Teaching Categories to Human Learners with Visual Explanations

Oisin Mac Aodha

University of Edinburgh

www.oisin.info
@oisinmacaodha

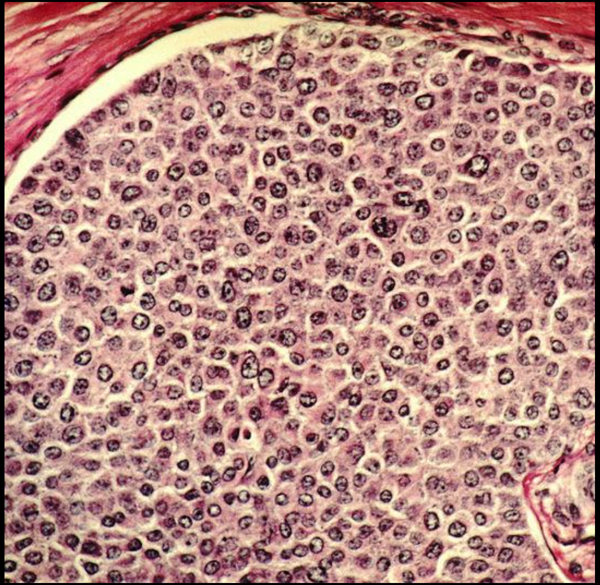
Can we design teaching algorithms that will enable humans to become better at visual categorization?

Why Visual Expertise?

What species?

Why Visual Expertise?

Cancerous?



Why Visual Expertise?

Poisonous?

Why Visual Expertise?

Forgery?

Challenges - 1 Visual Similarity

https://en.wikipedia.org/wiki/Grey_heron

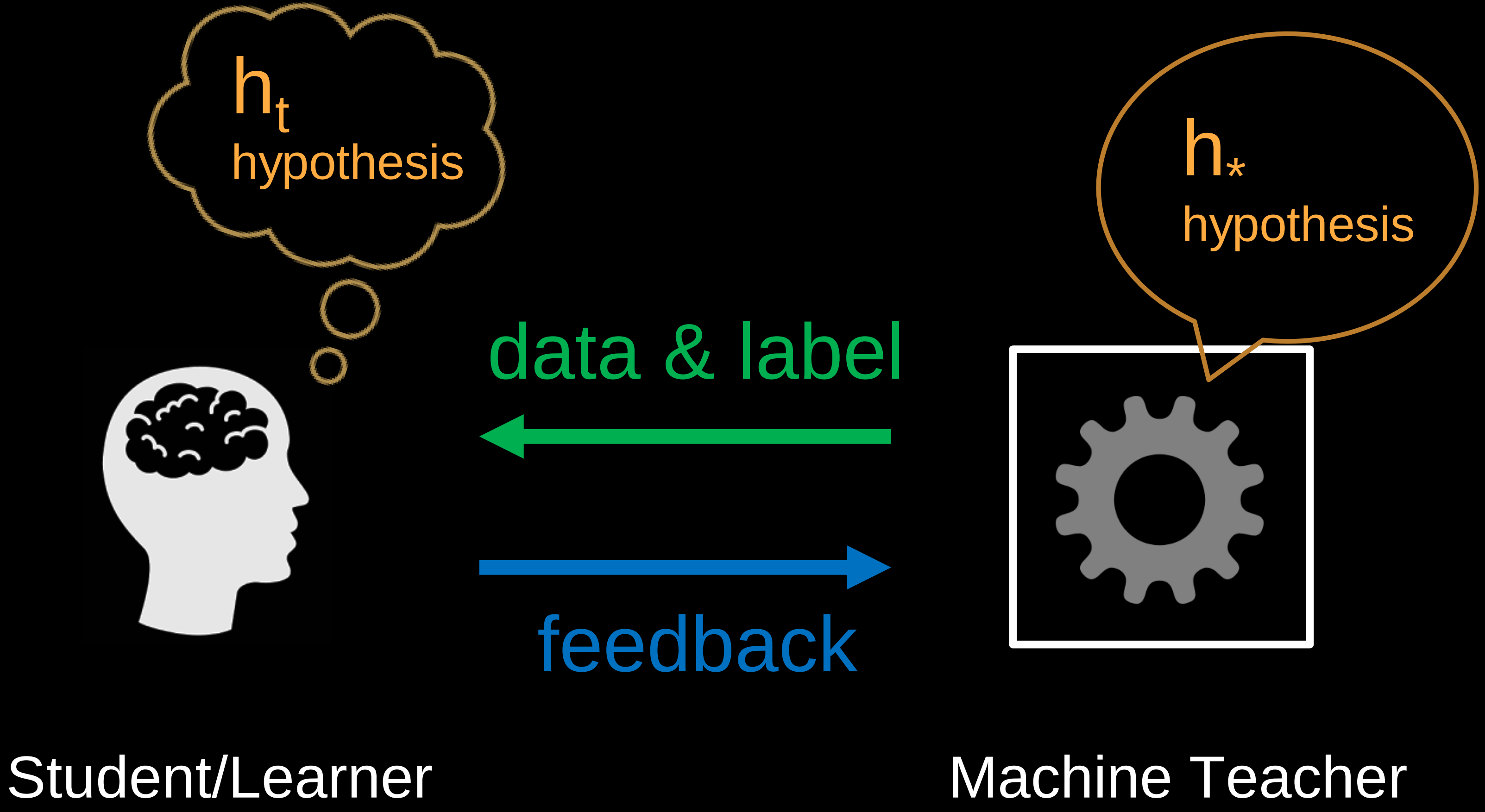
https://ebird.org/species/cocher1

Challenges - 2 Within Class Variation

Challenges - 3 “Attribution”

Which pixels “explain” the class label?

https://en.wikipedia.org/wiki/Grey_heron



h_t
hypothesis
h_*
hypothesis
data & label
feedback
Student/Learner
Machine Teacher

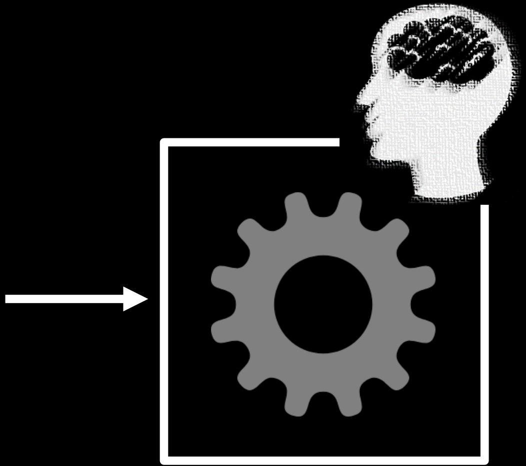
Teaching Visual Expertise

⋮

Set of images with class labels

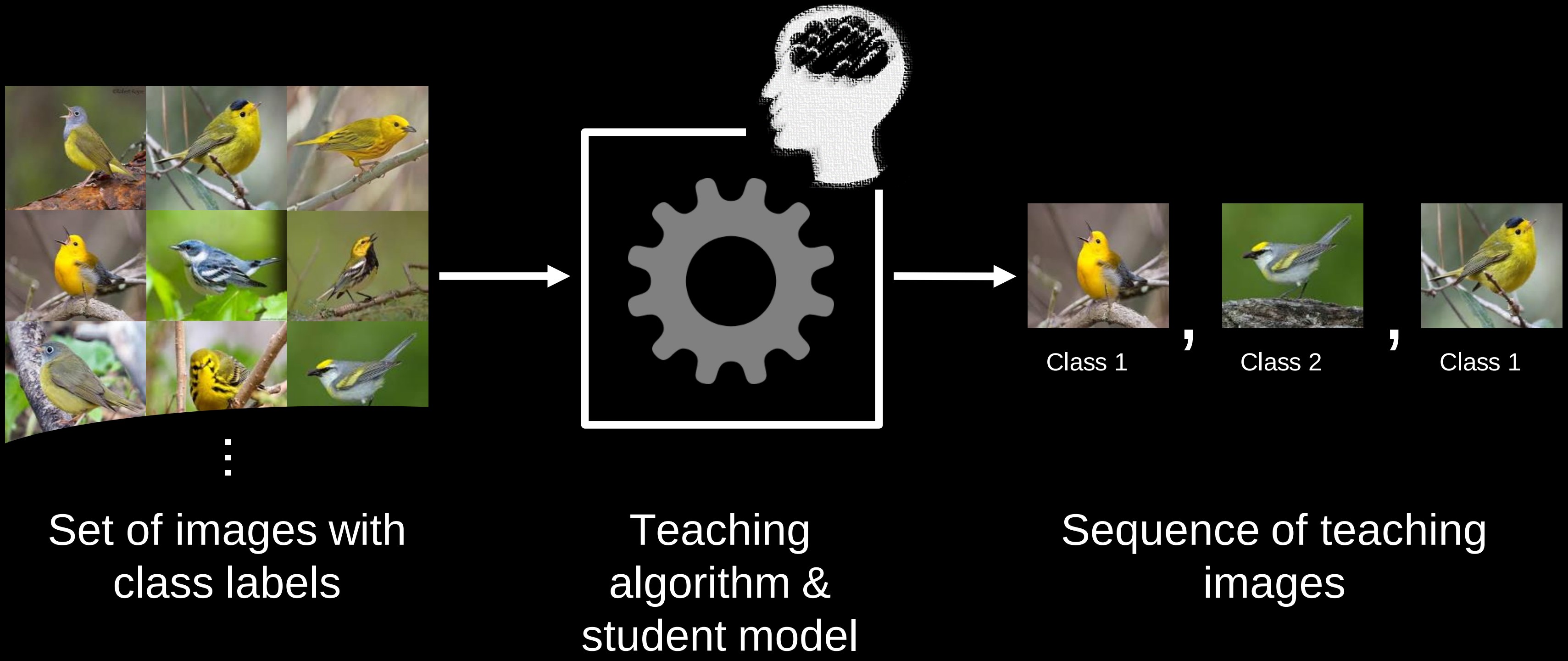
Teaching Visual Expertise

Set of images with class labels



Teaching algorithm & student model

Teaching Visual Expertise



Machine Teaching Landscape

Theoretical
Goldman & Kearns 1995
Zhu 2013
Chen et al. 2018
...

Spaced Repetition
Leitner 1972
Settles & Meeder 2016
Hunziker et al. 2019
Choffin et al. 2019
...

Decision Making
Bak et al. 2016
...

Visual Categories
Singla et al. 2014
Johns et al. 2015
Chen et al. 2018
...

Connecticut Warbler
or MacGillivray's Warbler

https://www.inaturalist.org/observations/9869215

Connecticut Warbler
~~or MacGillivray's Warbler~~

https://www.inaturalist.org/observations/9869215

Connecticut Warbler

MacGillivray's Warbler

https://www.inaturalist.org/observations/9869215
https://www.inaturalist.org/observations/3949369

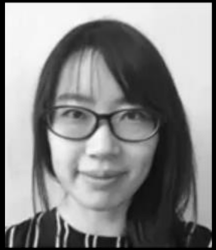
Connecticut Warbler

MacGillivray's Warbler

https://www.inaturalist.org/observations/9869215
https://www.inaturalist.org/observations/3949369

Teaching Categories to Human Learners with Visual Explanations

CVPR 2018

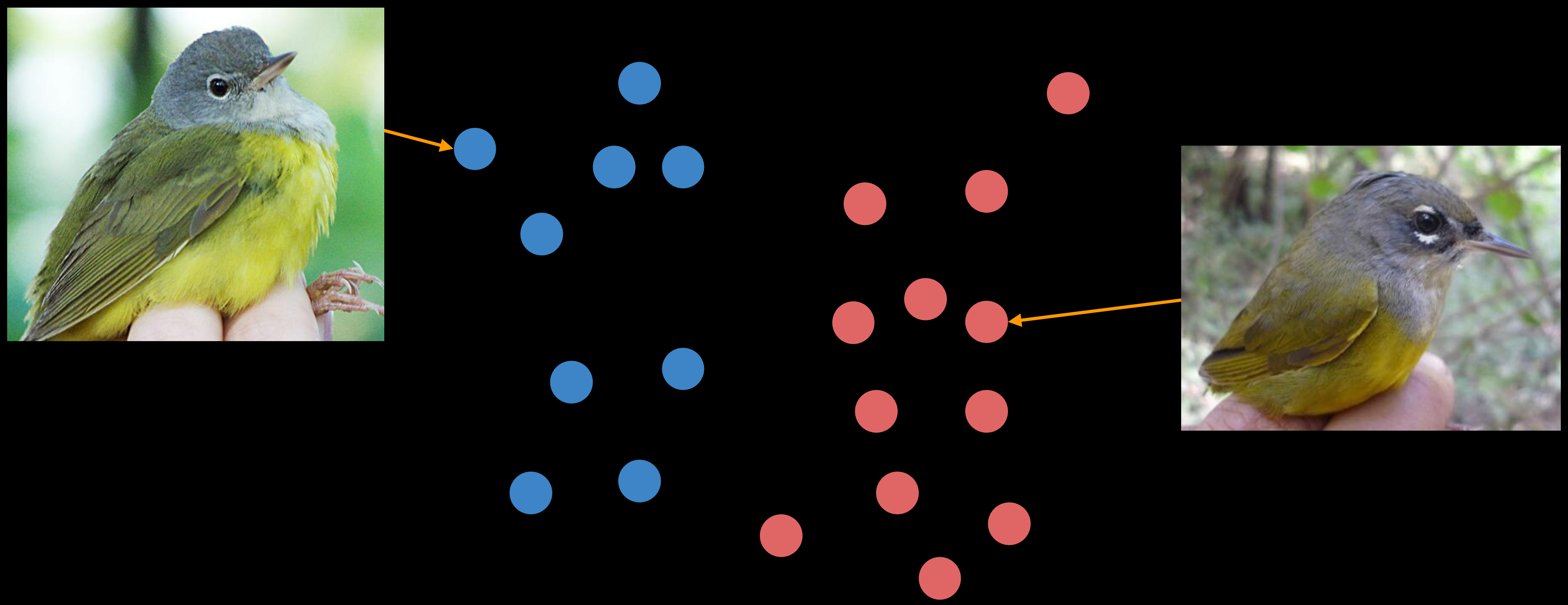
Yuxin Chen
Uni. of Chicago

Shihan Su
Caltech

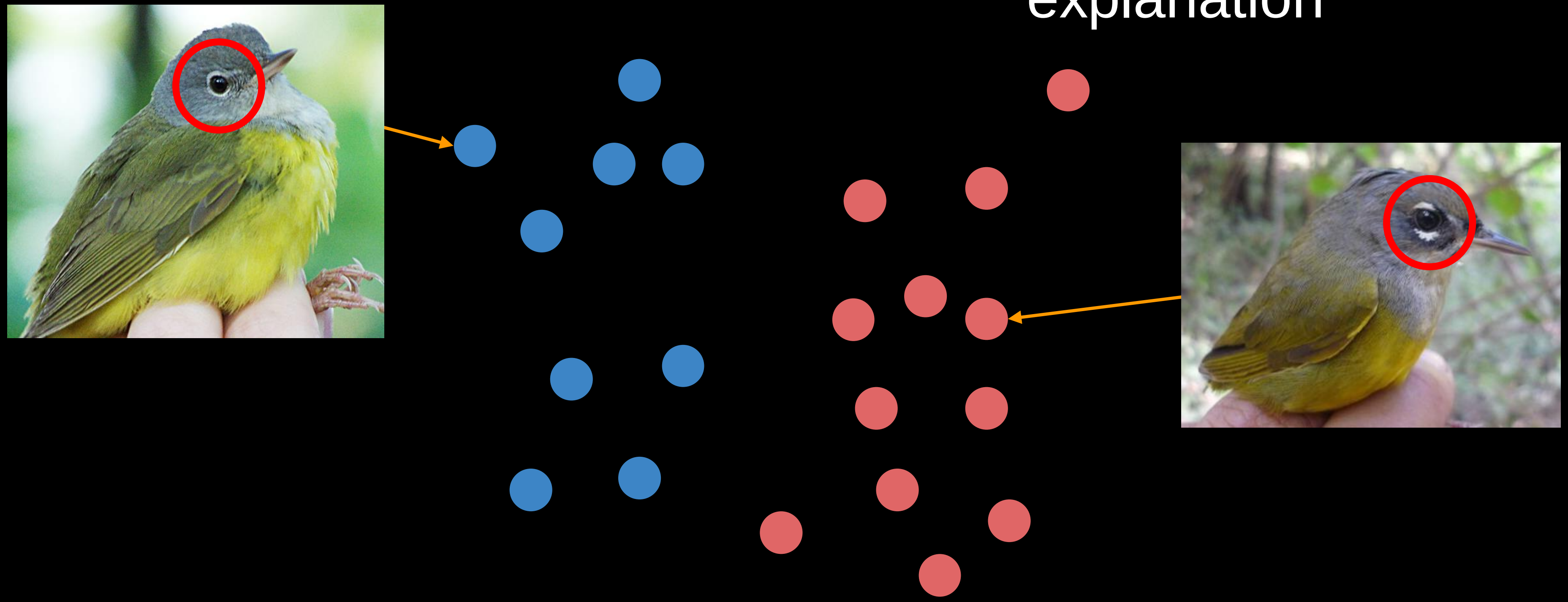
Pietro Perona
Caltech

Yisong Yue
Caltech

x is an image



e is an associated explanation



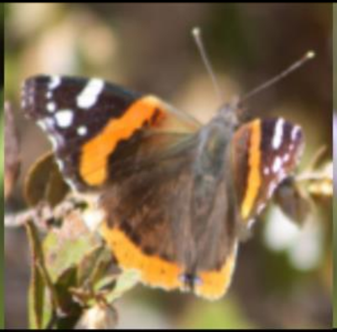
Visual “Explanations”

Monarch

Viceroy

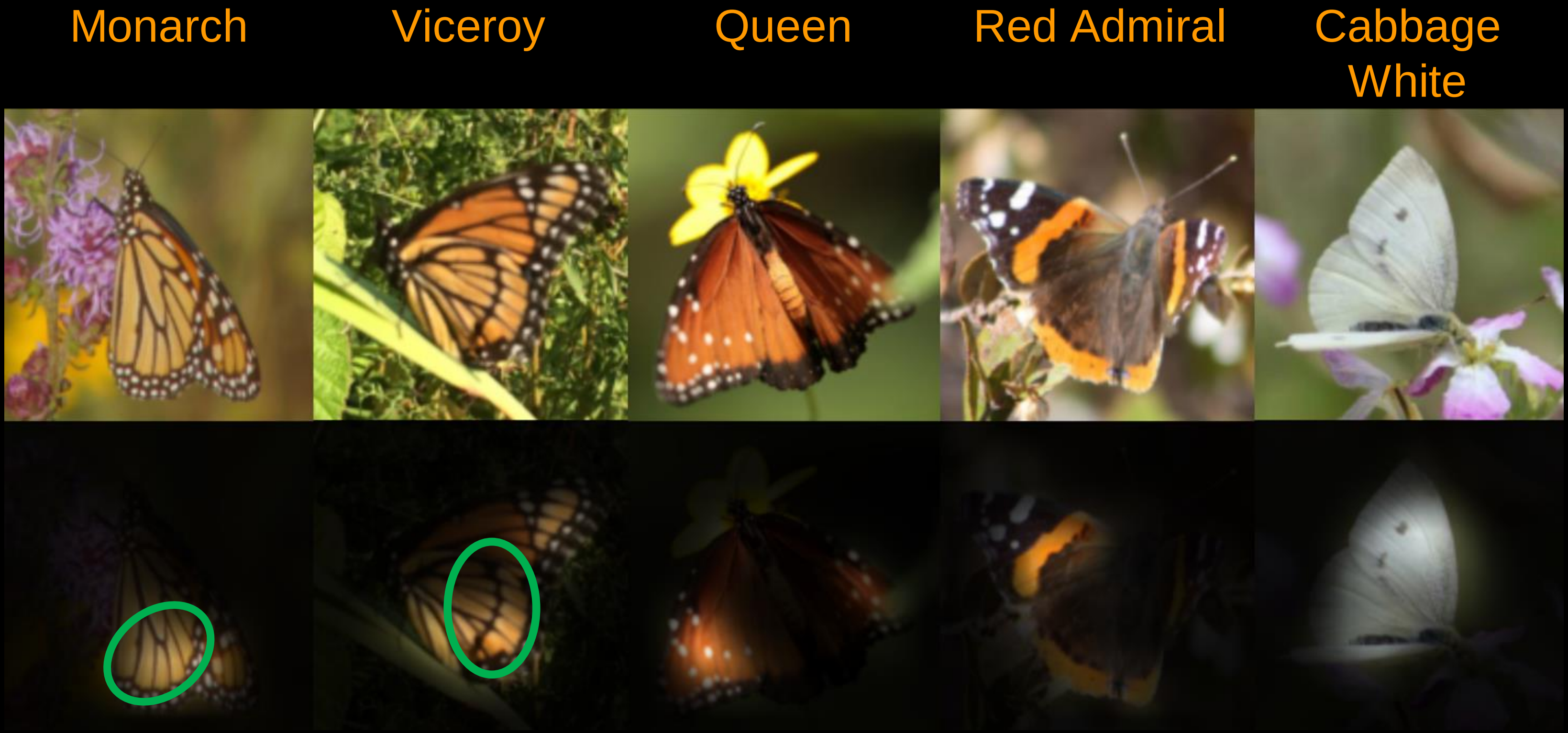
Queen

Red Admiral

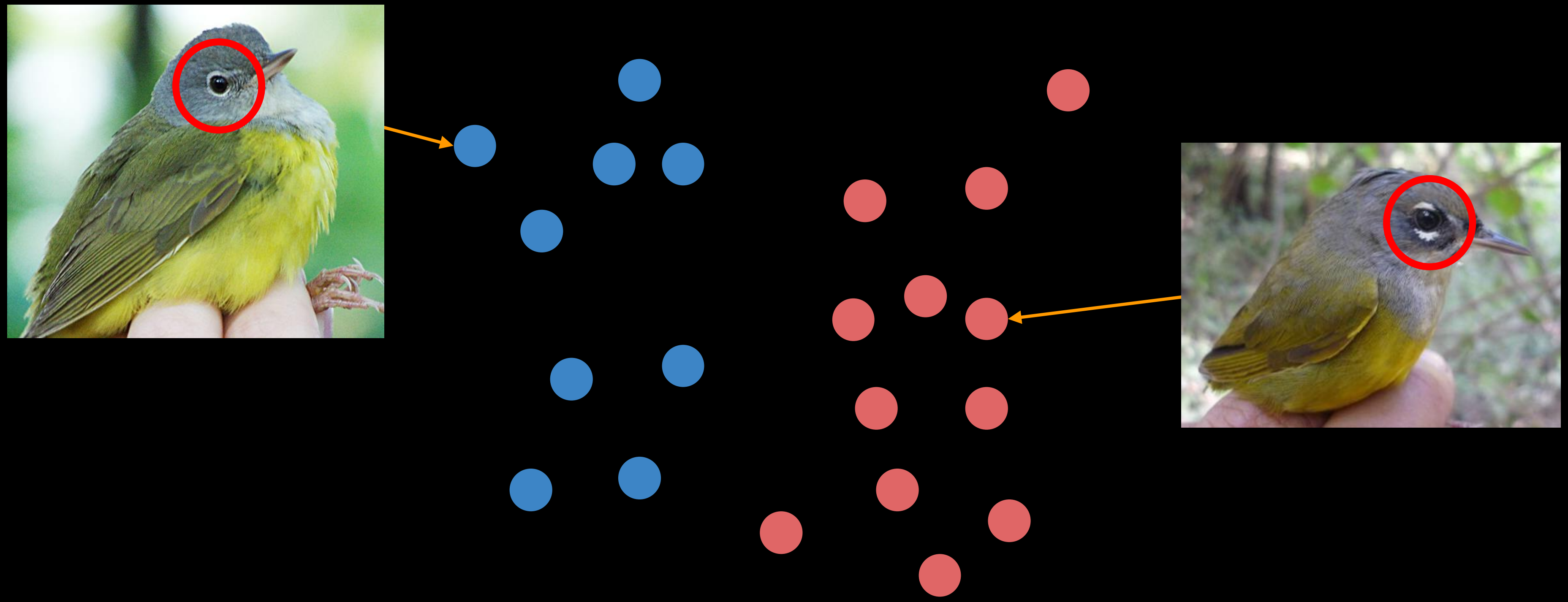


Cabbage White

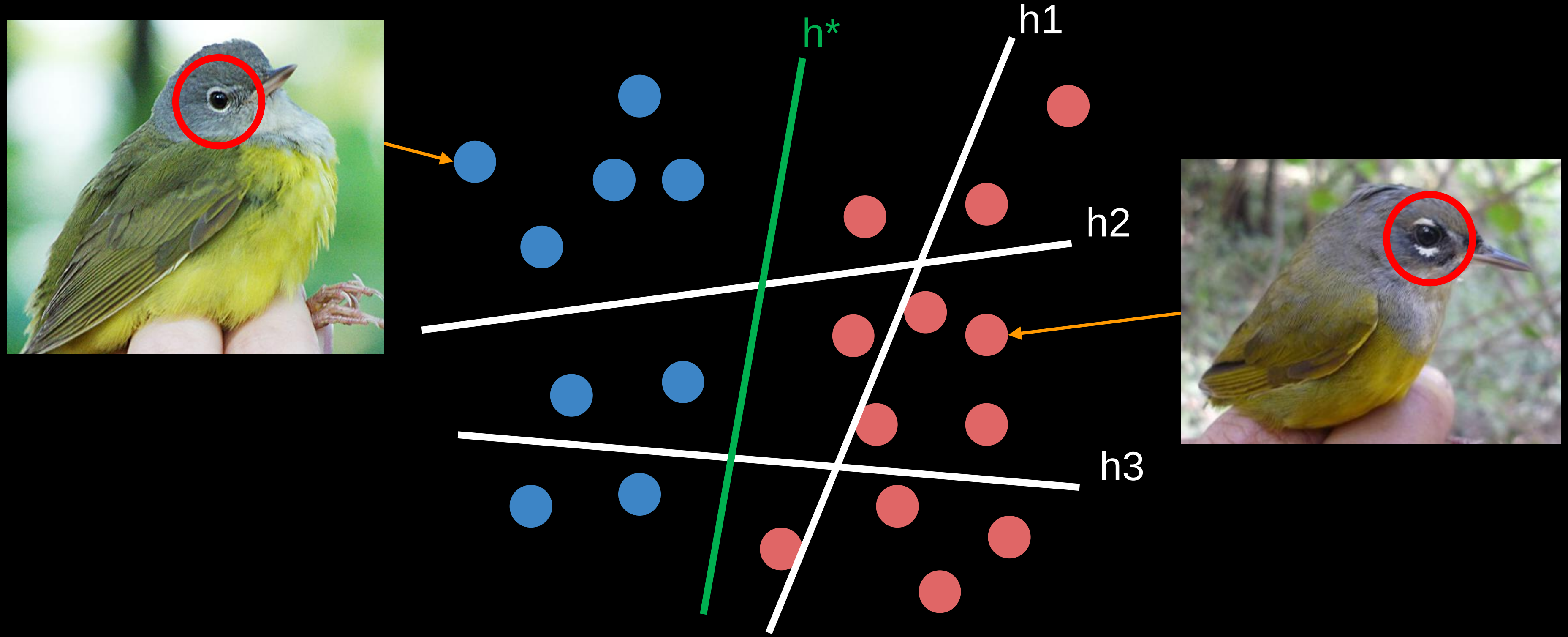
Visual “Explanations”

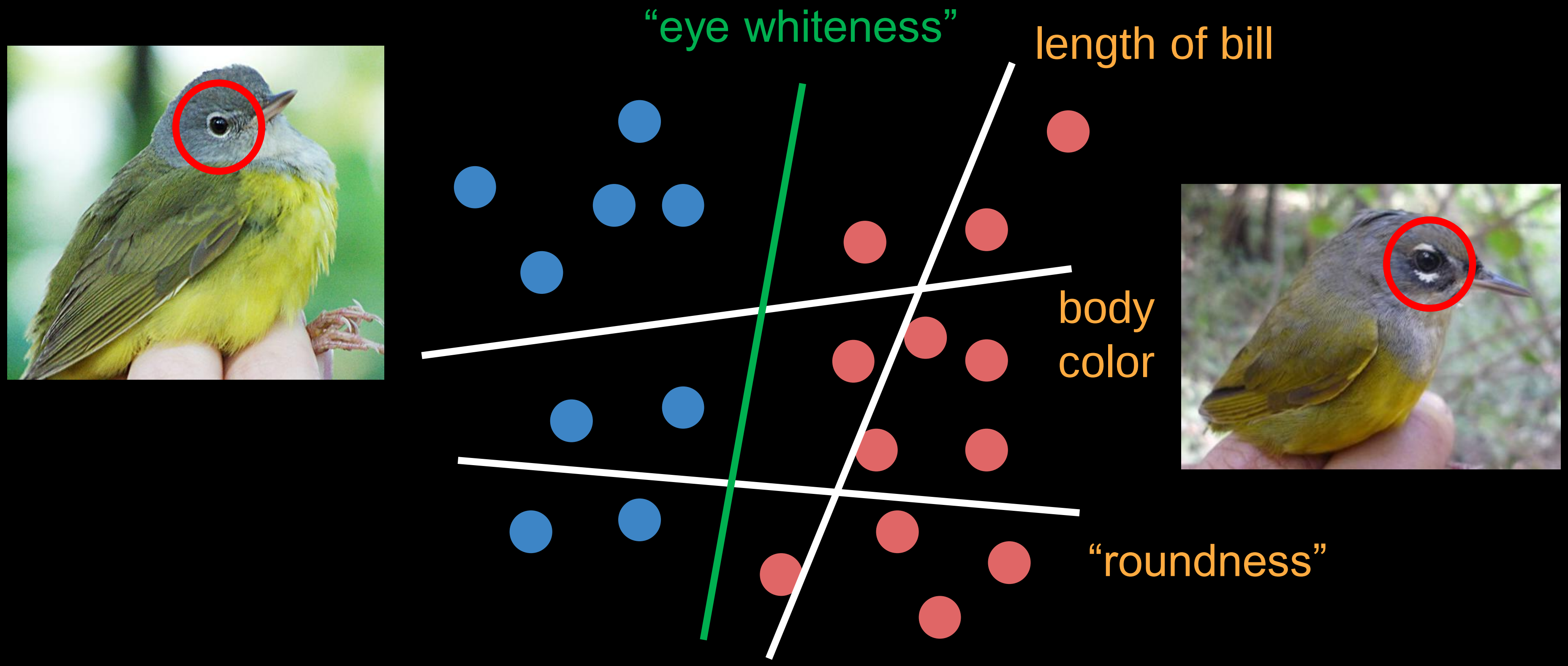


Learning Deep Features for Discriminative Localization **CVPR 2016**

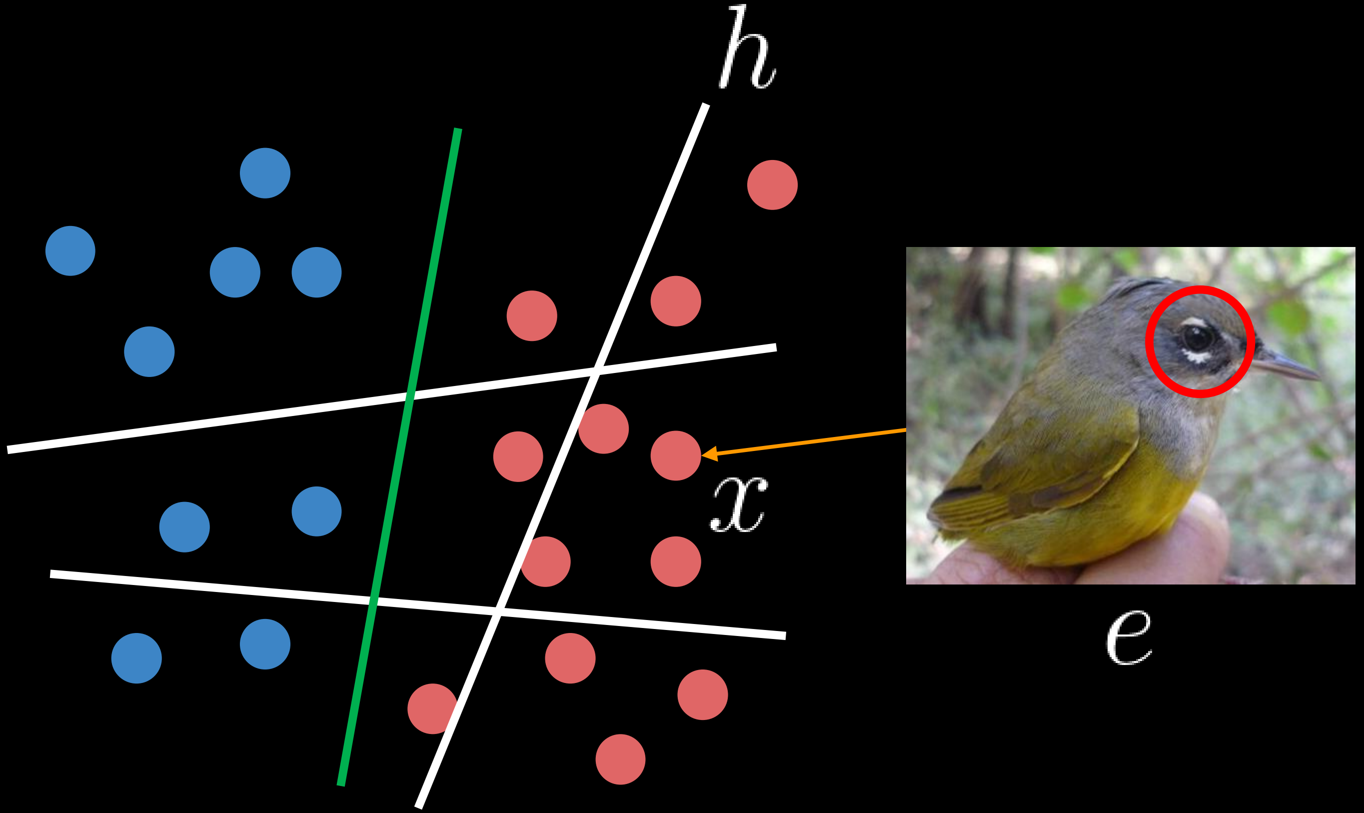


h is a hypothesis



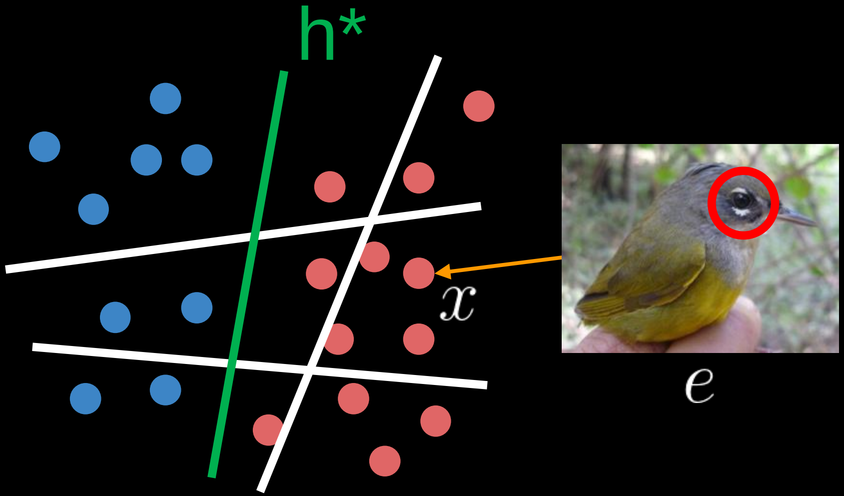


“eye whiteness”
length of bill
body color
“roundness”


h
x
e

How to Choose Teaching Set T to Teach h*?

$$T = \{(x_1, y_1, e_1), ..., (x_n, y_n, e_n)\}$$



Student Model

$$P(h|T)$$

Singla et al. Near-Optimally Teaching the Crowd to Classify **ICML 2014**

Student Model

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t)$$

"win stay, lose switch"

Singla et al. Near-Optimally Teaching the Crowd to Classify **ICML 2014**

Student Model

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t)$$

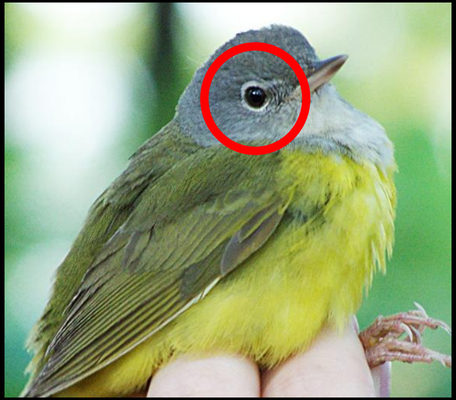
"win stay, lose switch"

$$S(y_t|h, x_t) = \begin{cases} 1 & \text{if} \quad y_t = \hat{y}_t^h \\ \frac{1}{1+\exp(-\alpha h(x_t) y_t)} & \text{otherwise} \end{cases}$$

Student Model - With Explanations

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t | h, x_t)$$

“Good”

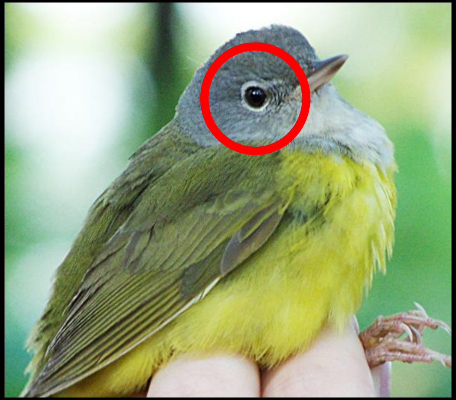


“Bad”

Student Model - With Explanations

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t) \prod_{x_t, e_t \in T} \boxed{(E(e_t)} D(x_t))$$

“Good”



“Bad”

Student Model - With Explanations

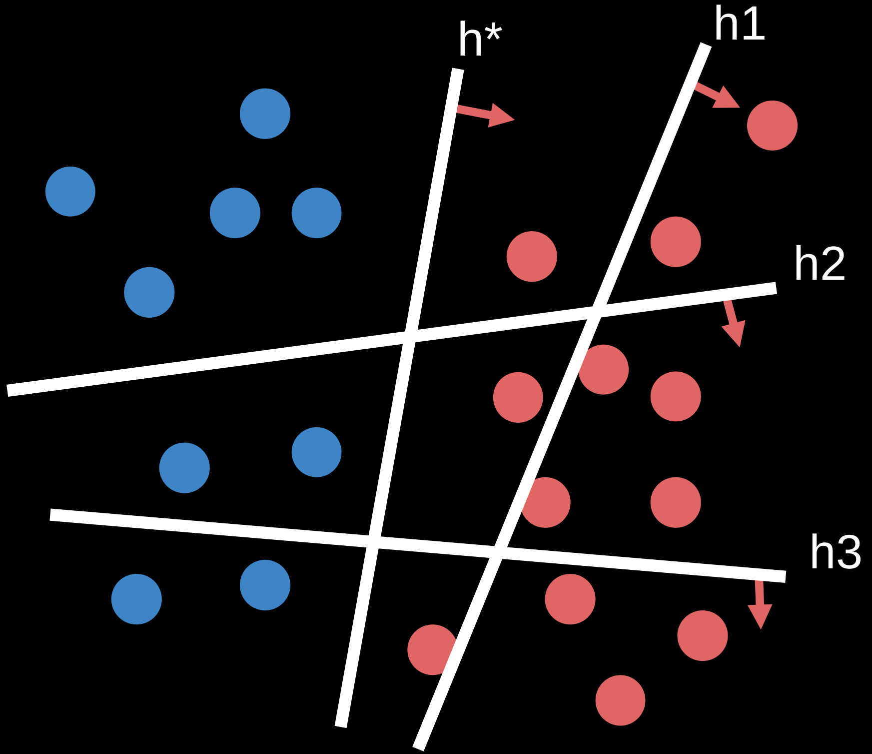
$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t) \prod_{x_t, e_t \in T} \boxed{(E(e_t)} D(x_t))$$

$$E(e_t) = \frac{1}{1 + \exp(-\beta \text{ diff}(e_t))}$$

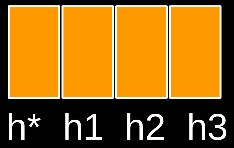
Selecting the Teaching Set *T*

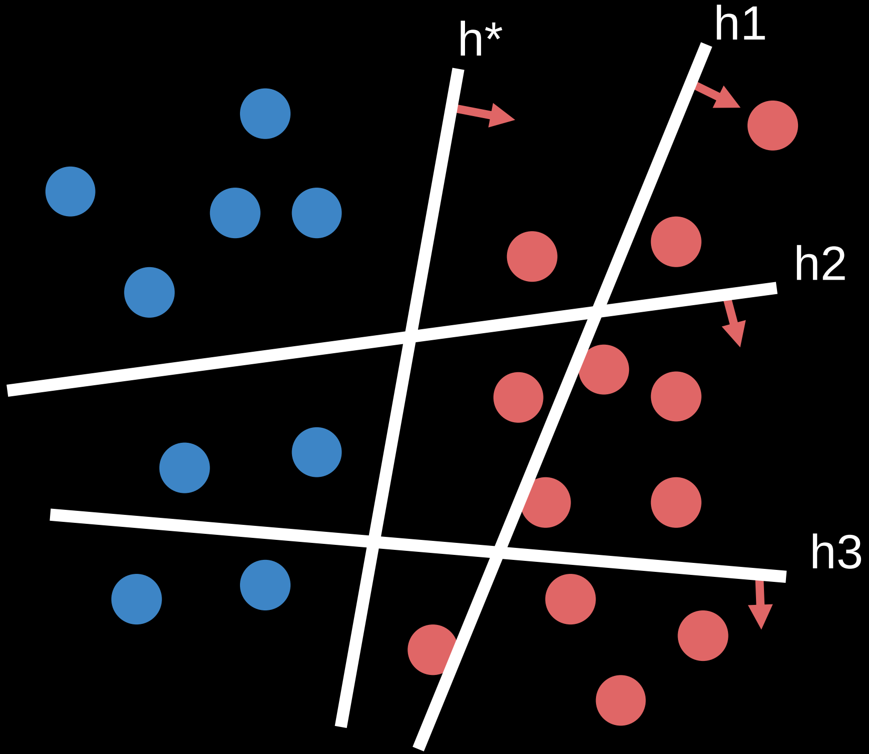
Select for largest reduction in expected error

$$\mathbb{E}[err(h)|T] = \sum_{h \in \mathcal{H}} P(h|T)err(h)$$


h*
h1
h2
h3

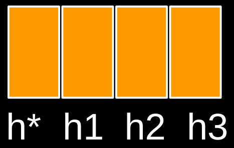
P(h) =

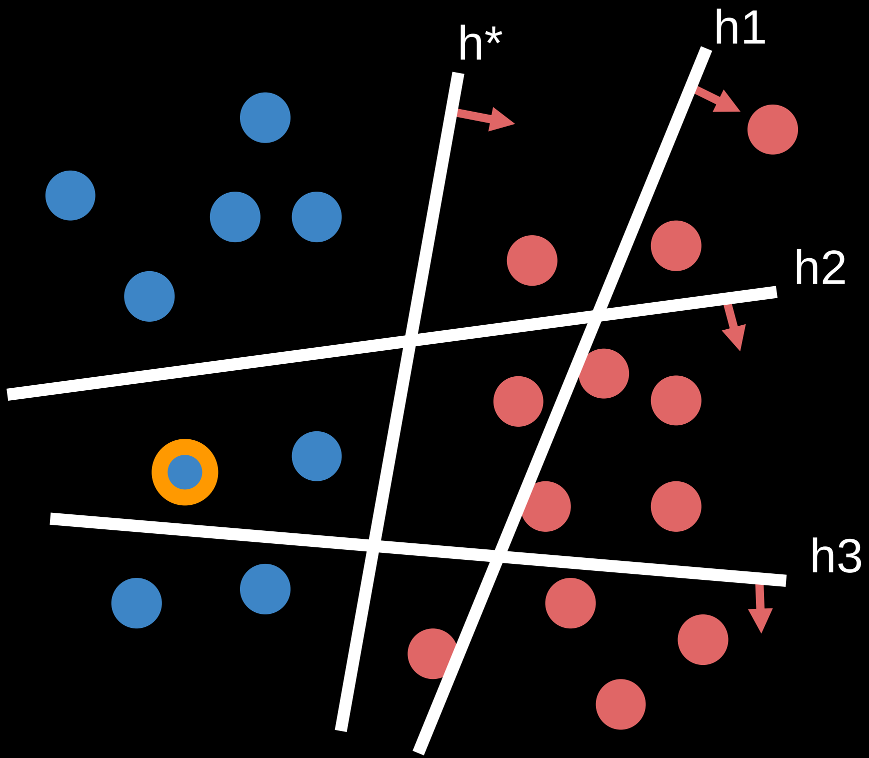




Select Teaching Example 1

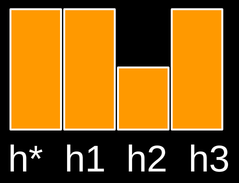
$P(h) =$

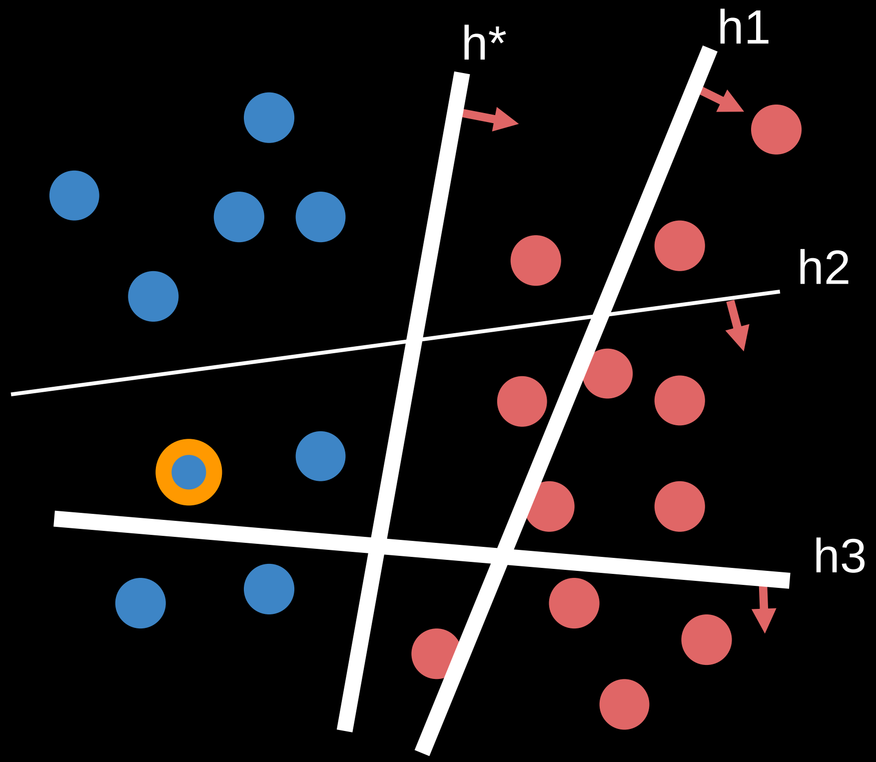




Update Model

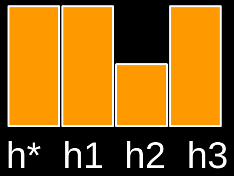
$P(h|x_1) =$

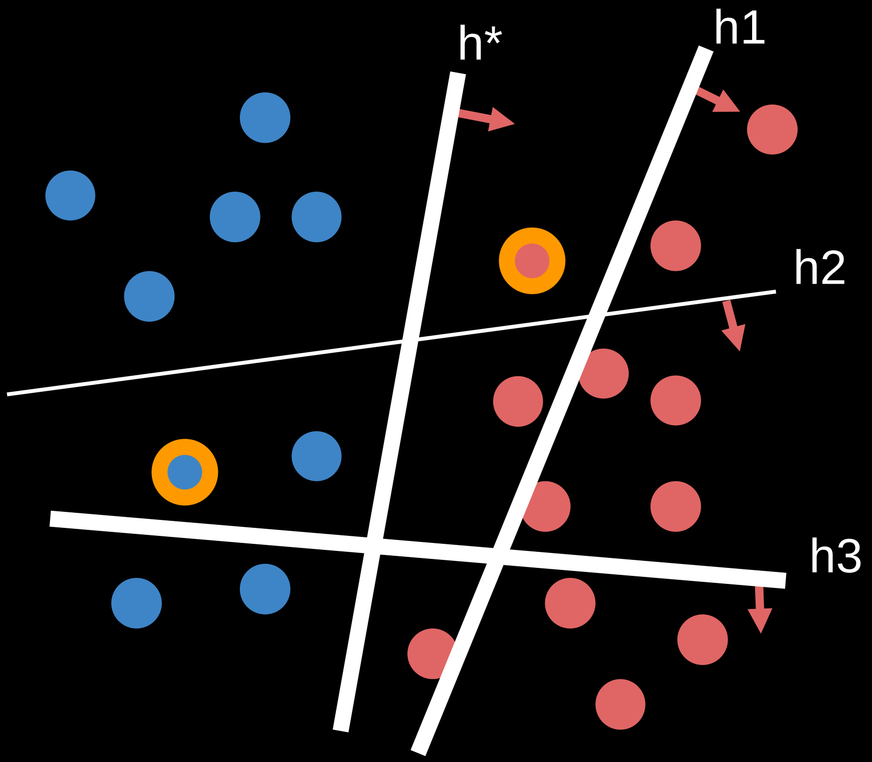




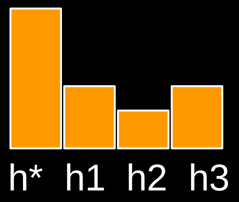
Select Teaching Example 2

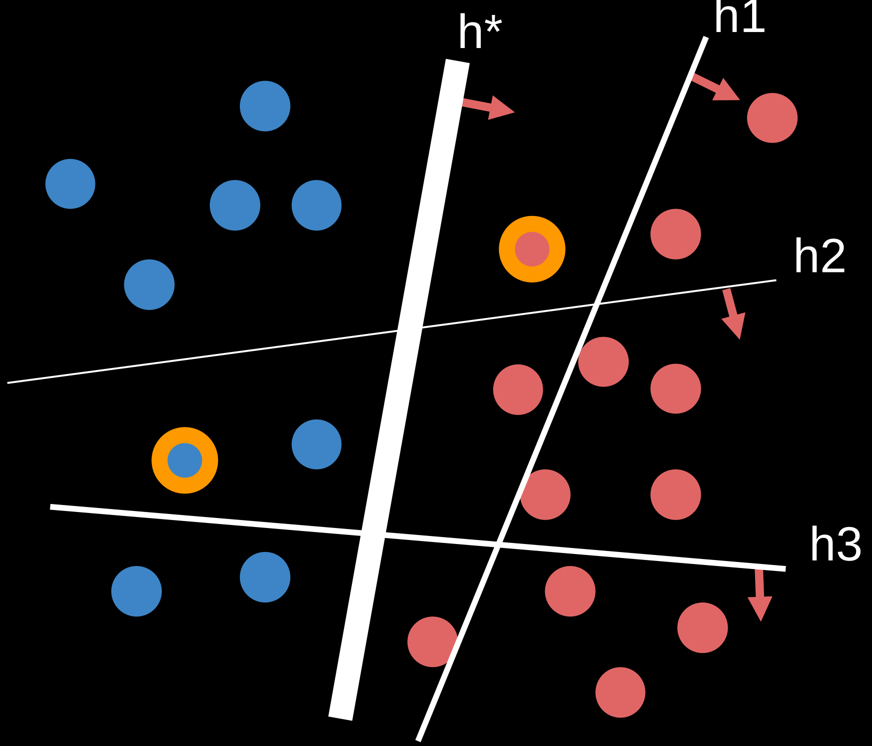
$P(h|x_1) =$





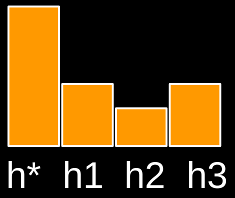
Update Model

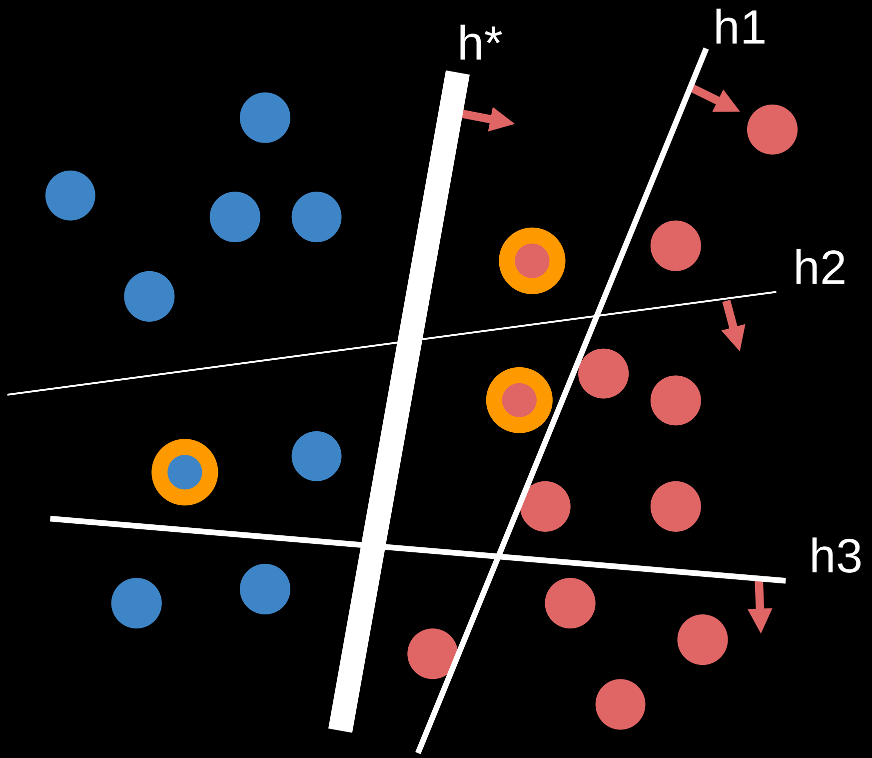
$P(h|x_1, x_2) =$ 



Repeat …

$P(h|x_1, x_2) =$





Multiclass Teaching

Independent posterior per class

$$\frac{1}{C}\sum_{c}\sum_{h\in\mathcal{H}} P_c(h|T)err_c(h)$$

Experimental Setup

Familiarize participants with interface

Teach for 20 iterations

Test for 20 iterations (to measure performance)

Step 1 - Query Learner

Which Species is Present?

A) Viceroy

B) Monarch

C) Queen

D) Red Admiral

Step 2 - Get Learner Response

Which Species is Present?

- **A) Viceroy**
- **B) Monarch**
- **C) Queen**
- **D) Red Admiral**

Step 3 - Provide Feedback

Which Species is Present?

A) Viceroy

B) Monarch

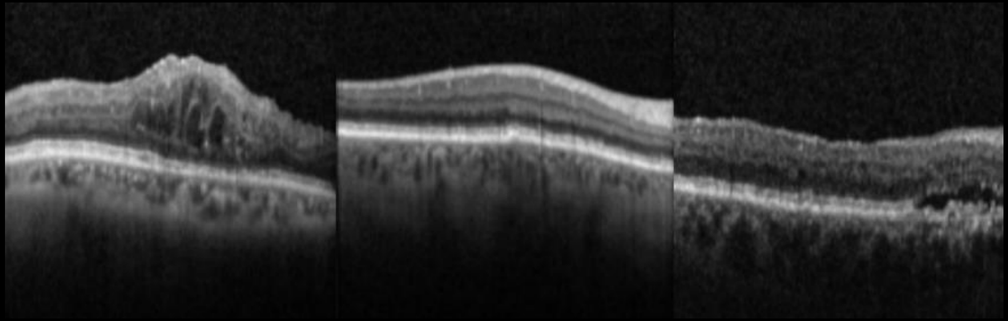
C) Queen

D) Red Admiral

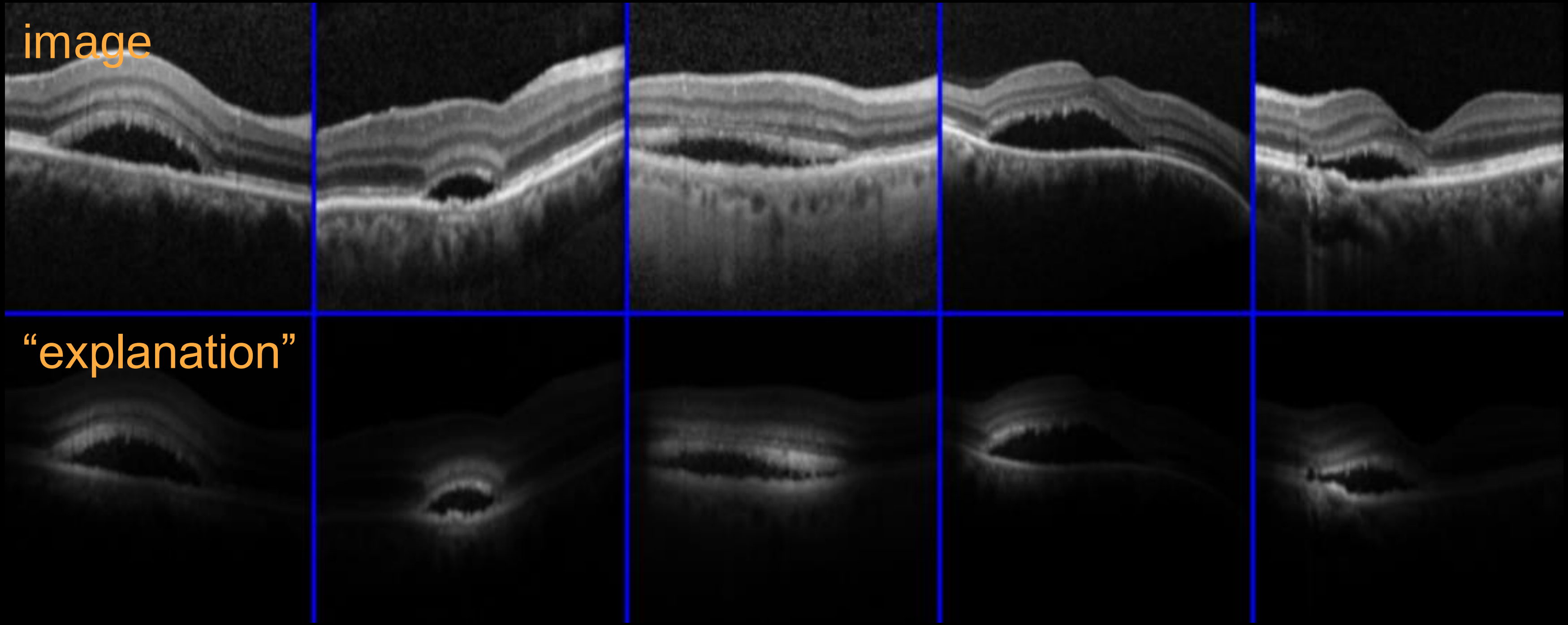
Retina Images

1125 images, 3 classes

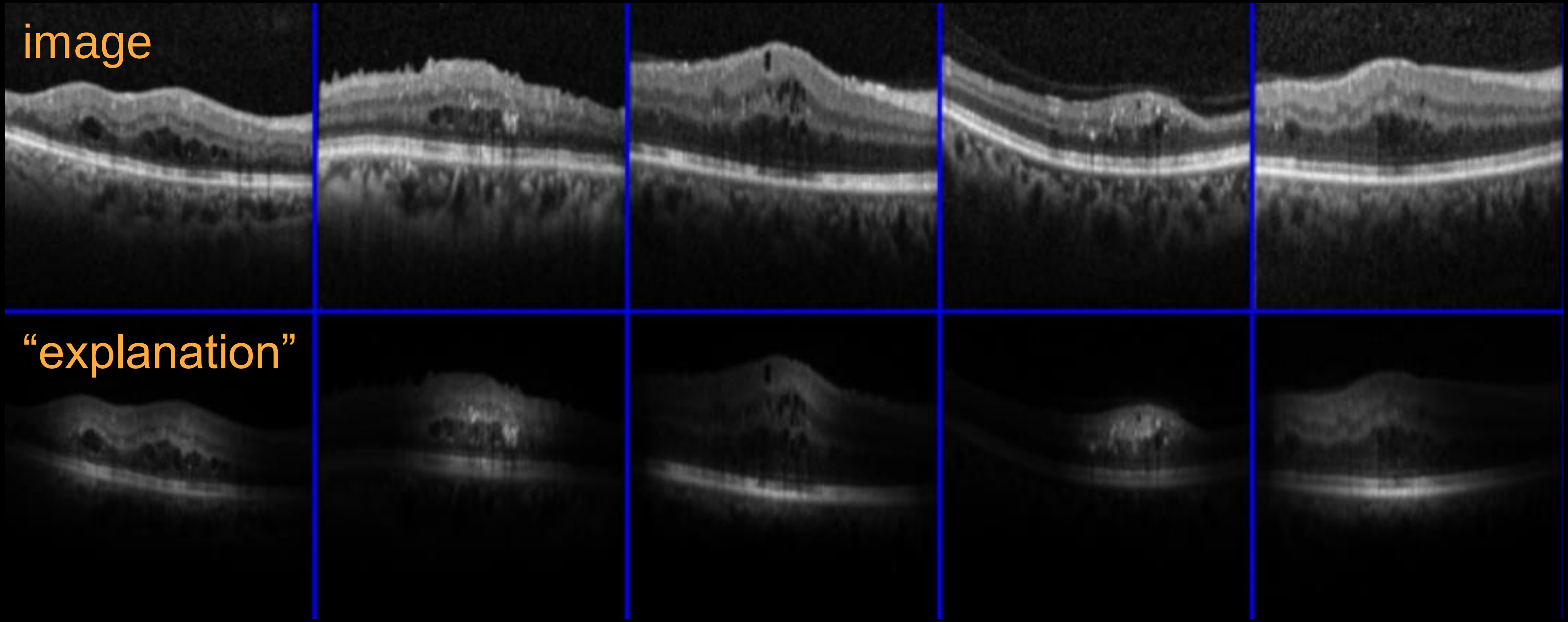
Macular Edema | Normal | Subretinal Fluid



~ 40 participants per dataset per teaching algorithm

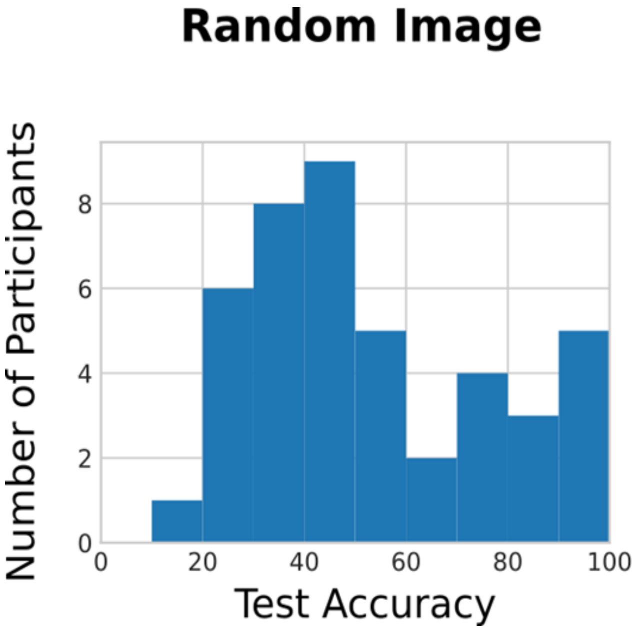


Subretinal fluid

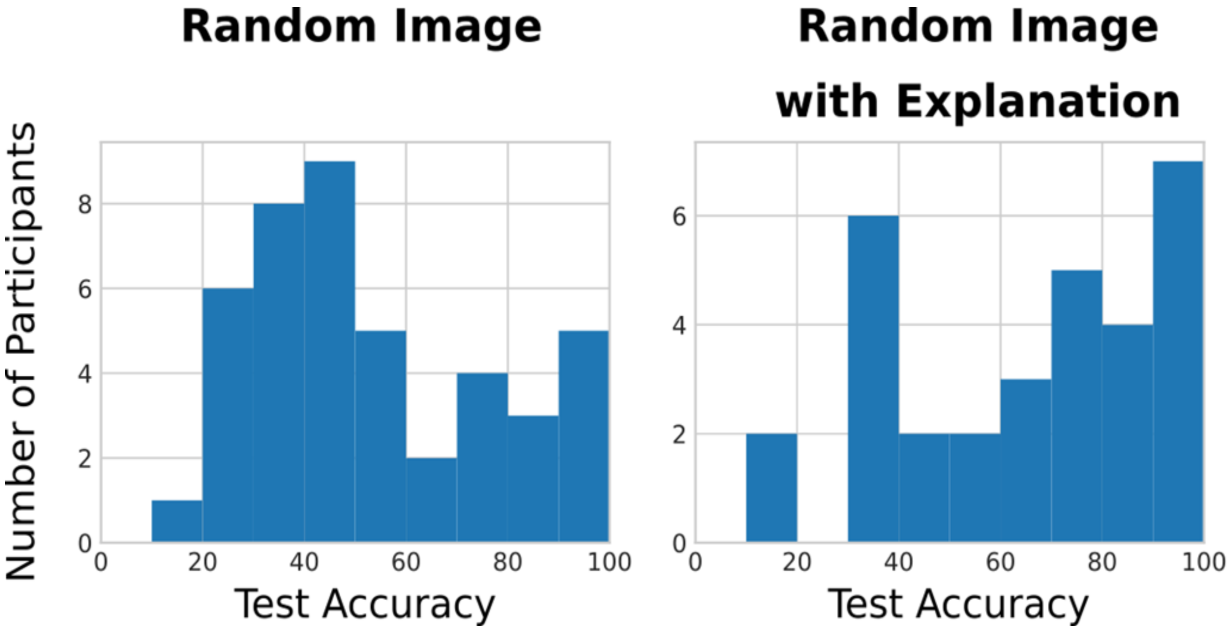


Macular Edema

Results for Retina Images



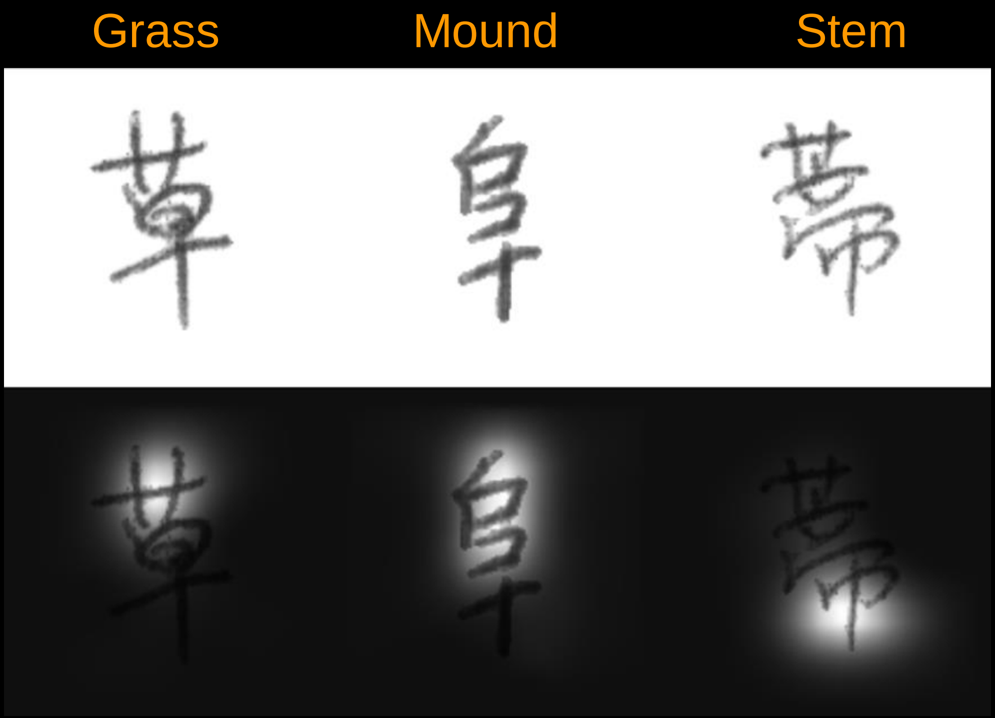
Results for Retina Images



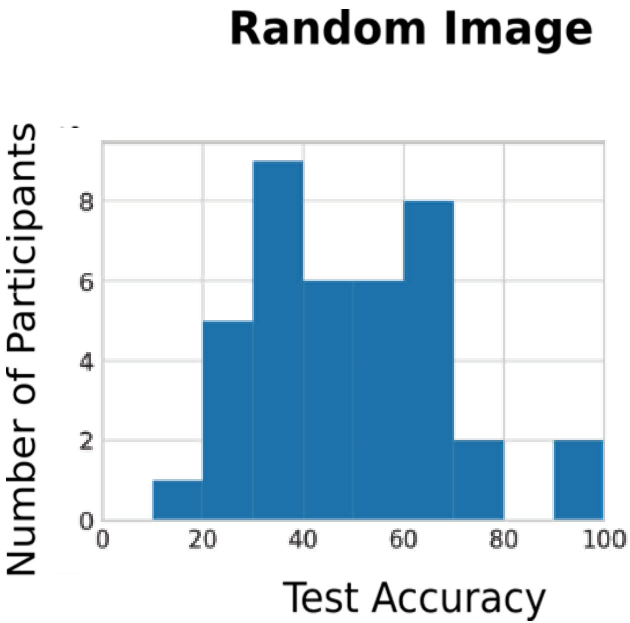
Results for Retina Images

Chinese Characters

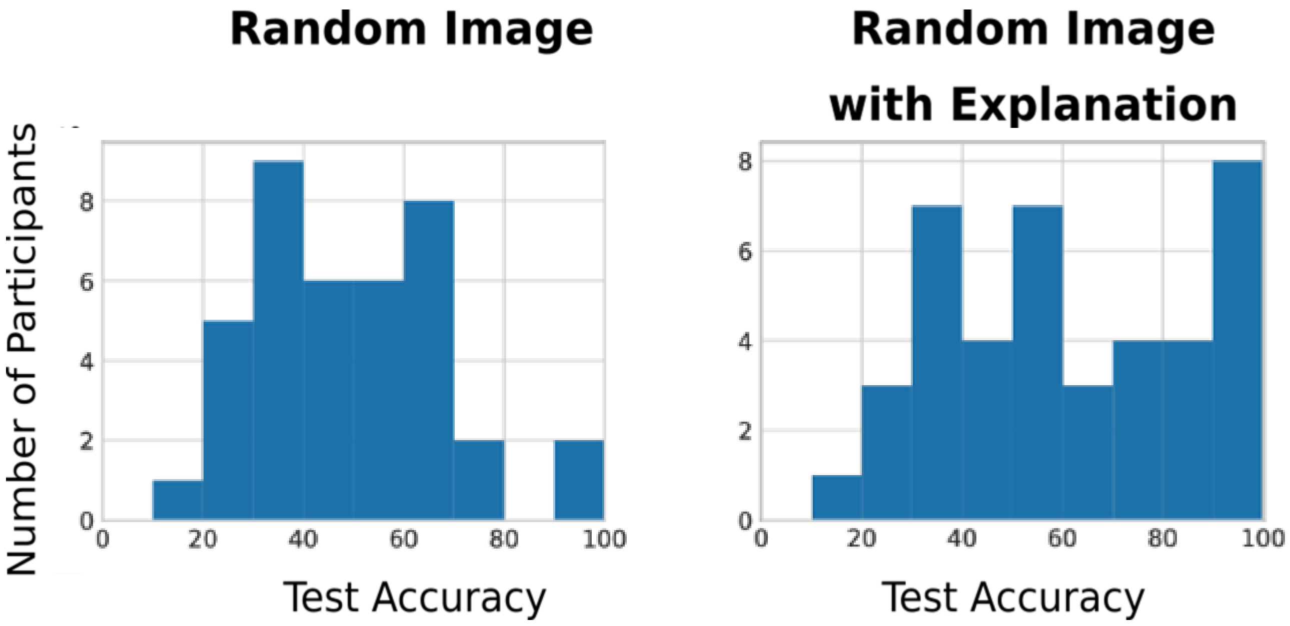
717 images, 3 classes



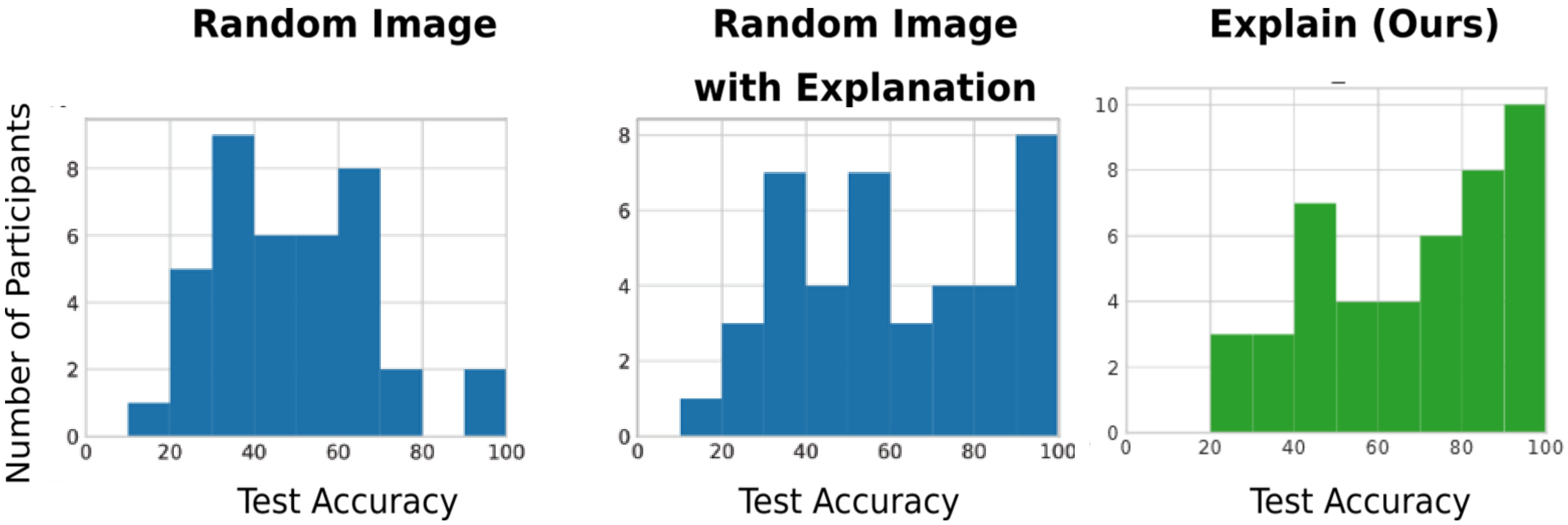
Results for Chinese Characters

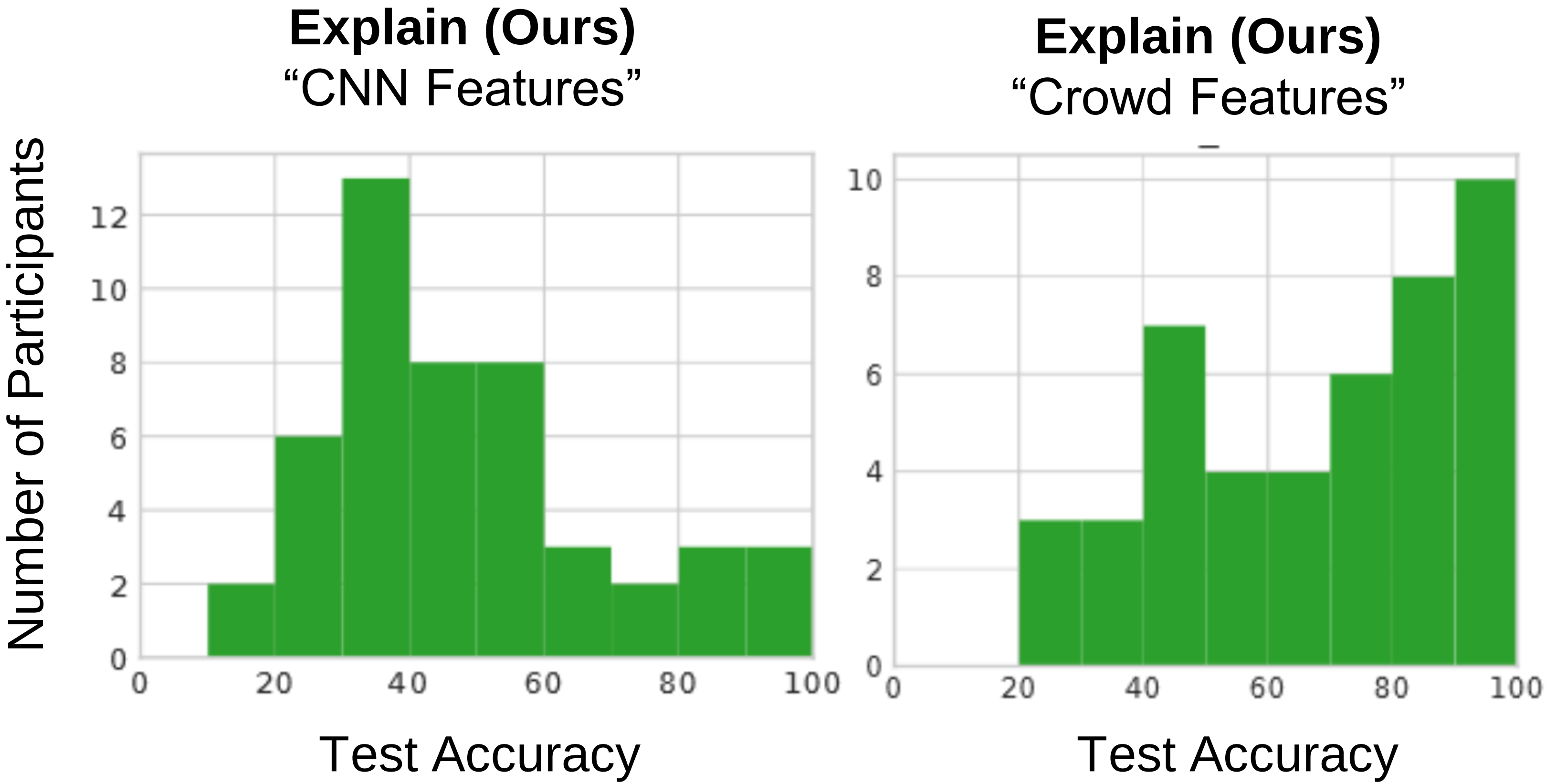


Results for Chinese Characters



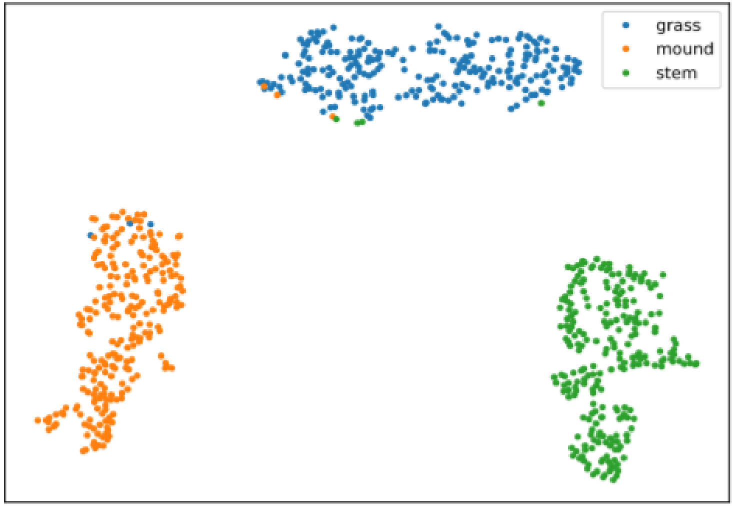
Results for Chinese Characters



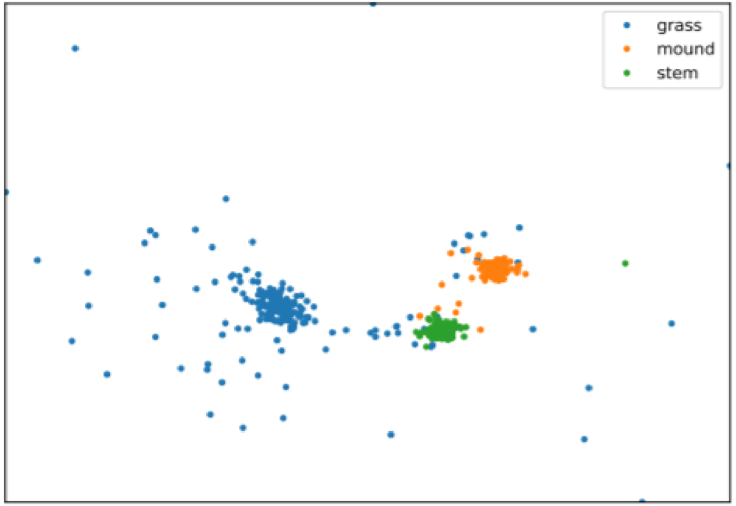


Explain (Ours)
“CNN Features”
Explain (Ours)
“Crowd Features”
Number of Participants
0
2
4
6
8
10
12
0
20
40
60
80
100
Test Accuracy
0
2
4
6
8
10
0
20
40
60
80
100
Test Accuracy

“CNN Features”



“Crowd Features”



Grass

Mound

Stem

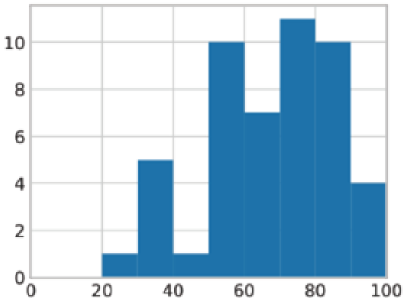
Butterflies

2,224 images, 5 classes

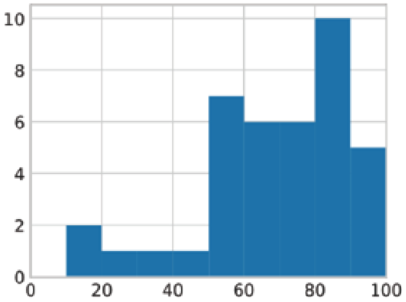
Monarch | Viceroy | Queen | Red Admiral | Cabbage White

Results for Butterflies

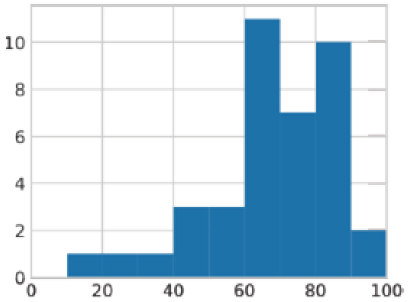
Random Label Only



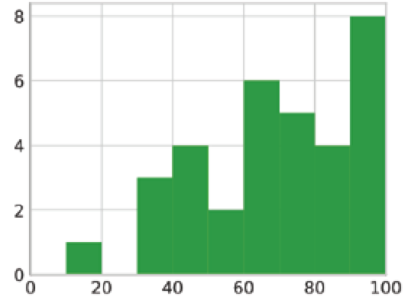
Random w/ Explain



Greedy Label Only

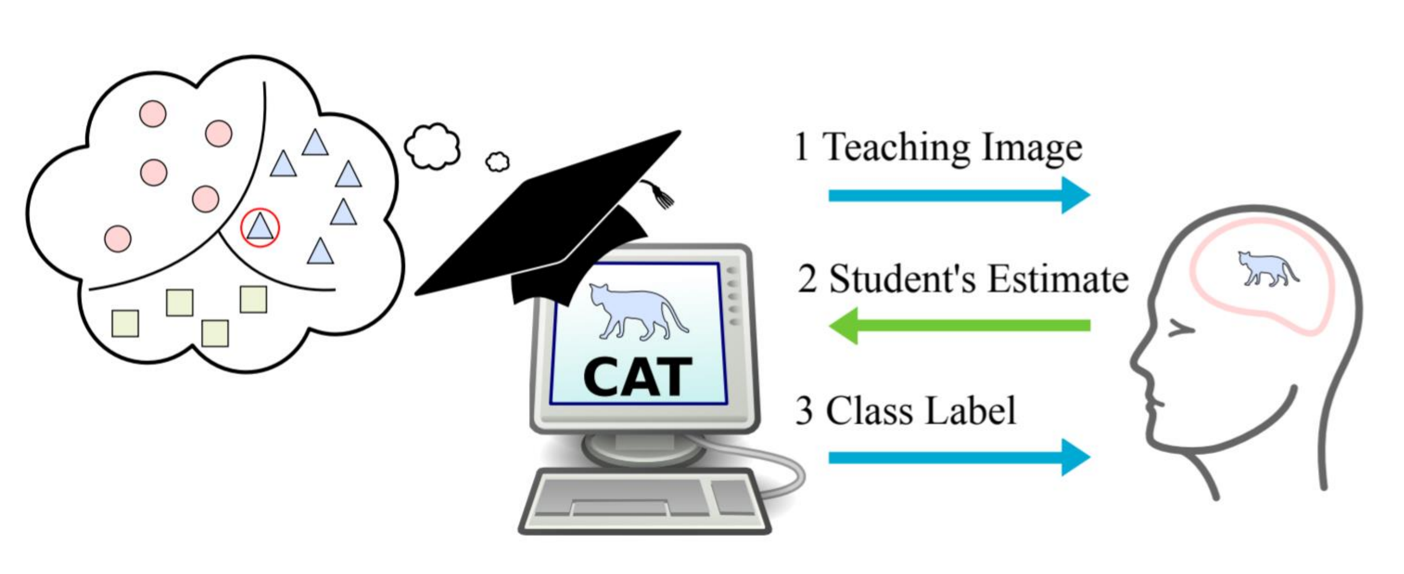


Our Approach



Next steps for teaching visual knowledge ….

Interactive Teaching



Becoming the Expert: Interactive Multi-Class Machine Teaching **CVPR 2015**
Johns, Mac Aodha, Brostow

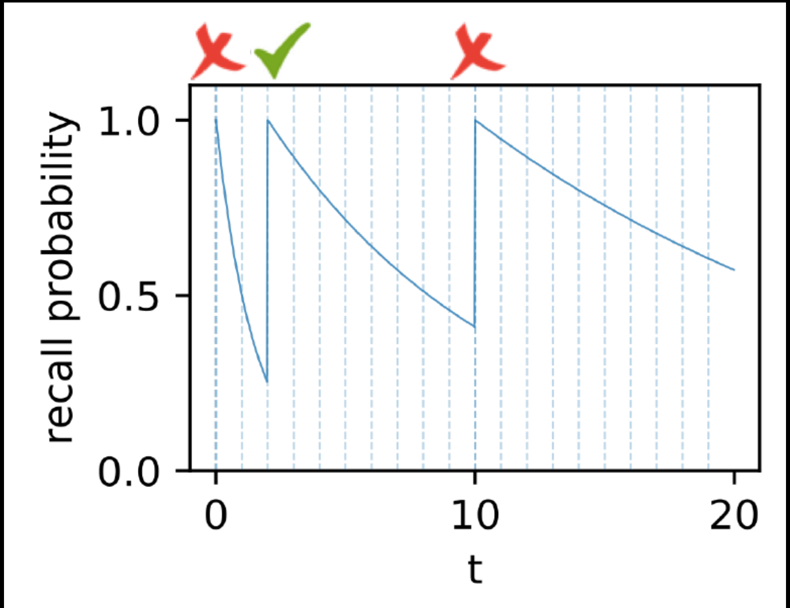
Understanding the Role of Adaptivity in Machine Teaching: The Case of Version Space Learners **NeurIPS 2018**
Chen, Singla, Mac Aodha, Perona, Yue

Modelling Learner Memory Decay

Memory decays over time

Spaced repetition model

Estimate learner recall



Teaching Multiple Concepts to Forgetful Learners **NeurIPS 2019**

Hunziker, Chen, Mac Aodha, Gomez Rodriguez, Krause, Perona, Yue, Singla

Scaling Up Visual Teaching - ebird.org/quiz

eBird Submit Explore My eBird Science About News Help

Donate Oisin Mac Aodha (macaodha) Language

211 S Wilson Ave, Pasadena, CA 91106, USA

What bird is this? 1 of 20 Home

- **Pacific Wren**
- **Marsh Wren**
- **Bewick's Wren**
- **Cactus Wren**
- **None of the above**

Not identifiable due to:

- Multiple species
- No bird

© Lawrence Haller

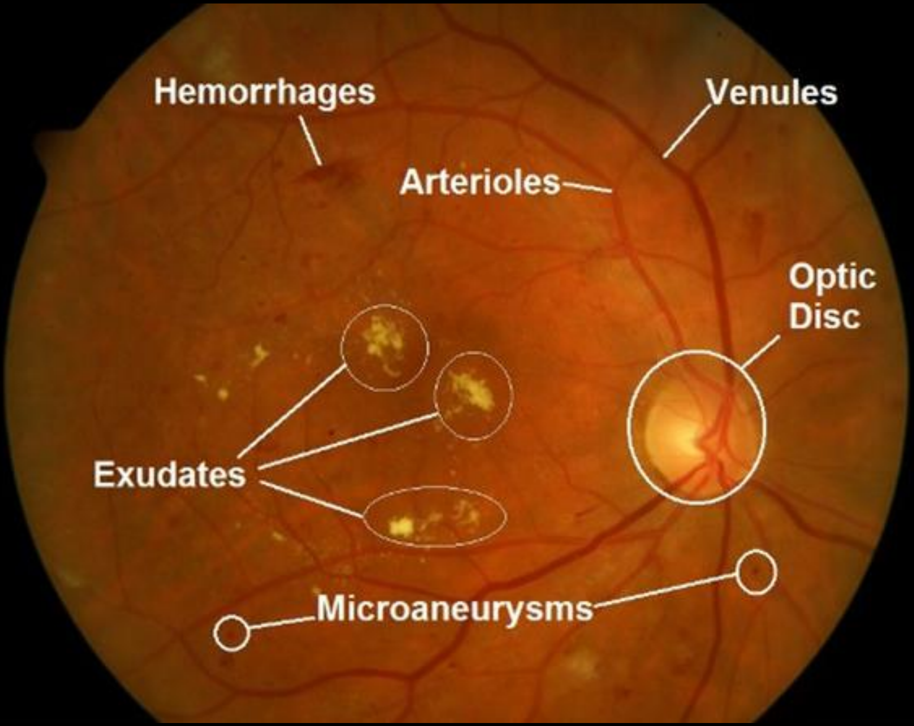
Williamson, Texas, United States | 7 Aug 2016

Macaulay Library ML32156901 | eBird Checklist S31011305 | TheCornellLab

Report...

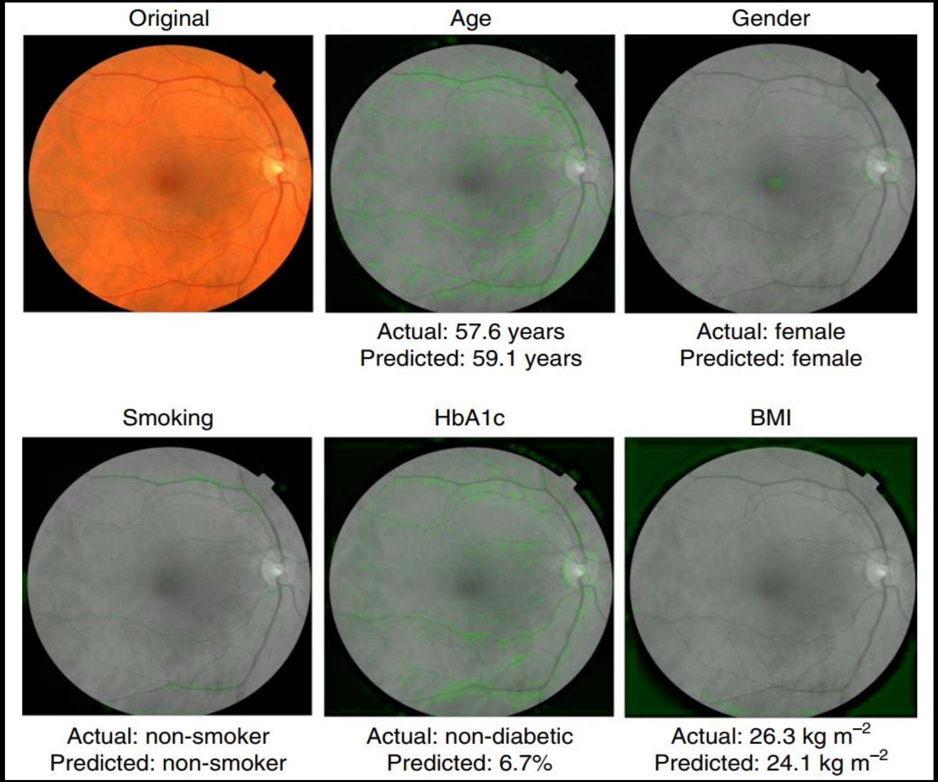
Teaching Fine-Grained Detail

Learning explanations through teaching



Closing the Loop

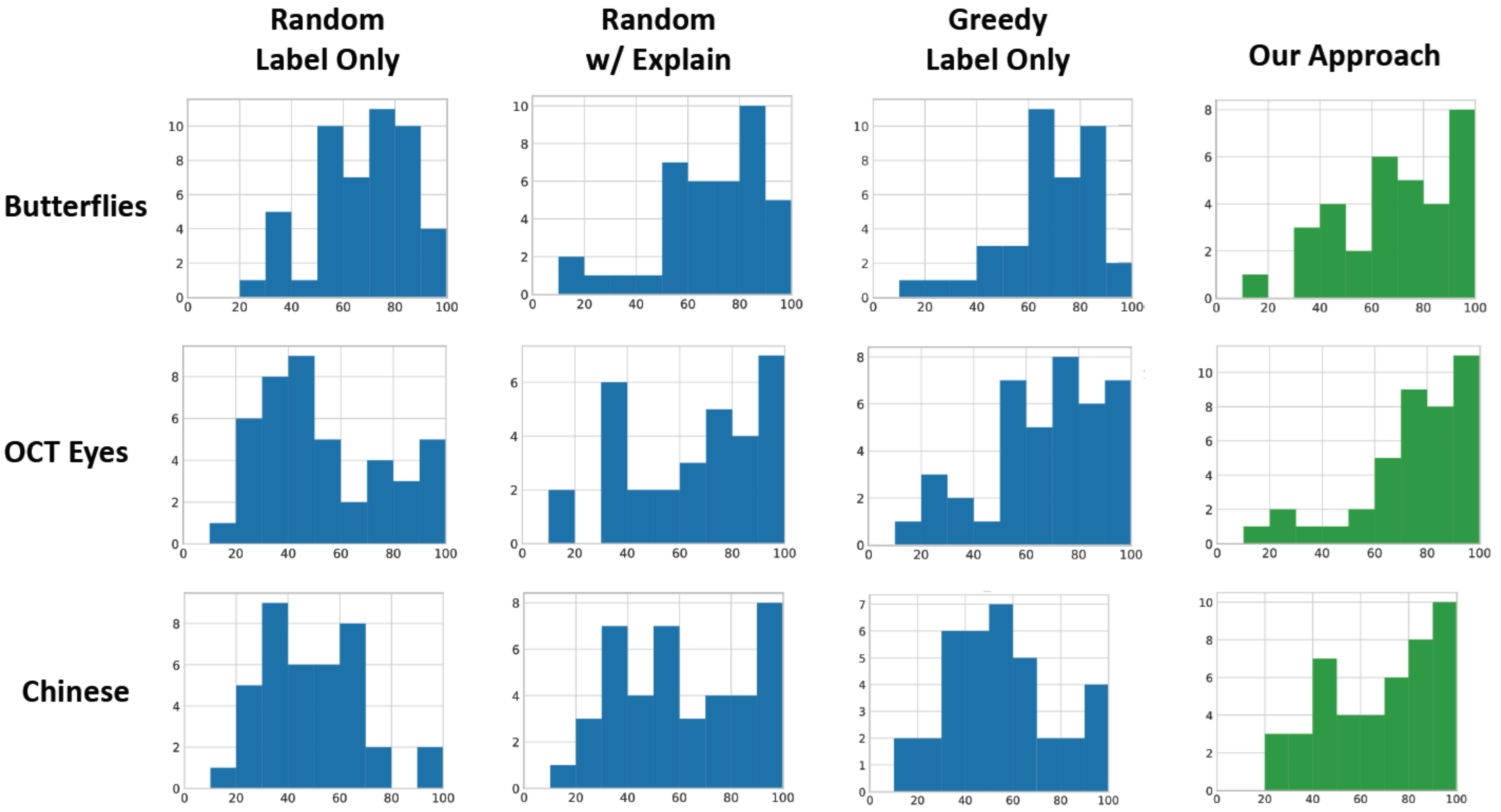
Teaching super human image understanding

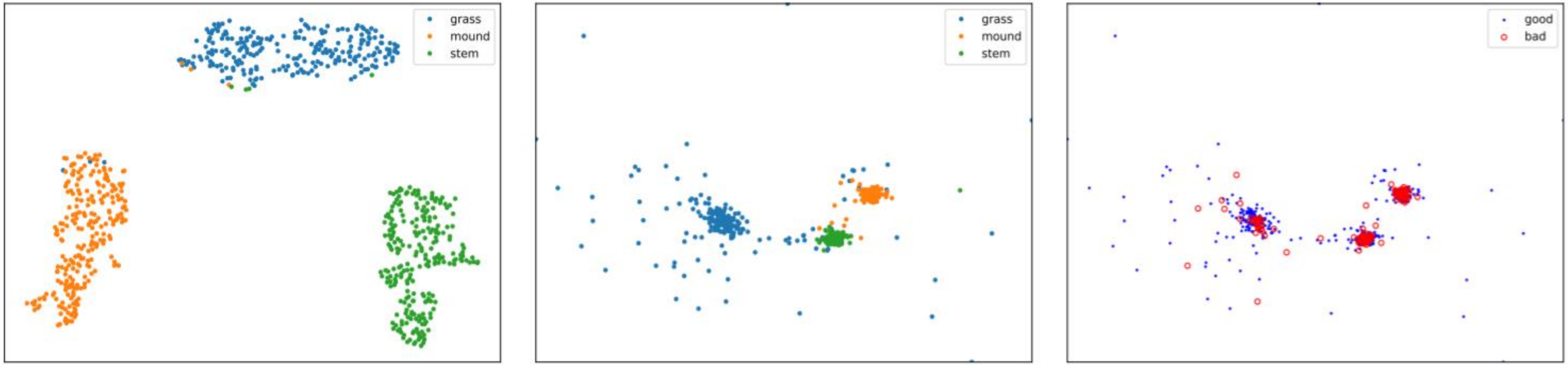


Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning
Poplin et al. **Nature Biomedical Engineering 2018**

Questions

Teaching GUI, model code, and data:
https://github.com/macaodha/explain_teach


Random
Label Only
Random
w/ Explain
Greedy
Label Only
Our Approach
Butterflies
OCT Eyes
Chinese

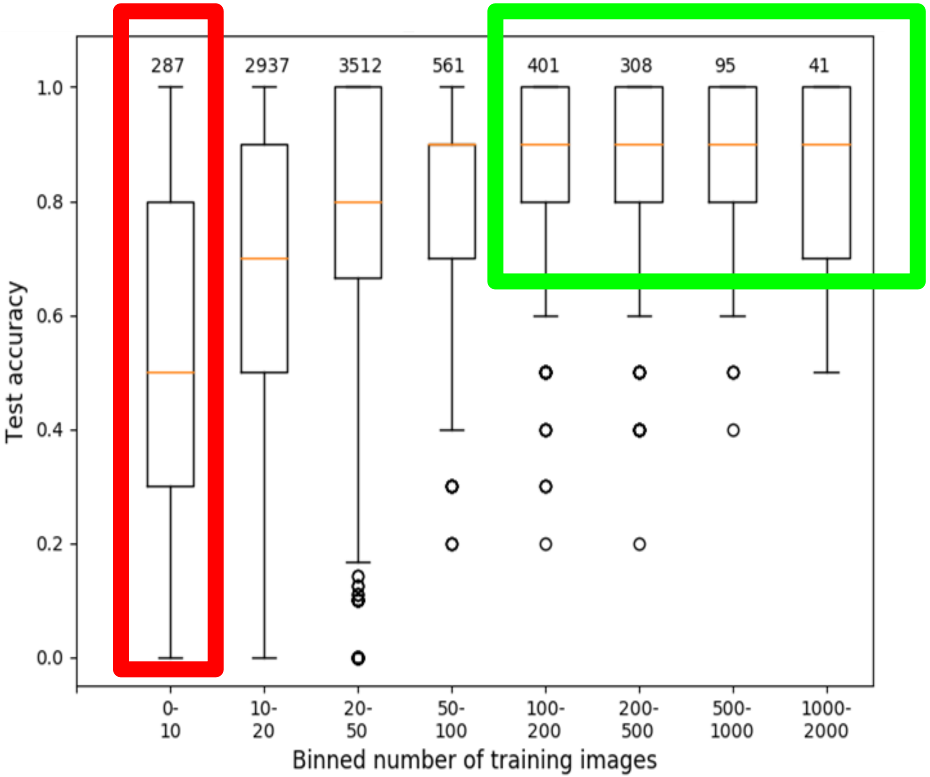


A) CNN Embedding B) Crowd Embedding C) Explanation Difficulty

Learning How to Perform Low Shot Learning

iNaturalist Dataset

8,142 classes
>400K images



The iNaturalist Species Classification and Detection Dataset **CVPR 2018**
Van Horn, Mac Aodha, Song, Cui, Sun, Shepard, Adam, Perona, Belongie