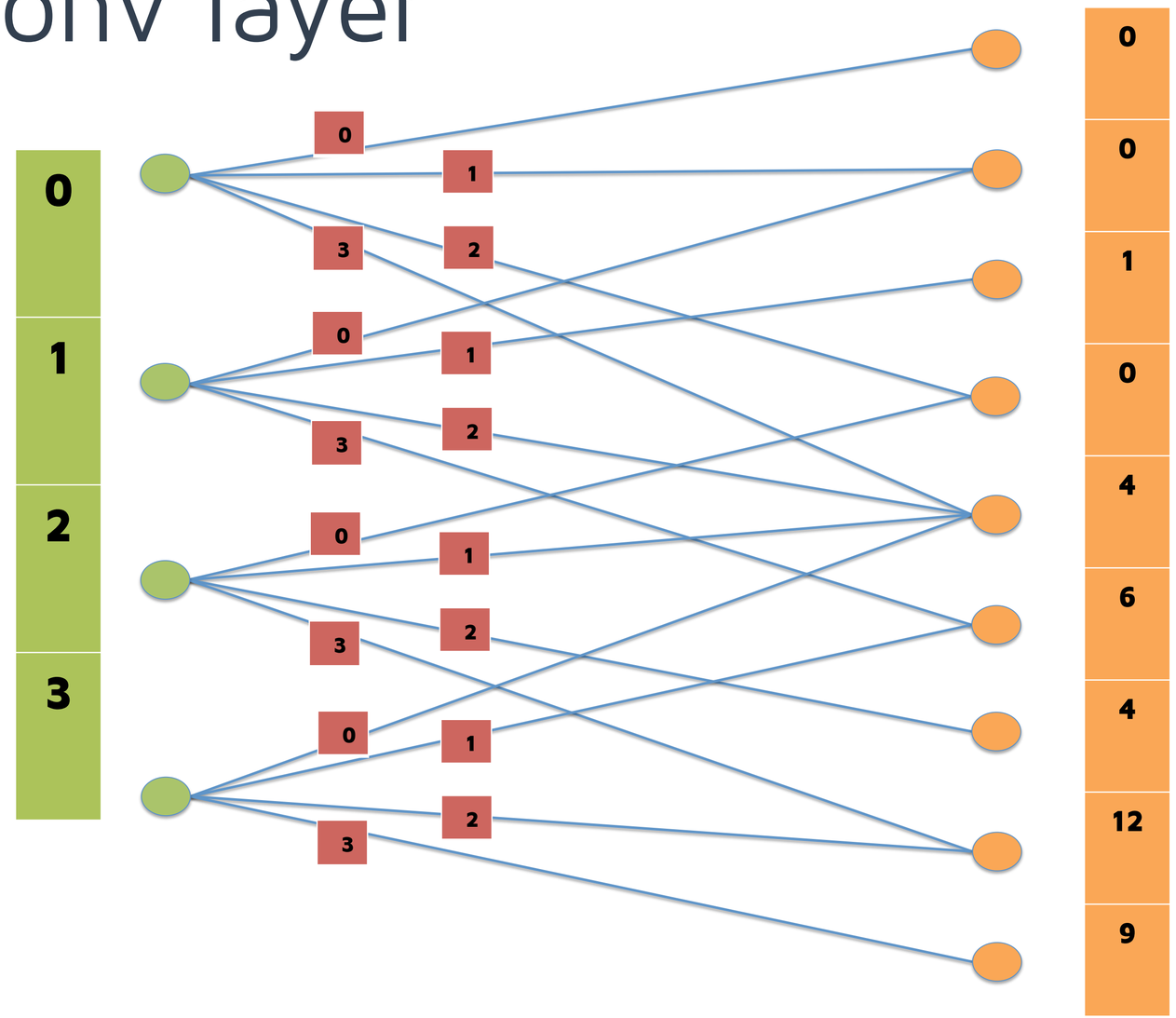
Deconvolution

- Ill posed problem, but we can approximate
- Scatter versus gather
- Used in convlayer backprop
- Equivalent to convolution with a flipped kernel on zero padded input
- Useful for convolutional autoencoders

Deconv layer



Deconv layer

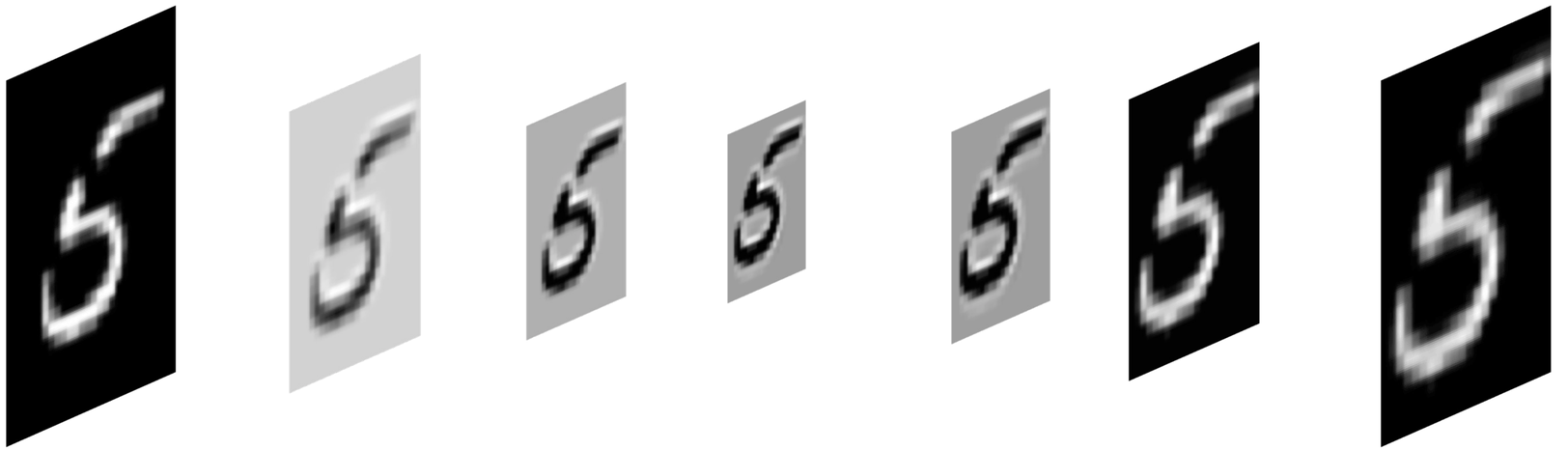


1 x 3 +

3 x 1 = 6

nervana

Convolutional autoencoder



Input	Conv1	Conv2	Conv3	Deconv1	Deconv2	Deconv3

RIGHT WHALE RECOGNITION

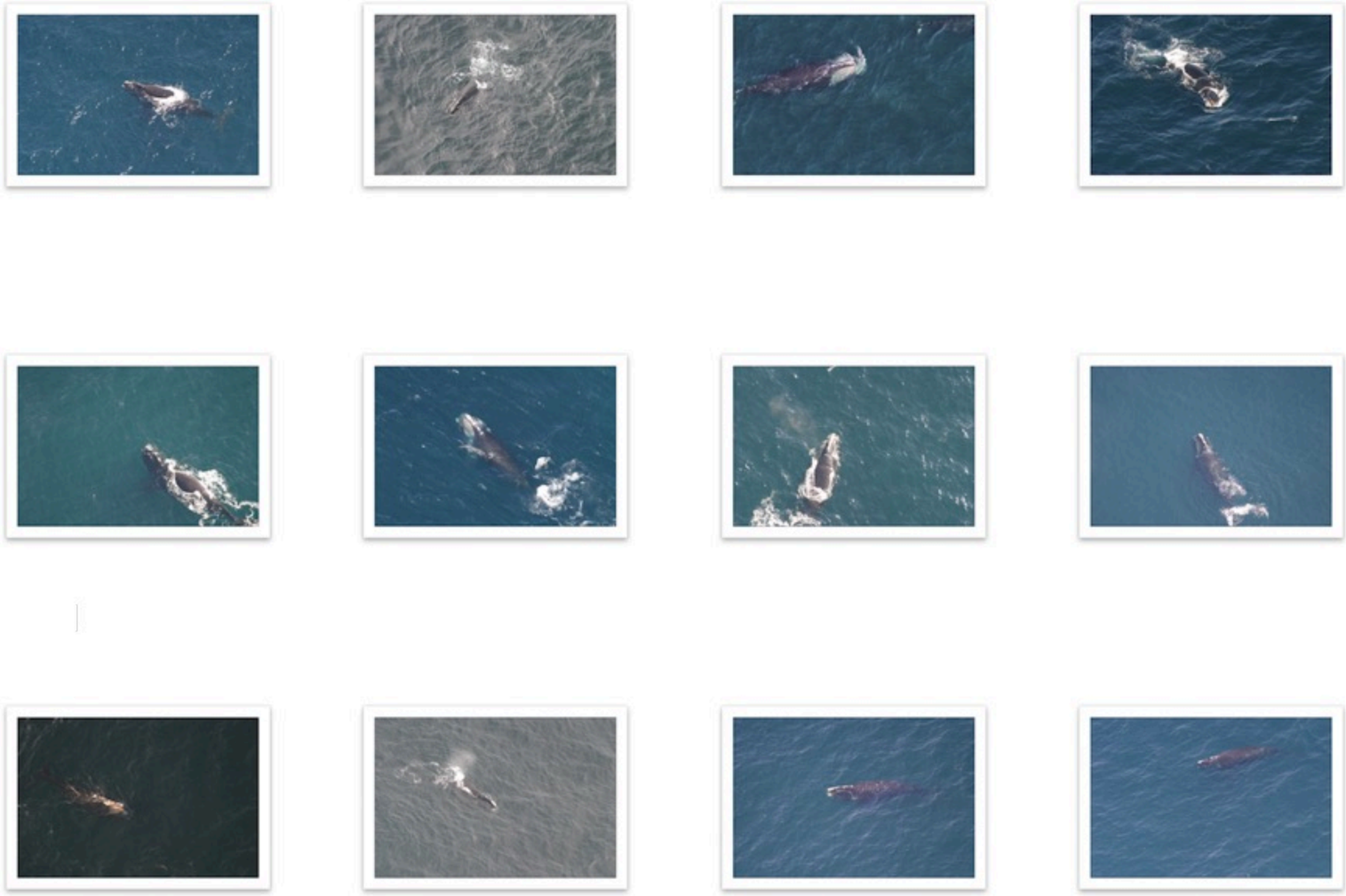
“Face” recognition for whales

- Identify whales in aerial photographs
- ~4500 labeled images, ~450 whales
- ~7000 test images
- Pictures taken over 10 years
- Automating the identification process will aid conservation efforts
- https://www.kaggle.com/c/noaa-right-whale-recognition
- $10,000 prize pool

Right whales

- One of the most endangered whales
- Fewer than 500 North Atlantic right whales left
- Hunted almost to extinction
- Makes a V shaped blow
- Has the largest testicle in the animal kingdom
- Eats 2000 pounds of plankton a day

“All y’all look alike!”



New England Aquarium Special Section:

The North Atlantic Right Whale Catalog

Scroll Through All Whales | Search For Individual Whales | Return to Right Whale Projects

41 of 696 whales

<< Previous **Catalog No:** 1102 Go Next >>

Whale Summary

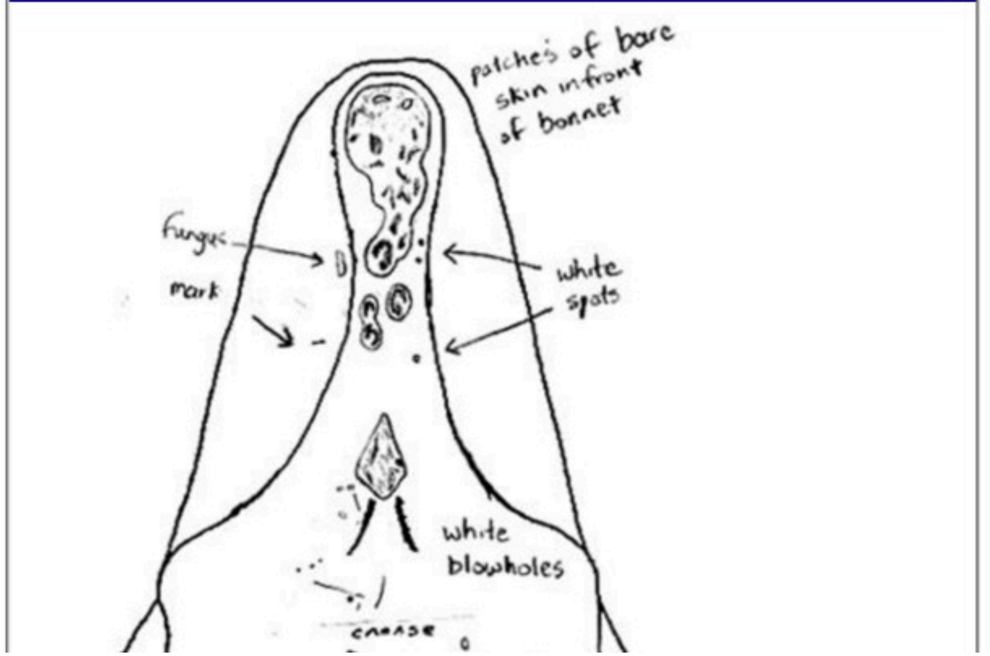
Catalog No:	1102	**Whale Name:**	CHURCHILL	**Sex:**	Male
Calving Female:	No	**Mother:**		**Last Year Seen:**	2001

Year of Birth:

Death Year:

Whale Composite

View Sightings



Whale Images (Click an image for a larger view)

Body Part Filter: <All Body Parts>

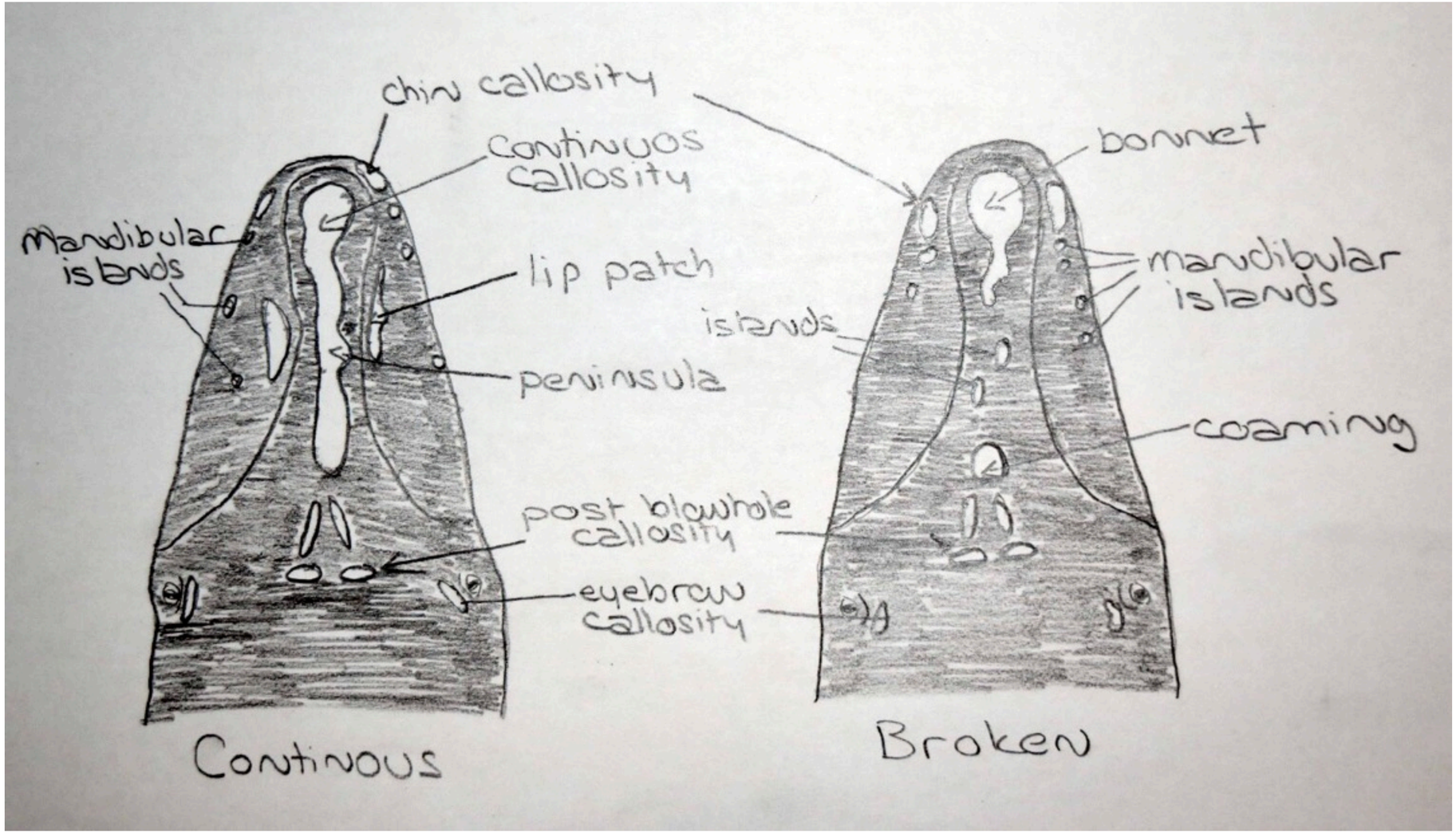
2001/06/09 CCS

2001/06/09 CCS

2001/06/09 CCS

2001/06/09 CCS

Source: http://rwcatalog.neaq.org/



Source: https://teacheratsea.files.wordpress.com/2015/05/img_2292.jpg

Brute force approach

Churchill

Quasimodo

Aphrodite

?

*Not actual names

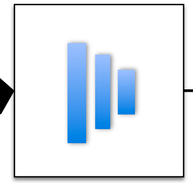
A better method

Quasimodo

Aphrodite

Churchill

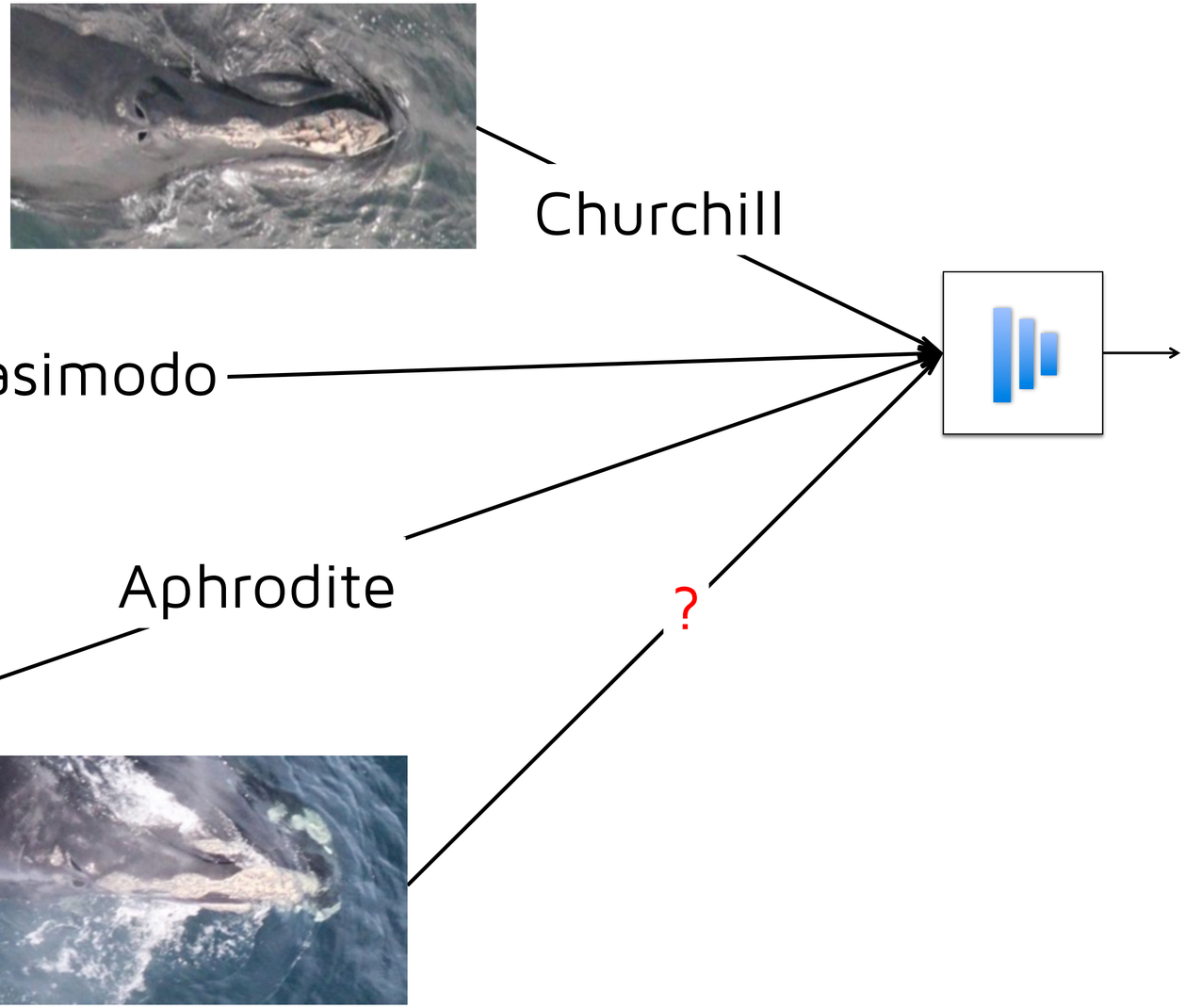
?



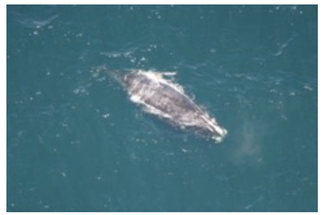
Object localization

- Many approaches in the literature
- Overfeat (http://arxiv.org/pdf/1312.6229v4.pdf)
- R-CNN (http://arxiv.org/pdf/1311.2524v5.pdf)

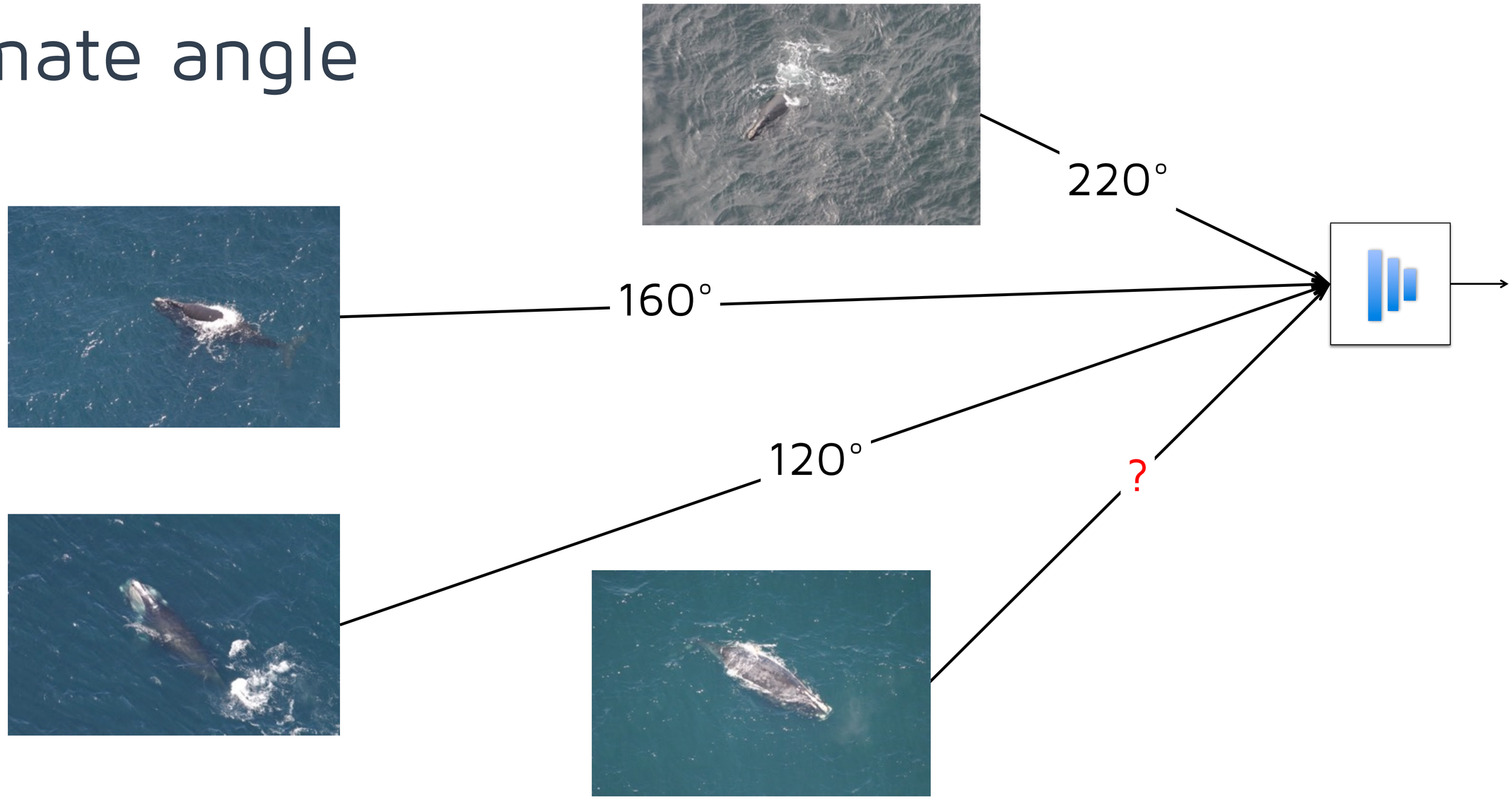
Even better!



Getting mugshots

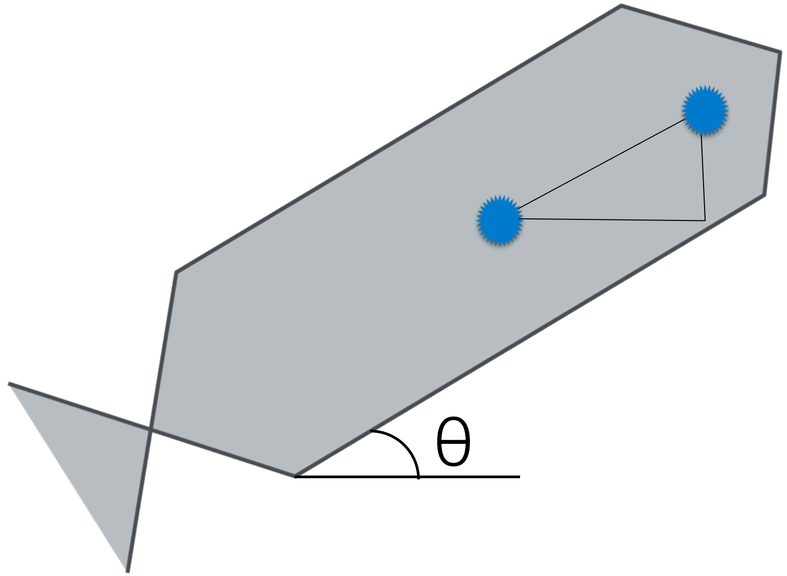
- How to go from  to  ?
- Training set can be manually labeled
- No manual operations allowed on test set!
- Estimate the heading (angle) of the whale using a CNN?

Estimate angle

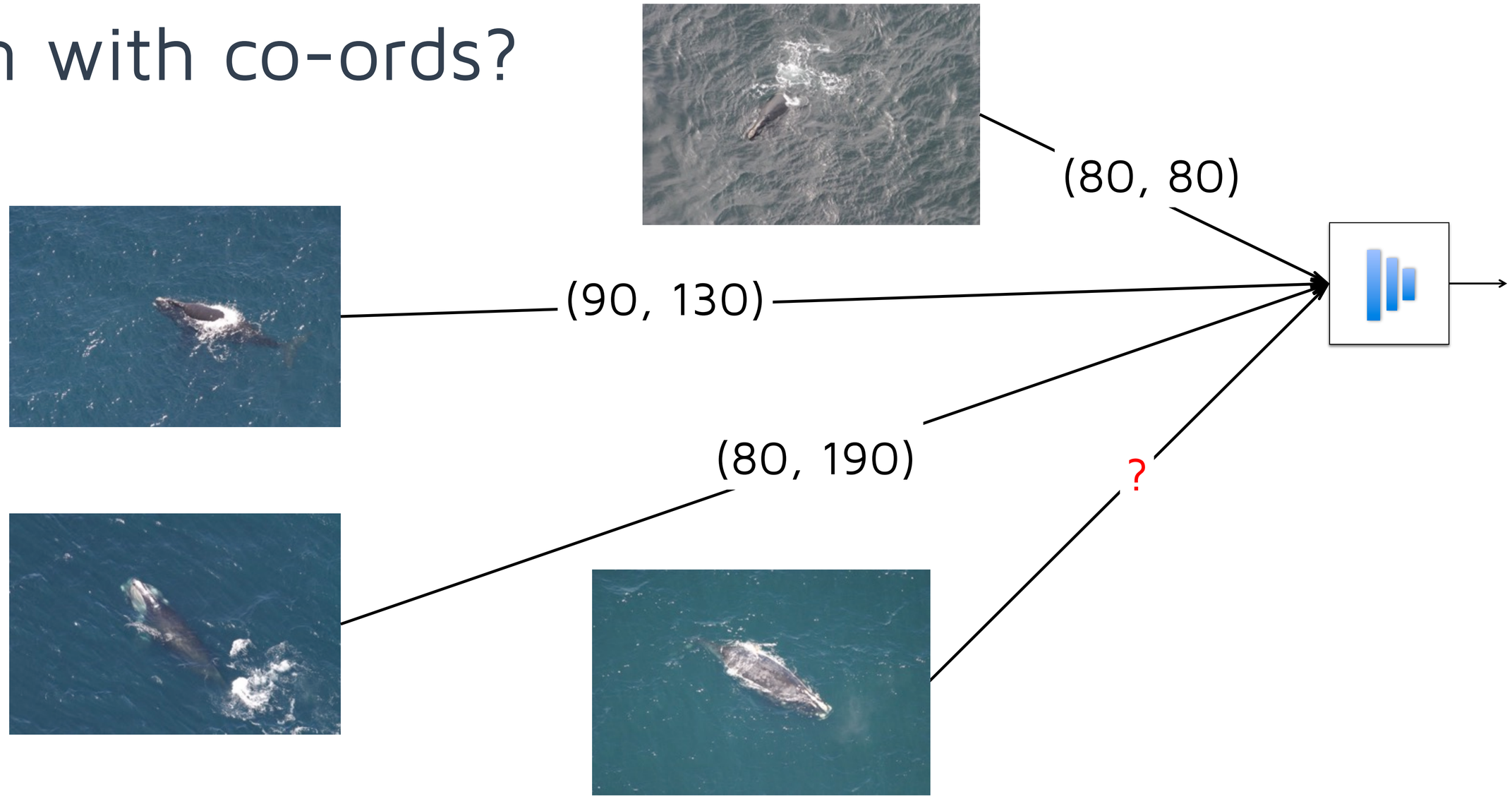


An easier way to estimate angle

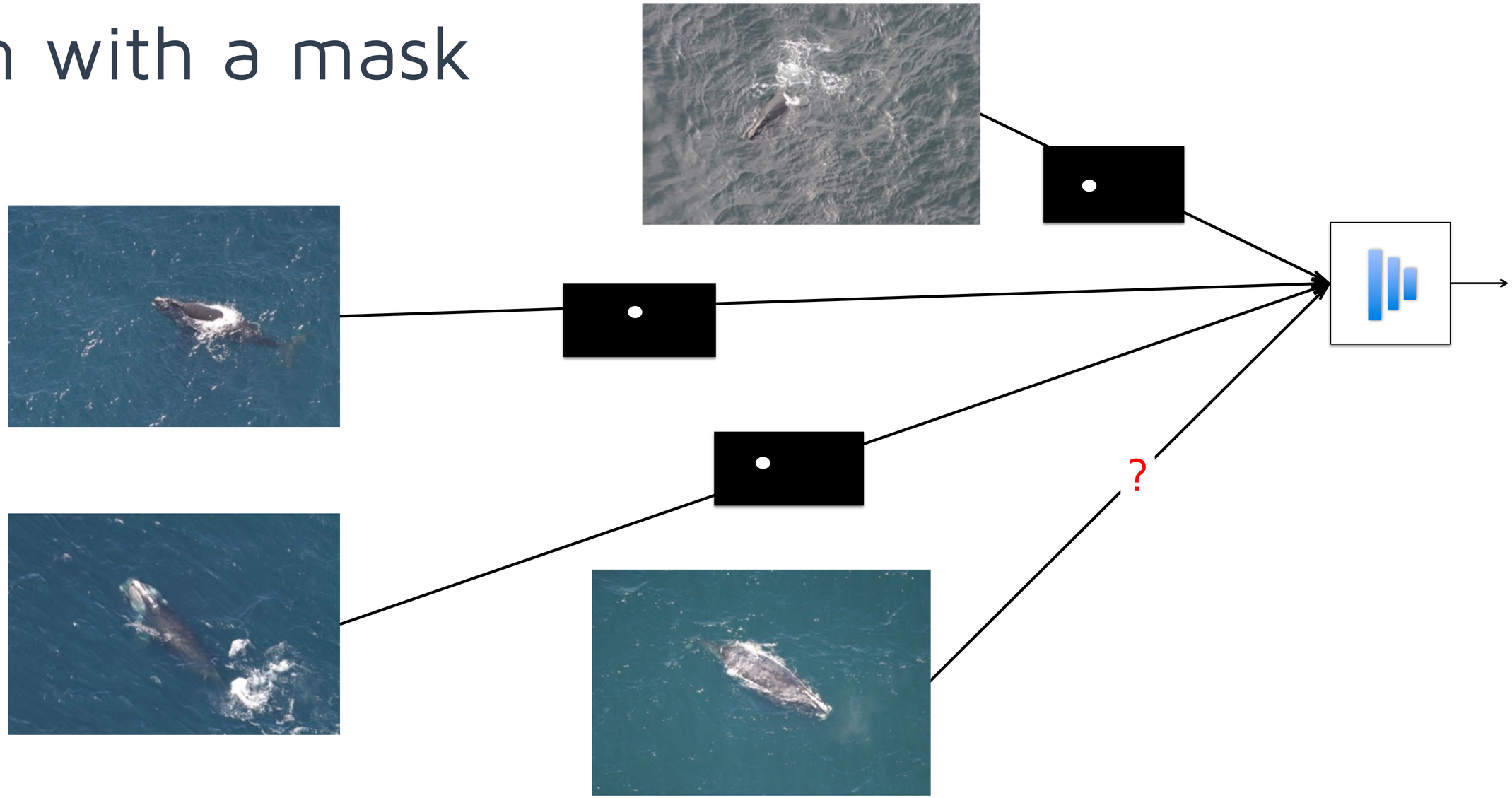
- Find two points along the whale's body
- $\theta = \arctan((y1 - y2) / (x1 - x2))$
- But how do you label the test images?



Train with co-ords?



Train with a mask



Code for convolutional encoder

```
init = Gaussian(scale=0.1)
opt = Adadelta(decay=0.9)
common = dict(init=init, batch_norm=True, activation=Rectlin())

layers = []
nchan = 128
layers.append(Conv((2, 2, nchan), strides=2, **common))
for idx in range(16):
    layers.append(Conv((3, 3, nchan), **common))
    if nchan > 16:
        nchan /= 2
for idx in range(15):
    layers.append(Deconv((3, 3, nchan), **common))
layers.append(Deconv((4, 4, nchan), strides=2, **common))
layers.append(Deconv((3, 3, 1), init=init))

cost = GeneralizedCost(costfunc=SumSquared())
mlp = Model(layers=layers)
callbacks = Callbacks(mlp, train, eval_set=val, **args.callback_args)
mlp.fit(train, optimizer=opt, num_epochs=args.epochs, cost=cost, callbacks=callbacks)
```

Code for classifier

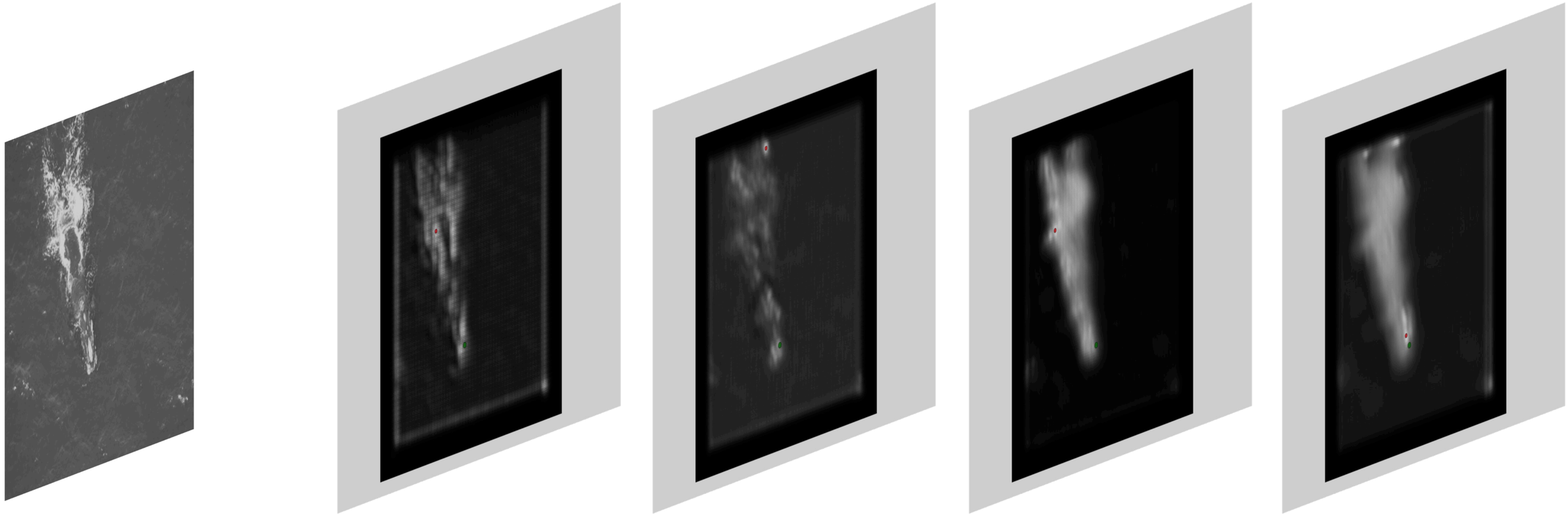
```
init = Gaussian(scale=0.01)
opt = Adadelta(decay=0.9)
common = dict(init=init, batch_norm=True, activation=Rectlin())

layers = []
nchan = 64
layers.append(Conv((2, 2, nchan), strides=2, **common))
for idx in range(6):
    if nchan > 1024:
        nchan = 1024
    layers.append(Conv((3, 3, nchan), strides=1, **common))
    layers.append(Pooling(2, strides=2))
    nchan *= 2
layers.append(DropoutBinary(keep=0.5))
layers.append(Affine(nout=447, init=init, activation=Softmax()))

cost = GeneralizedCost(costfunc=CrossEntropyMulti())
mlp = Model(layers=layers)
callbacks = Callbacks(mlp, train, eval_set=val, **args.callback_args)
mlp.fit(train, optimizer=opt, num_epochs=args.epochs, cost=cost, callbacks=callbacks)
```

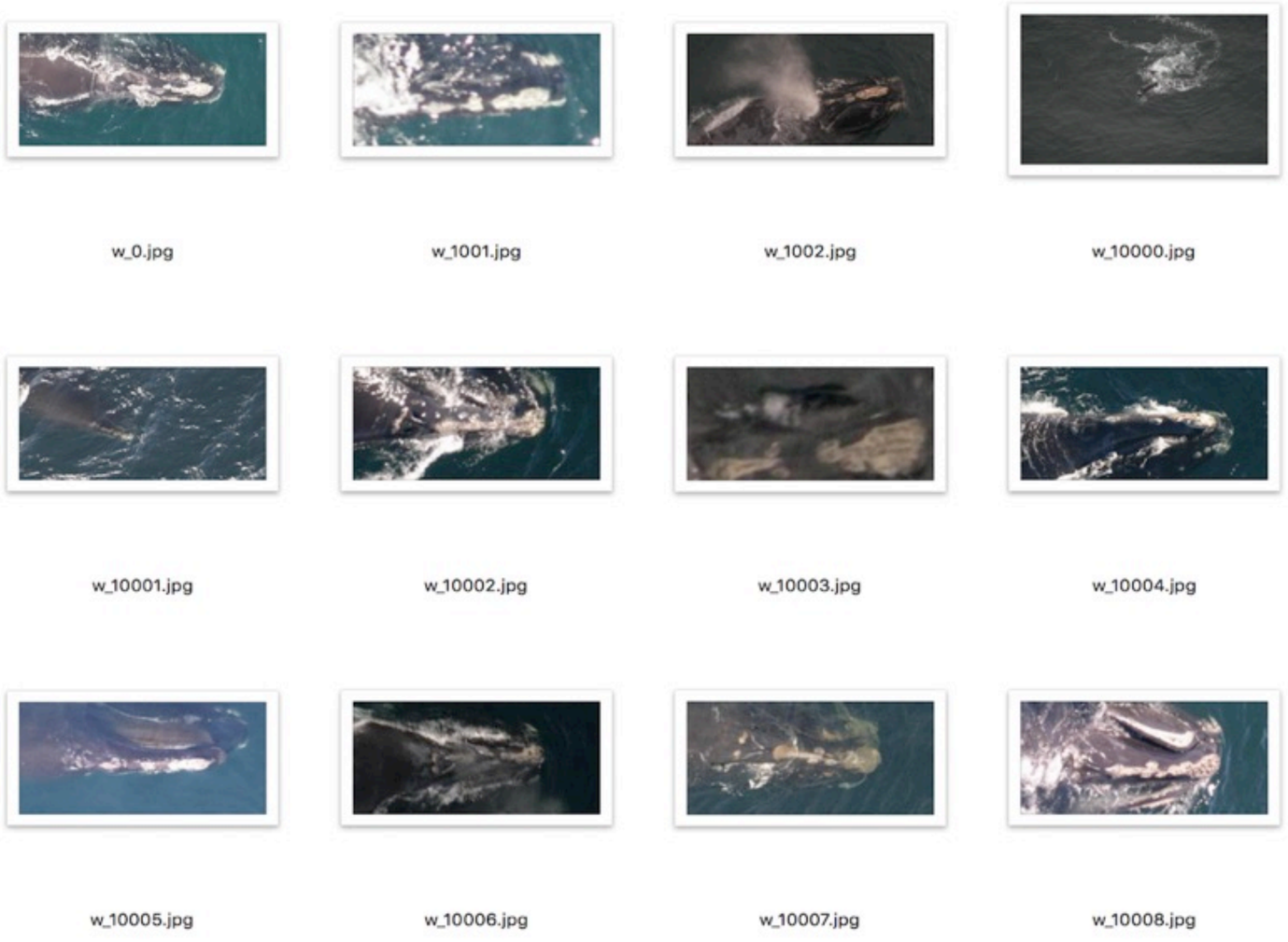
Results –heatmaps

Input	epoch 0	epoch 2	epoch 4	epoch 6



Prediction indicated by ●

Results –sample crops from test set



nervana

$10,000 • 178 teams

Right Whale Recognition

Merger and 1st Submission Deadline

Thu 27 Aug 2015 — Thu 7 Jan 2016 (52 days to go)

Dashboard

Public Leaderboard - Right Whale Recognition

This leaderboard is calculated on approximately 49% of the test data.
The final results will be based on the other 51%, so the final standings may be different.

See someone using multiple accounts?
Let us know.

#	Δ1w	Team Name * in the money	Score	Entries	Last Submission UTC (Best – Last Submission)
1	—	deepsense.io *	3.70568	19	Fri, 13 Nov 2015 22:03:45 (-7.2d)
2	—	Shen(Deep)-Jing(Whale) *	3.78693	14	Wed, 28 Oct 2015 13:31:05
3	—	Whalearning *	3.85244	26	Fri, 13 Nov 2015 07:11:30 (-45.2d)
4	new	**Anil Thomas**	4.09338	1	Mon, 16 Nov 2015 06:12:13
5	↓1	The Drunken Whalers	4.12315	43	Fri, 30 Oct 2015 15:24:07
6	↓1	toshi_k	4.50391	33	Sun, 08 Nov 2015 12:03:13 (-11.1d)
7	↓1	stas_sl	4.55536	12	Fri, 11 Sep 2015 10:03:46 (-4d)
8	↓1	Paul H	4.56856	17	Sun, 13 Sep 2015 00:31:54 (-4d)
9	↓1	Mike	4.58045	17	Fri, 25 Sep 2015 08:29:33
10	—	One Old Dog	4.62447	129	Mon, 16 Nov 2015 00:38:07

Acknowledgements

- NOAA Fisheries
- Kaggle
- Developers of sloth
- Playground Global

nervana

MAKING MACHINES SMARTER.