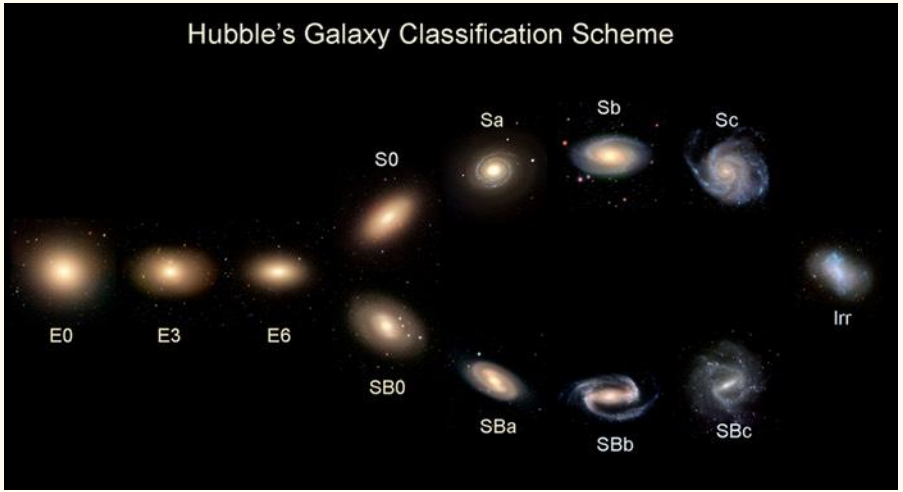
Cluster Galaxy Classification Using the Perseus Cluster With a Convolutional Neural Network

—

Jason Pruitt, Dr. Aaron Romanowsky
Dept. of Physics and Astronomy San José State University

Classifying Galaxies by Hand/Eye

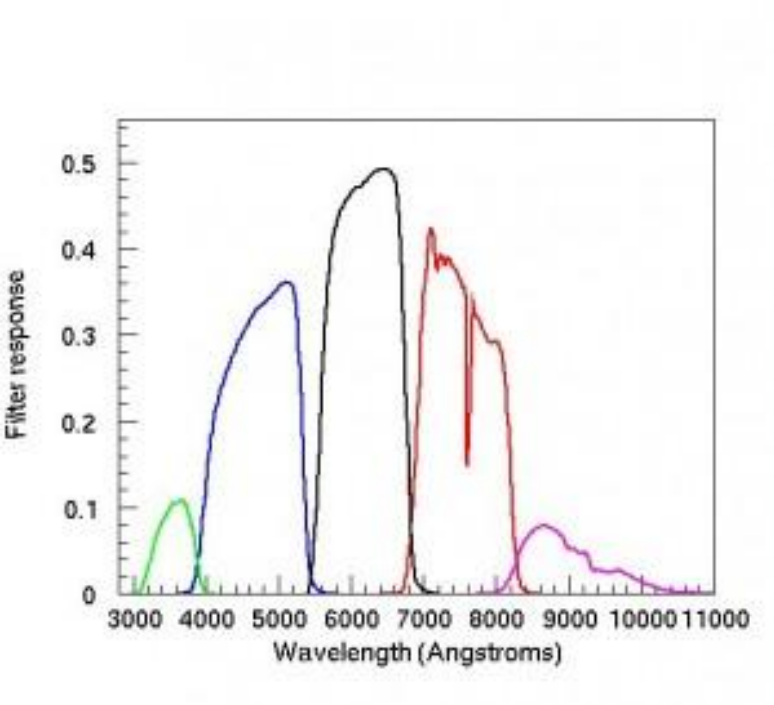
Astronomers use the Hubble tuning fork:



"Morphological Classification" - based off of shape, constituent components

https://itu.physics.uiowa.edu/labs/advanced/classifying-galaxies/part-1-hubbles-tuning-fork

Stellar Objects and Color



https://www.sdss4.org/instruments/camera/#Filters

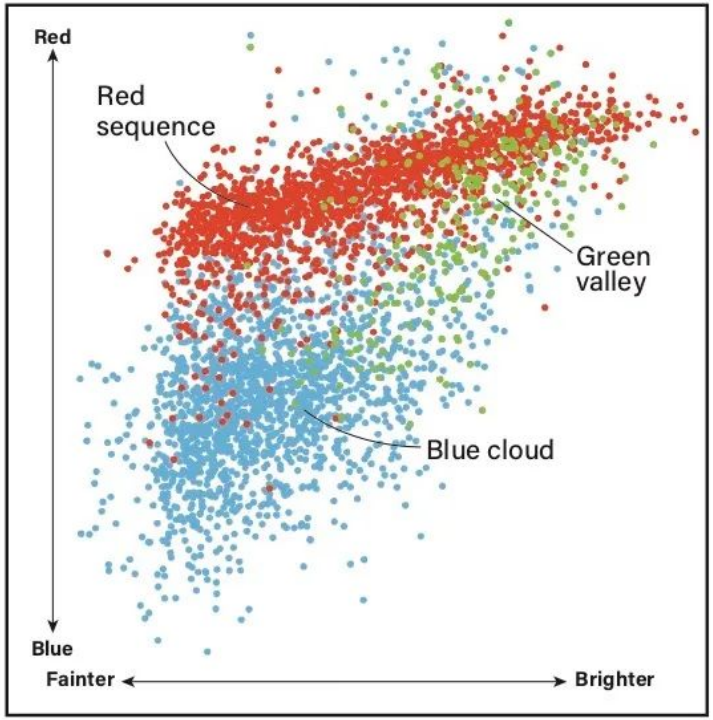
Magnitudes:

r, i, u, z, g

Colors:

(r-i), (r-u), (u-z), etc.

Color regions from color-magnitude relationships



(Bakich, 2021), modified from Gavazzi et al. (2010).

Red Sequence

Blue Cloud

Green Valley

Perseus Cluster Catalogue

The Perseus galaxy cluster

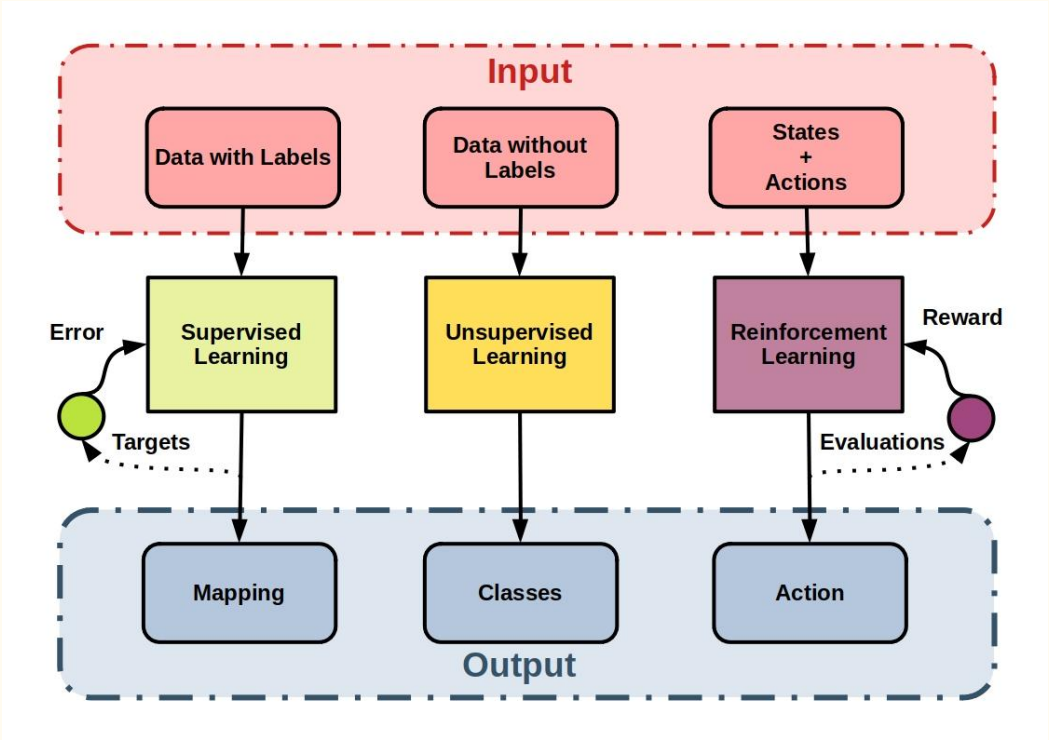
An example elliptical galaxy in the cluster

Tagged by Wittmann et al. 2019

https://chandra.harvard.edu/photo/2014/perseus/

Wittmann et al, A catalog of galaxies in the direction of the perseus cluster. The Astrophysical Journal Supplement Series, 245(1):10, nov 2019.

Types of Machine Learning



https://starship-knowledge.com/supervised-vs-unsupervised-vs-reinforcement

Machine Learning and Neural Networks

200x200x3
Pixel
Image
Tensor

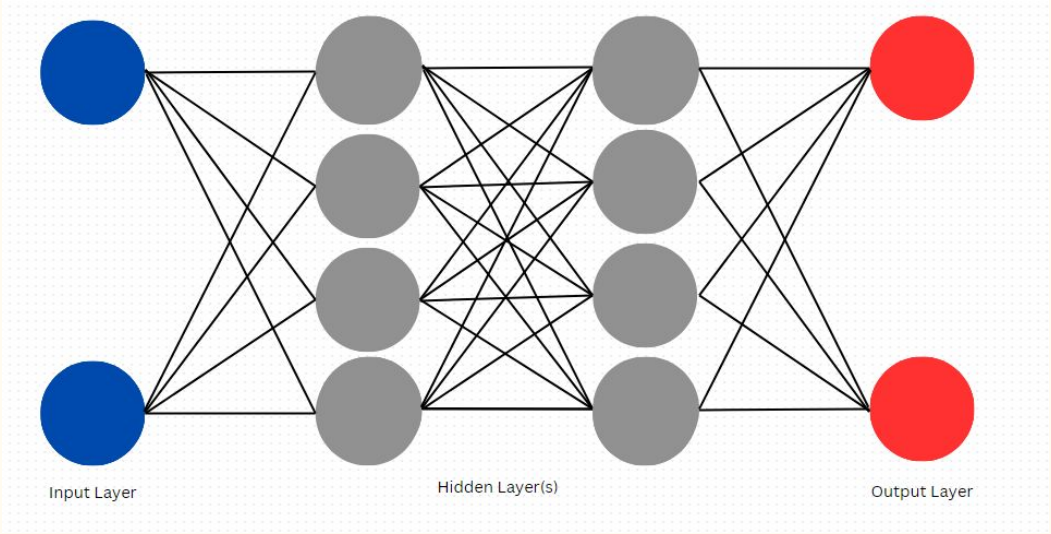
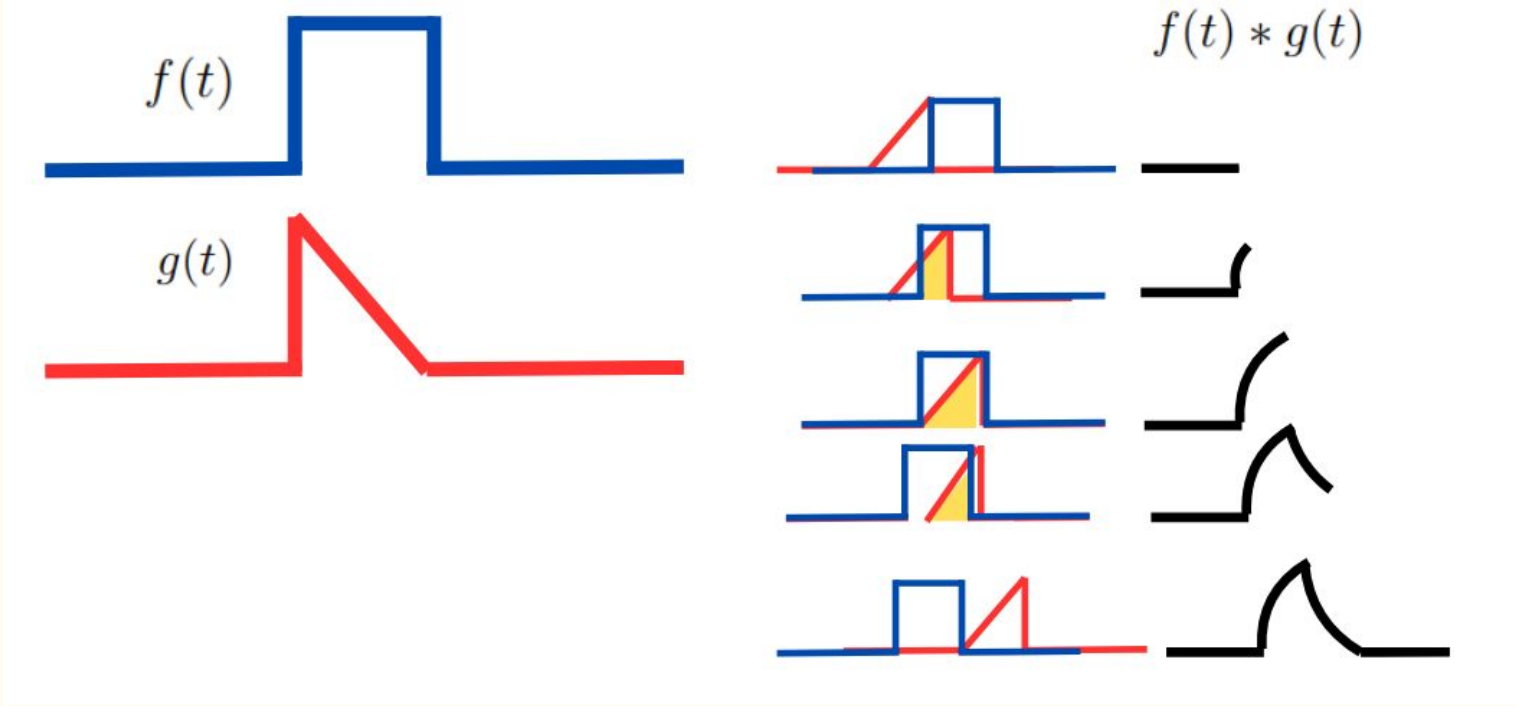


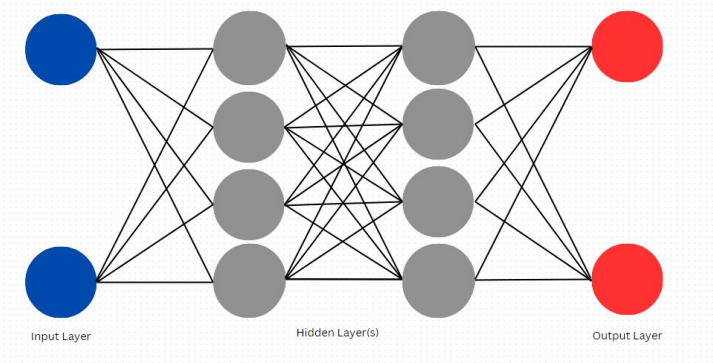
Image
Prediction

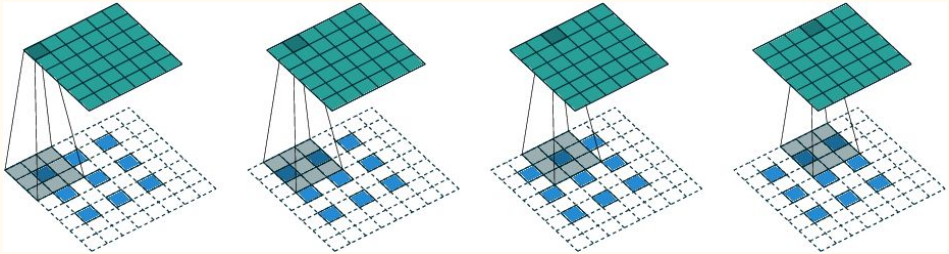
"Cluster member"

Function Convolution in Math



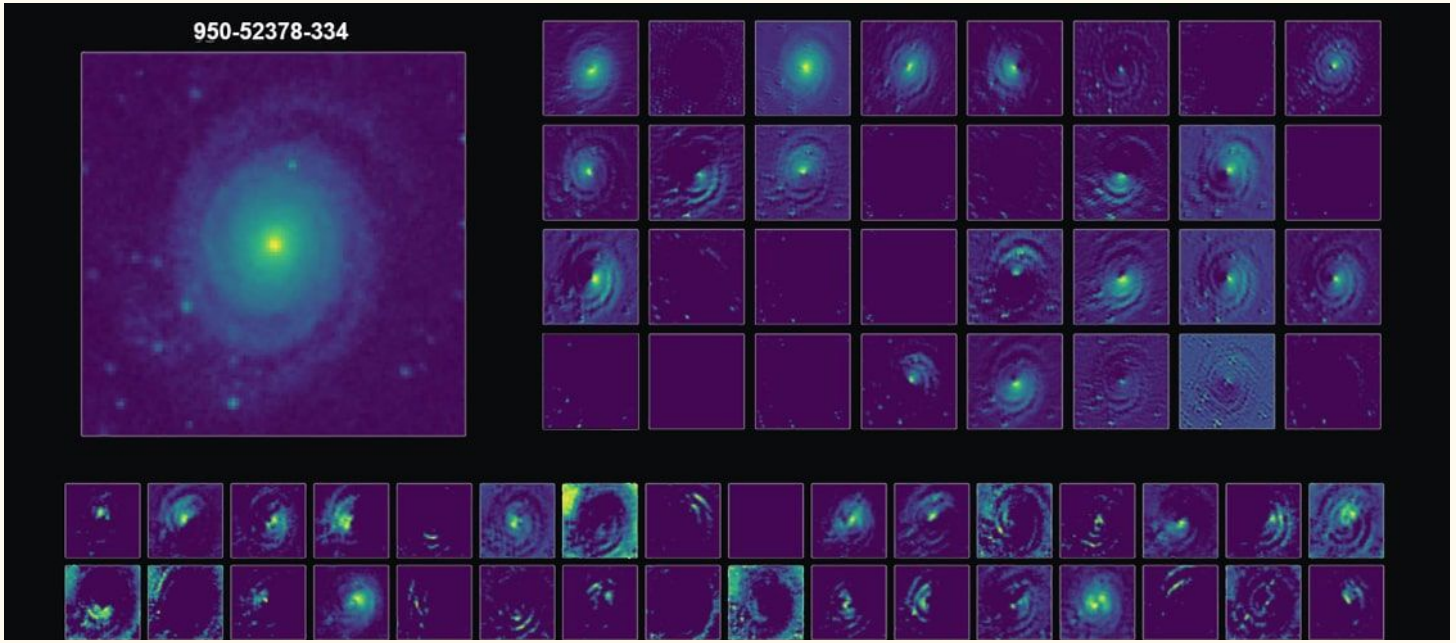
The Convolutional Layer





Dumoulin, Vincent and Francesco Visin. “A guide to convolution arithmetic for deep learning.” ArXiv abs/1603.07285 (2016): n. pag.

Convolution Builds a Feature Map

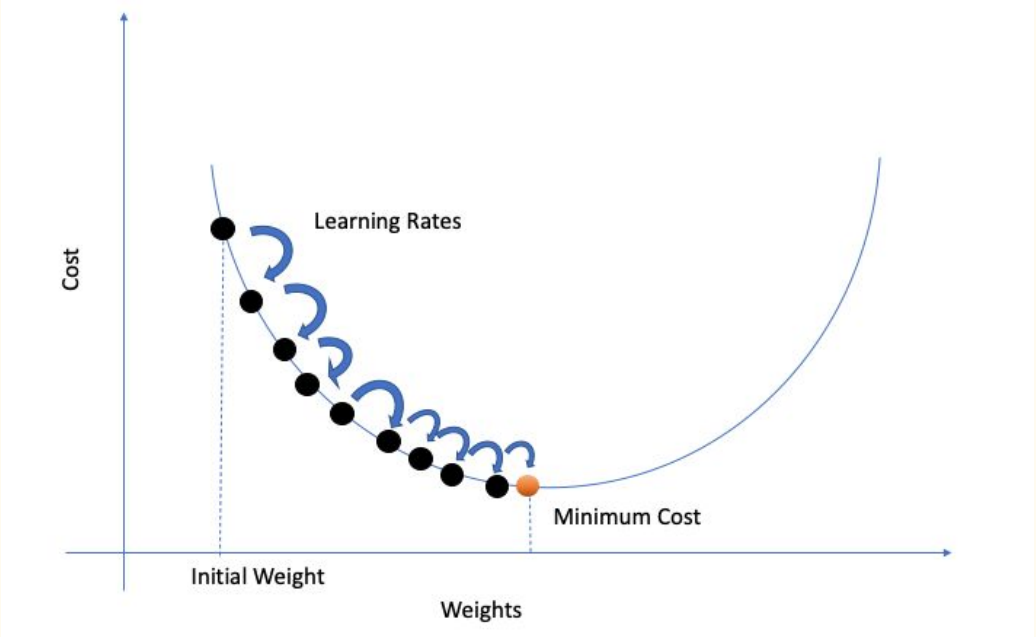


https://phys.org/news/2021-07-thousands-galaxies-eye.html

Large Networks Cause Problems

Balance between enough layers / parameters and the vanishing / exploding gradients

Can’t just add more layers blindly



Ghosh, Bhaskar et al.. (2020). An Empirical Analysis of Generative Adversarial Network Training Times with Varying Batch Sizes. 10.1109/UEMCON51285.2020.9298092.

He et al. - Residual Learning

Target Classification: $\mathcal{H}(x)$ ex. “Cluster galaxy”

He et al. - Residual Learning

Target Classification: $\mathcal{H}(x)$ ex. “Cluster galaxy”

Residual: $\mathcal{F}(x) = \mathcal{H}(x) - x$

He et al. - Residual Learning

Target Classification: $\mathcal{H}(x)$ ex. “Cluster galaxy”

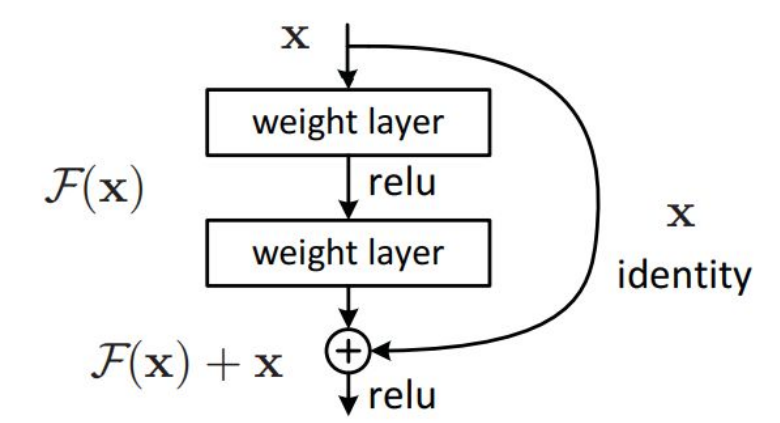
Residual: $\mathcal{F}(x) = \mathcal{H}(x) - x$

Remapping: $\mathcal{F}(x) + x = \mathcal{H}(x)$

Now the network only needs to learn $\mathcal{F}(x)$

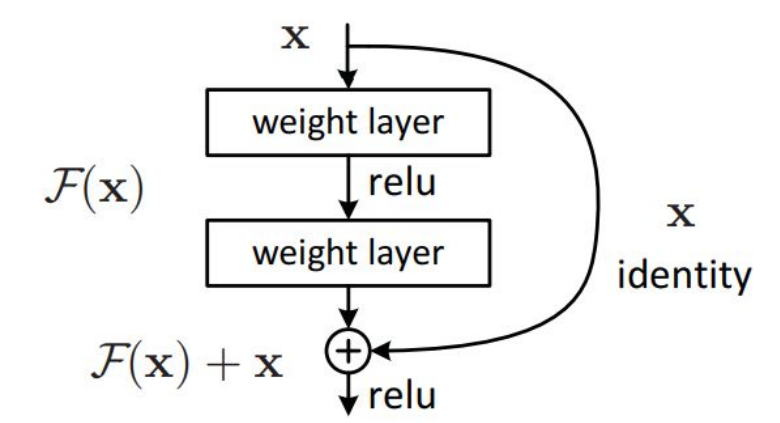
He et al. - ResNets

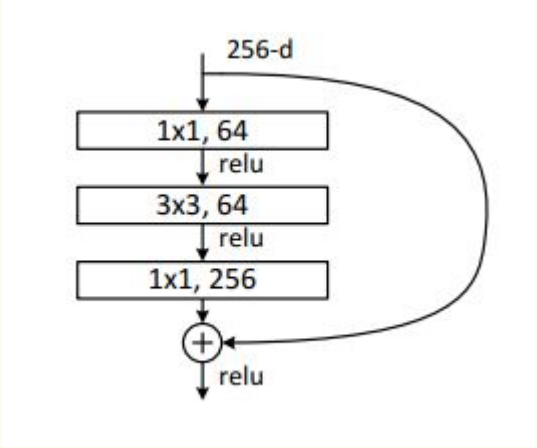
Remapping: $\mathcal{F}(x) + x = \mathcal{H}(x)$



He et al. - ResNets

Remapping: $\mathcal{F}(x) + x = \mathcal{H}(x)$

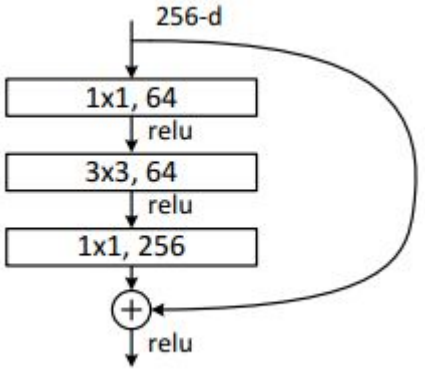




He et al. - ResNets

Remapping: $\mathcal{F}(x) + x = \mathcal{H}(x)$

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9



Important Metrics for Classification

First consider:

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

Important Metrics for Classification

First consider:

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

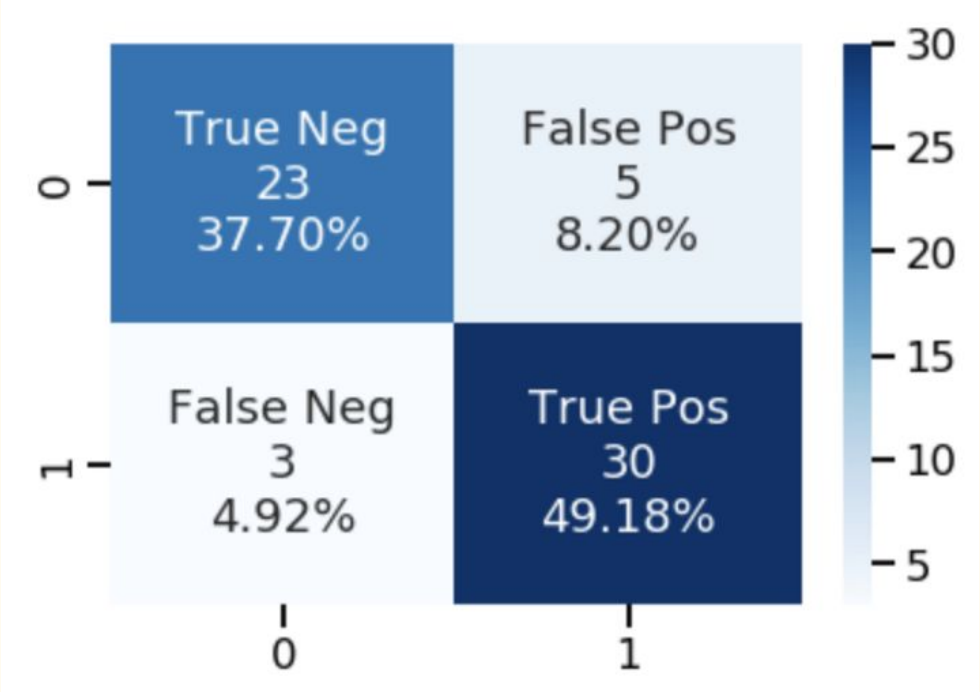
$$Accuracy = \frac{TP + TN}{P + N}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

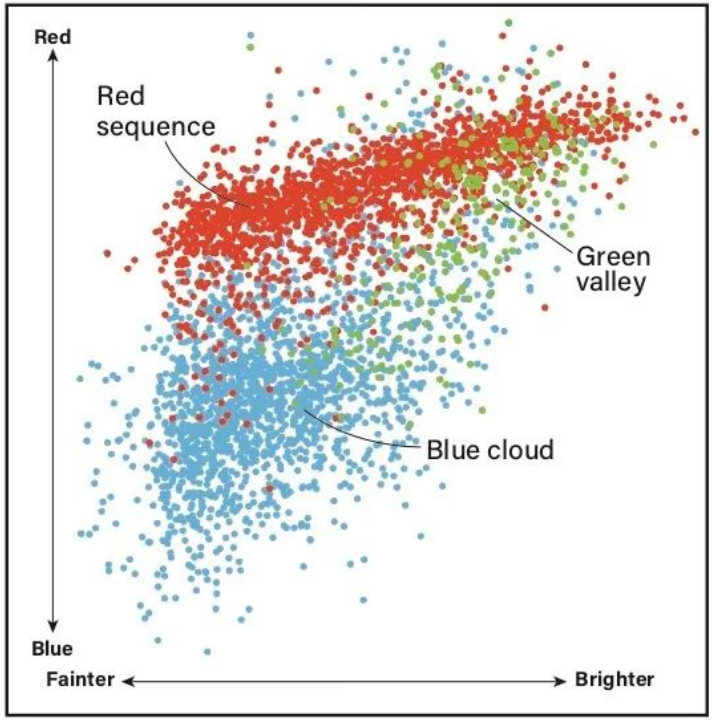
$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

All Together - Confusion Matrix



Binary classification - https://medium.com/@dtuk81/confusion-matrix-visualization-fc31e3f30fea

Problem to Solve - Cluster Galaxy Classification

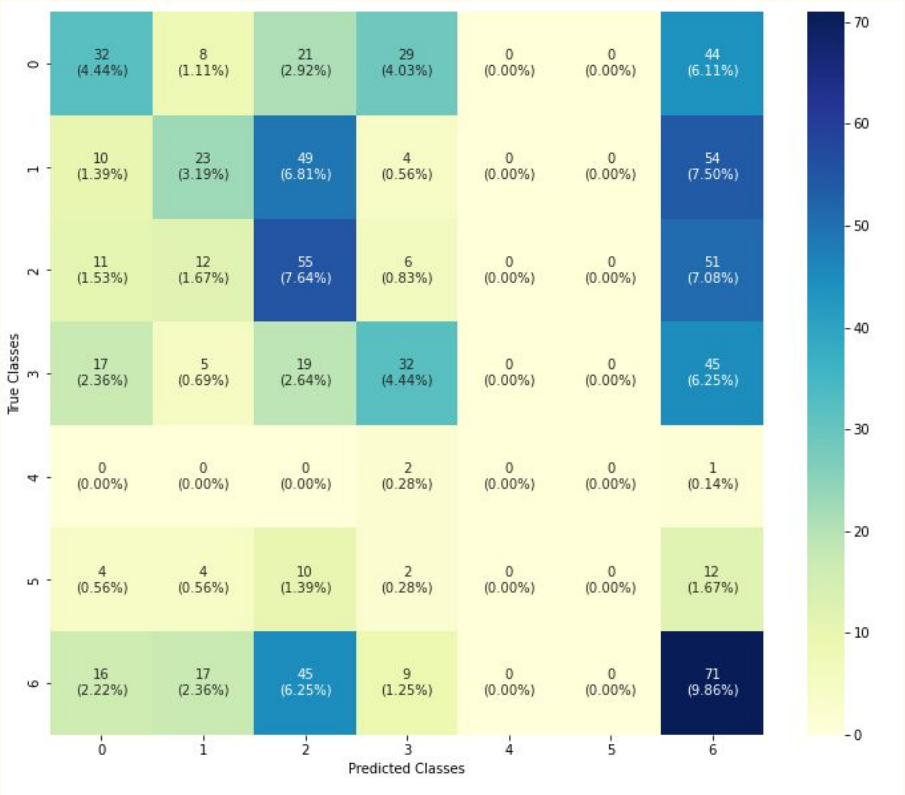


(Bakich, 2021), modified from Gavazzi et al. (2010).

Work from Wittmann et al. 2019 gives us a 5437 image training set with labels

Can we use this to classify cluster galaxies not in the red sequence?

Confusion Matrix for Visualization of Performance



Label mapping for 7 classes:

0 : Cluster or background LTG

1 : Likely background ETG or unresolved source

2 : Likely cluster or background edge-on disk galaxy

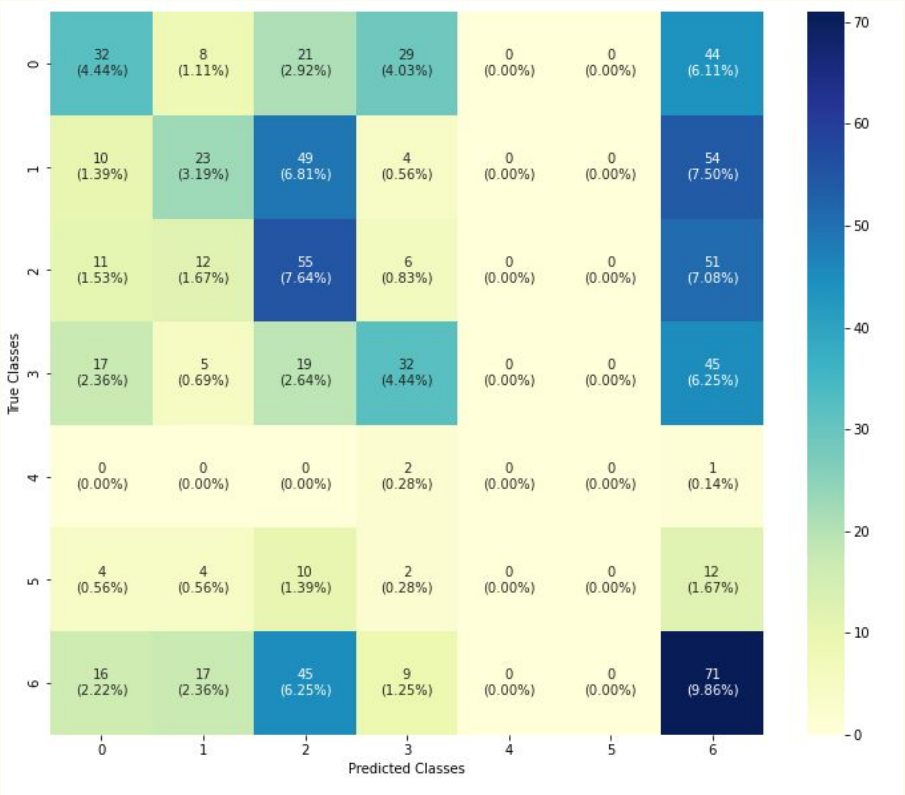
3 : Likely dE/ETGcluster candidate

4 : Likely merging system

5 : Possible dE/ETGcluster candidate

6 : Background galaxy with possibly weak substructure

Confusion Matrix for Visualization of Performance



Label mapping for 7 classes:

0 : Cluster or background LTG

0 : Likely background ETG or unresolved source

0 : Likely cluster or background edge-on disk galaxy

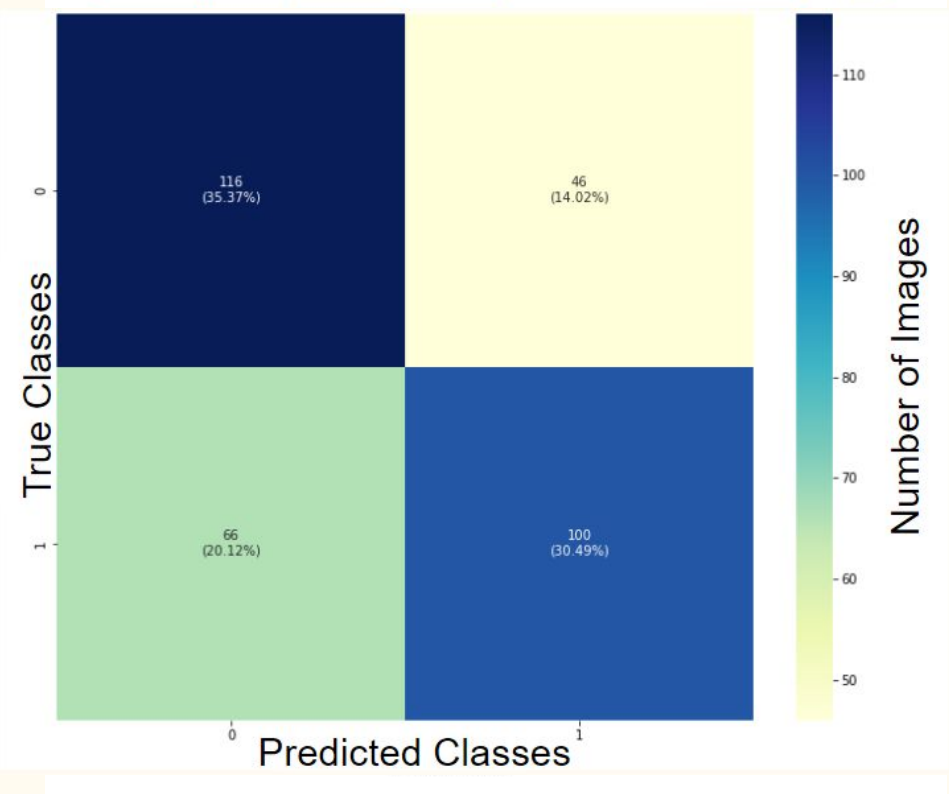
1 : Likely dE/ETGcluster candidate

~~4 : Likely merging system~~

1 : Possible dE/ETGcluster candidate

0 : Background galaxy with possibly weak substructure

Confusion Matrix for Visualization of Performance



Label mapping for 7 classes:

0 : Cluster or background LTG

0 : Likely background ETG or unresolved source

0 : Likely cluster or background edge-on disk galaxy

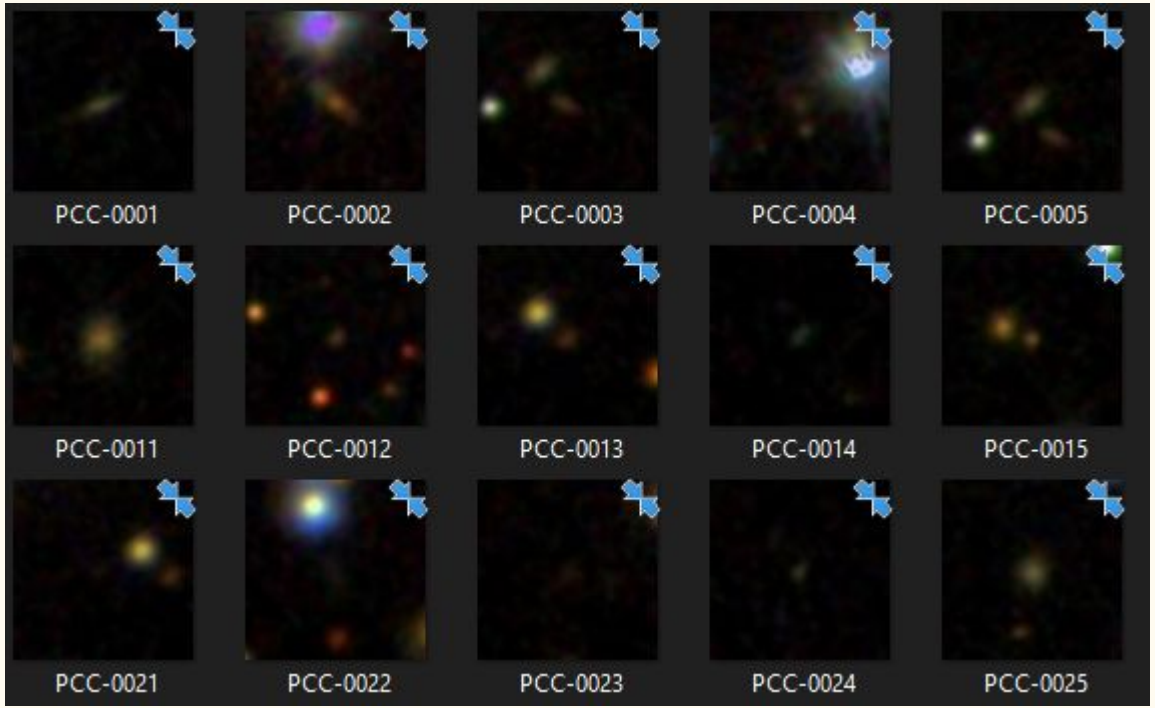
1 : Likely dE/ETGcluster candidate

~~4 : Likely merging system~~

1 : Possible dE/ETGcluster candidate

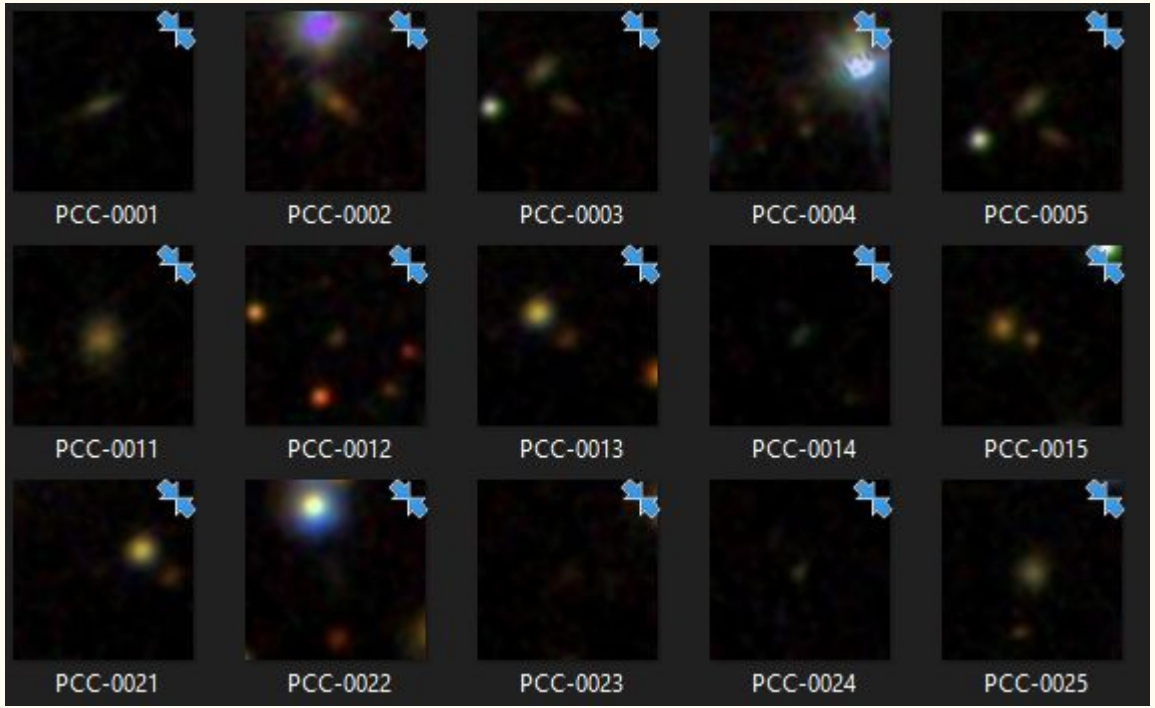
0 : Background galaxy with possibly weak substructure

Data Cleaning - Brightness Threshold



PCC-0023

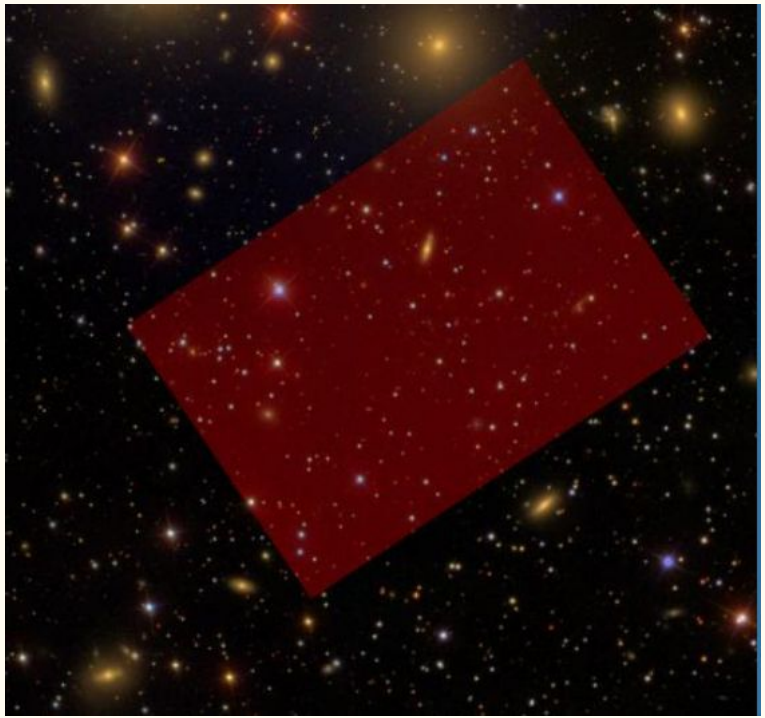
Data Cleaning - Brightness Threshold

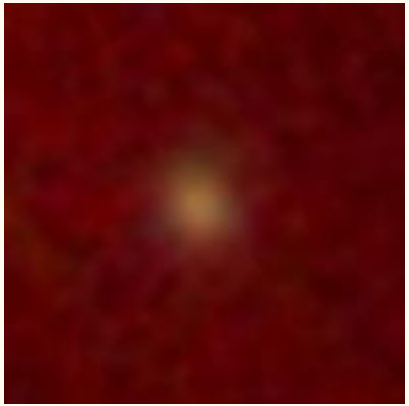


PCC-0023

$r < 19.4$, Reduced train set to 272 objects

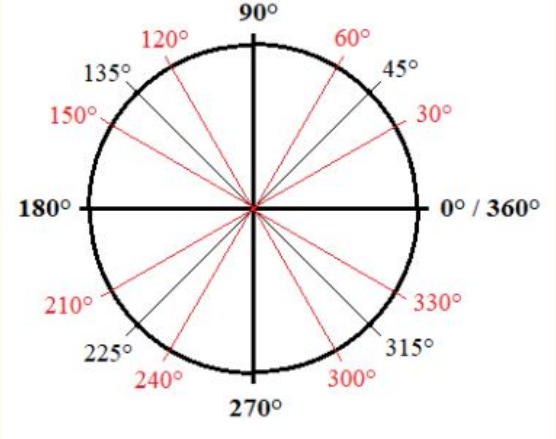
Data Cleaning - Red Objects





Reduced train set to 230 objects

Data Augmentation - Rotations



Multiplied train set by 15
(Now 3450 images)

Hyperparameters

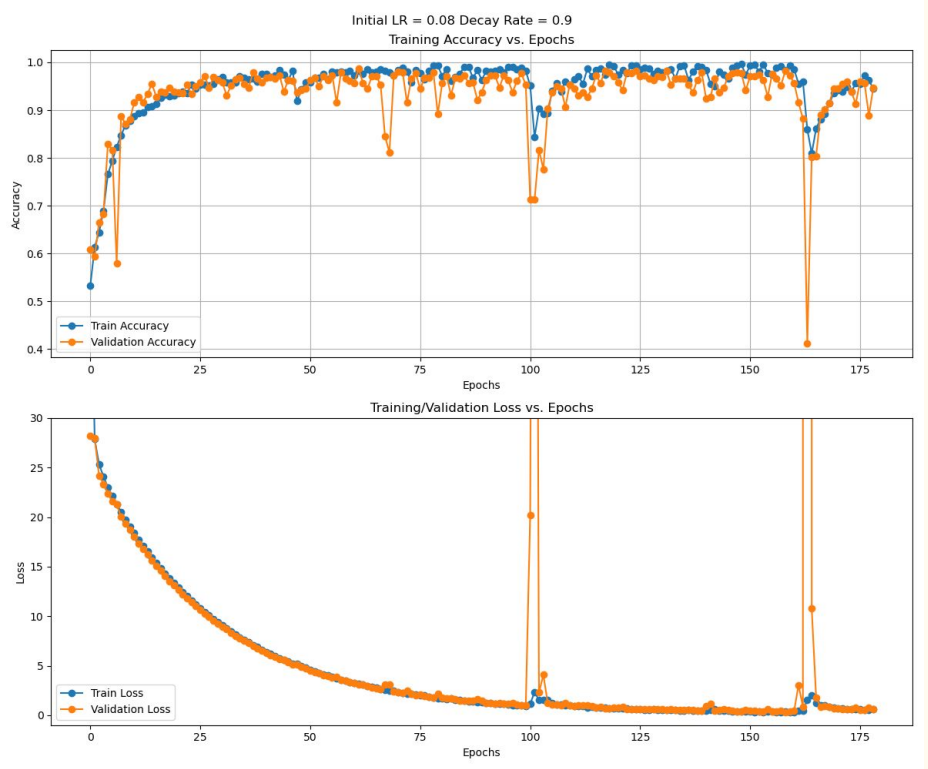
Train/test split of 75/25

Train/valid split of 75/25

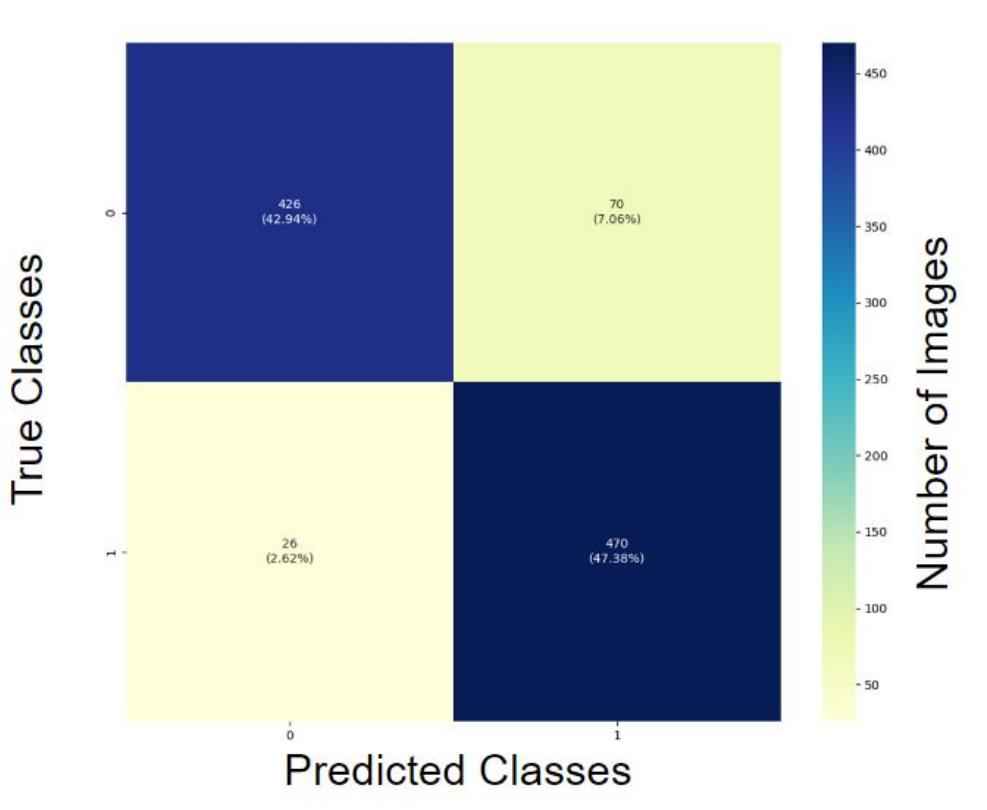
20 image batch size train and valid

200 epochs

$$LR = (0.8)e^{-0.9t}$$

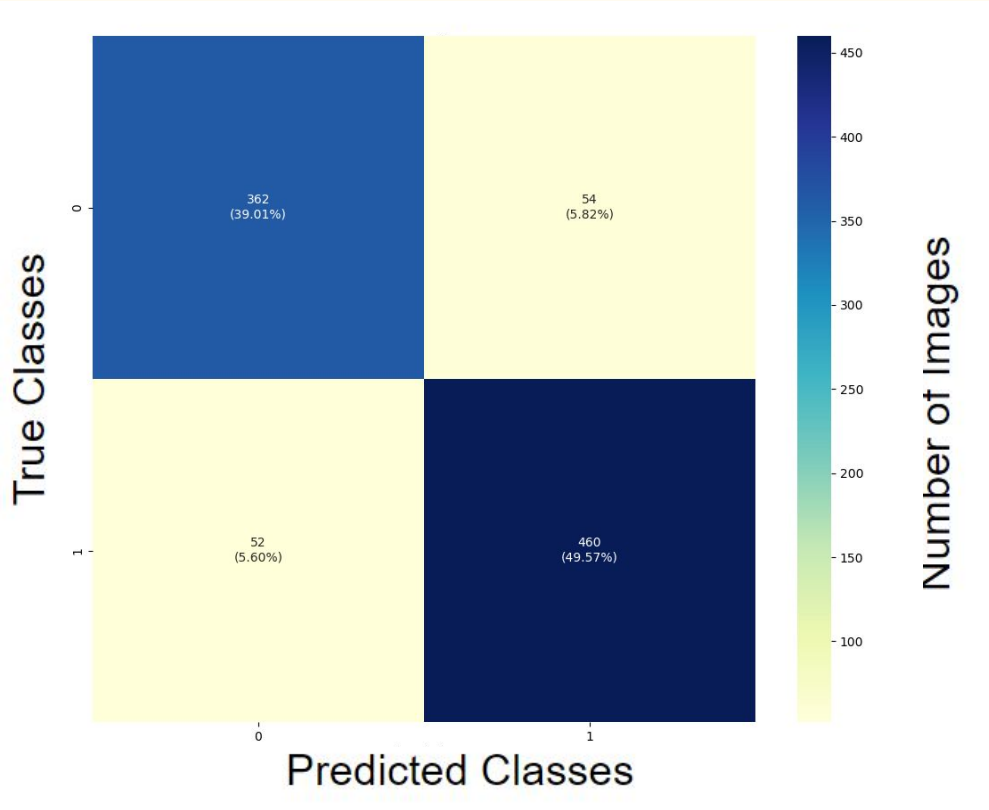


Performance on PCC Brights



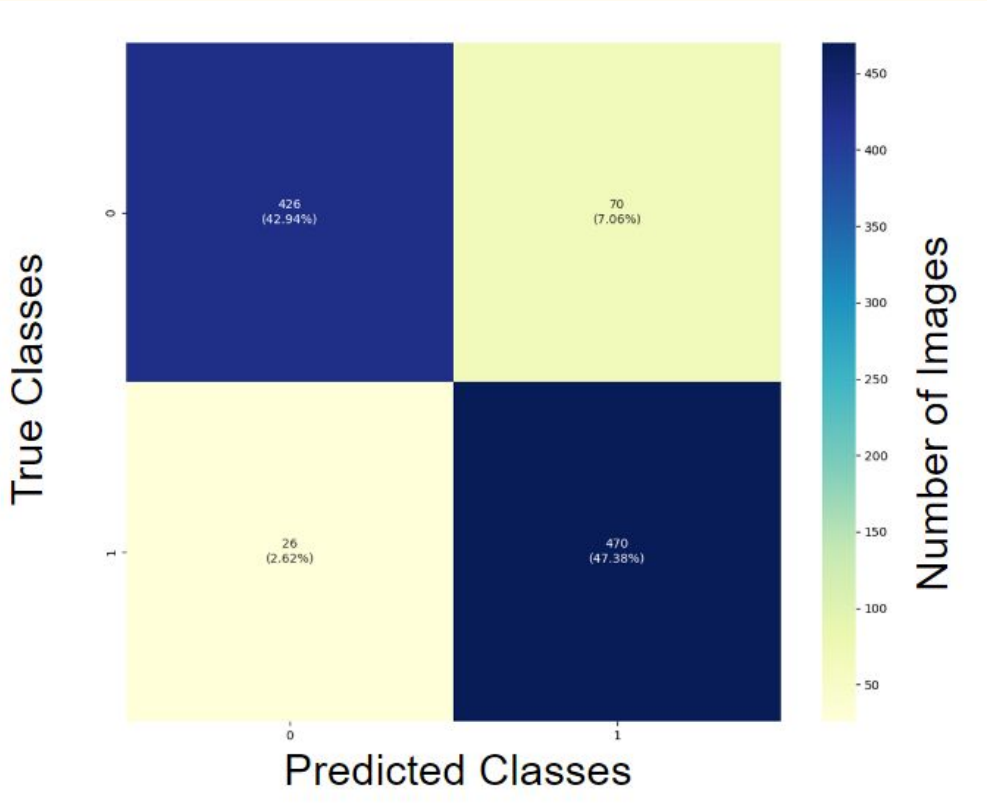
An average of 95% accuracy from every model on the test set from train/test split (balanced classes)

Performance on PCC Brights (grayscale)



An average of ~5% accuracy dip when trained on grayscale versions of images

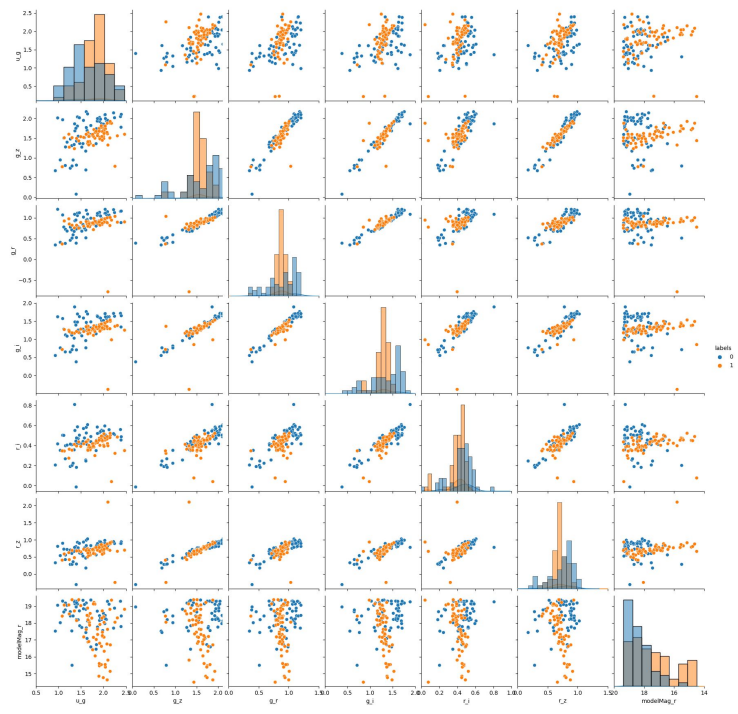
Performance on PCC Brights



An average of 95% accuracy from every model on the test set from train/test split (balanced)

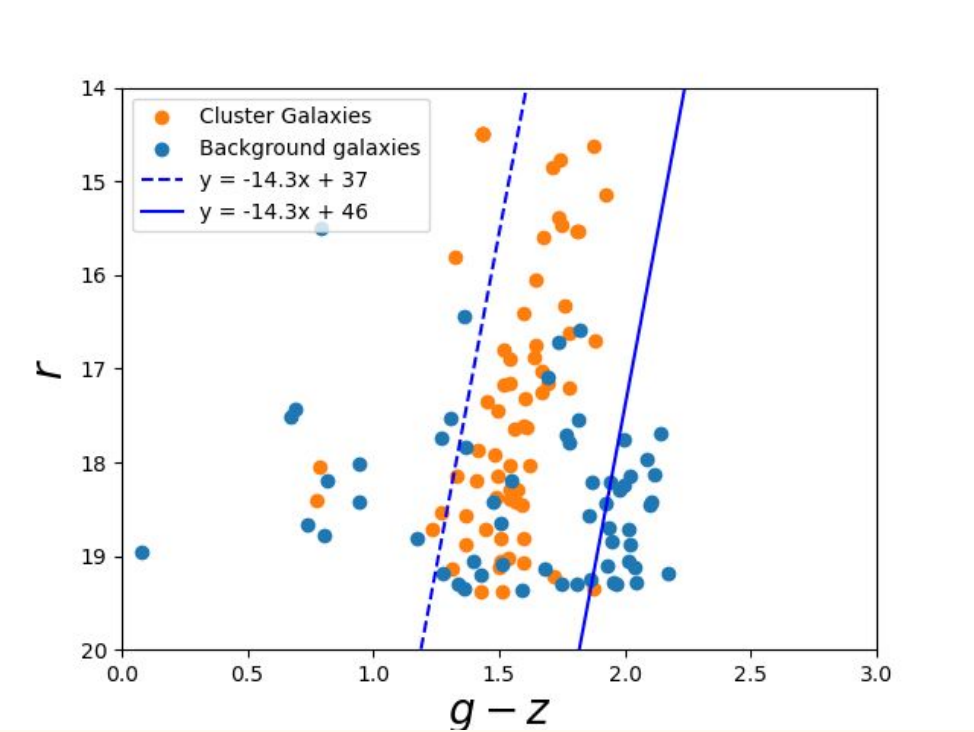
Needed a way to independently verify

Finding a New Test Set - Color Pairplots



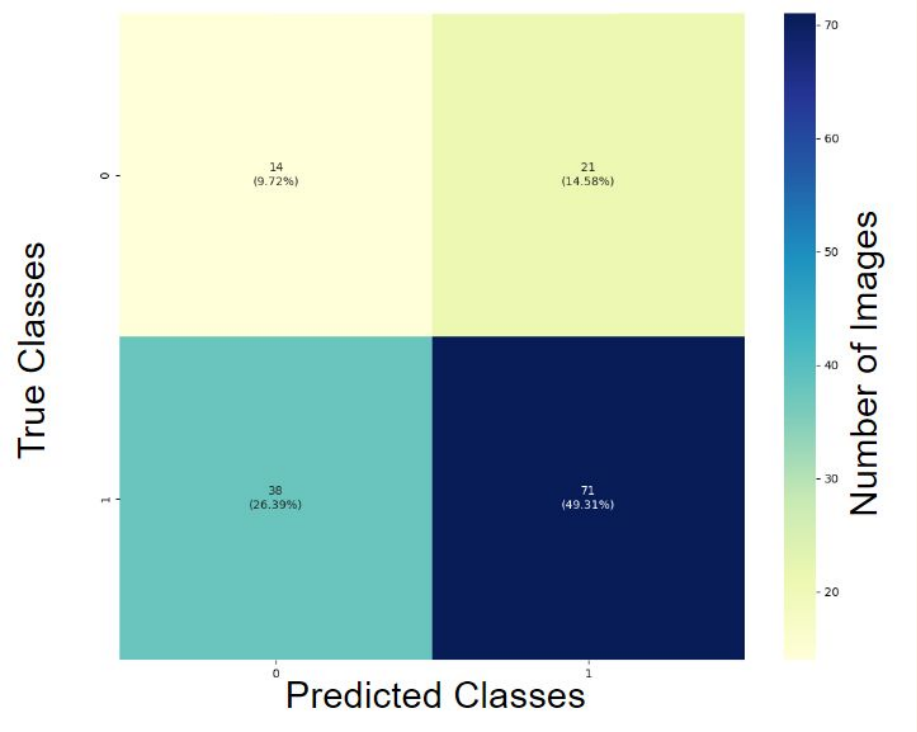
Finding an independent test set - any dividing lines between the classes?

Finding a New Test Set - Color Pairplots



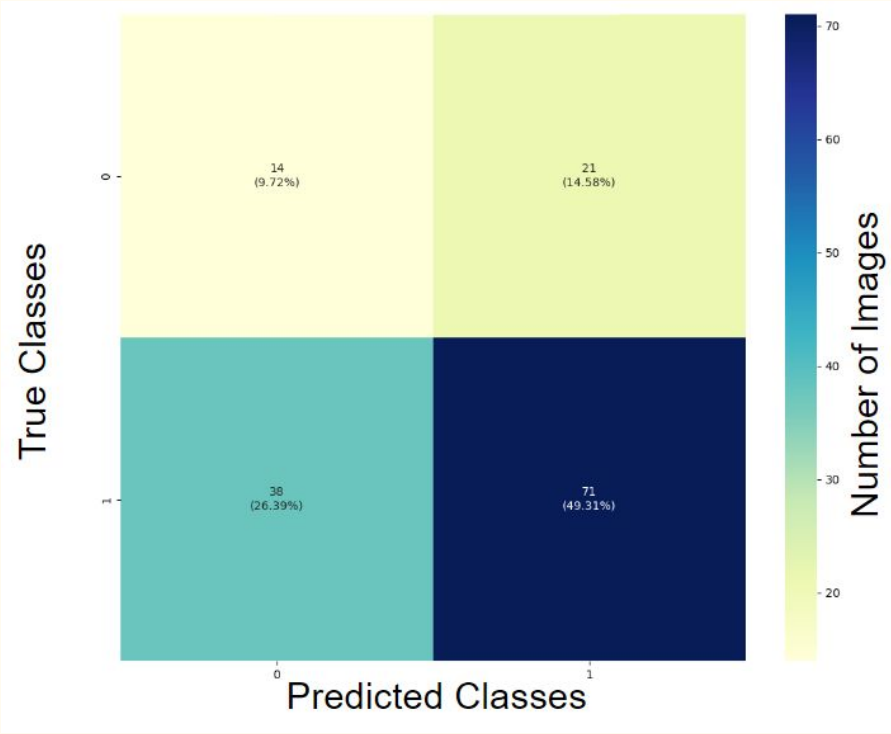
Red Sequence most obvious in r vs (g-z) color magnitude relation

First Independent Test - Low Performance

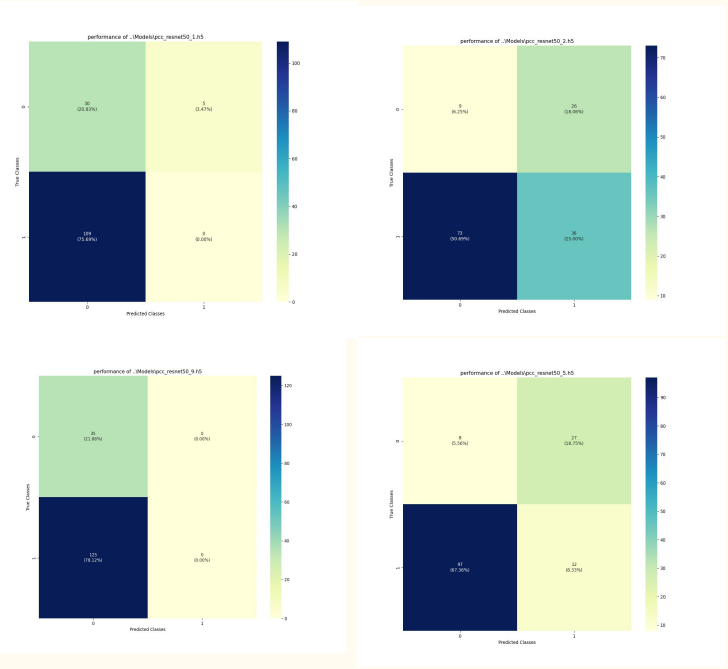


Off diagonal elements take up 40% of the total predictions

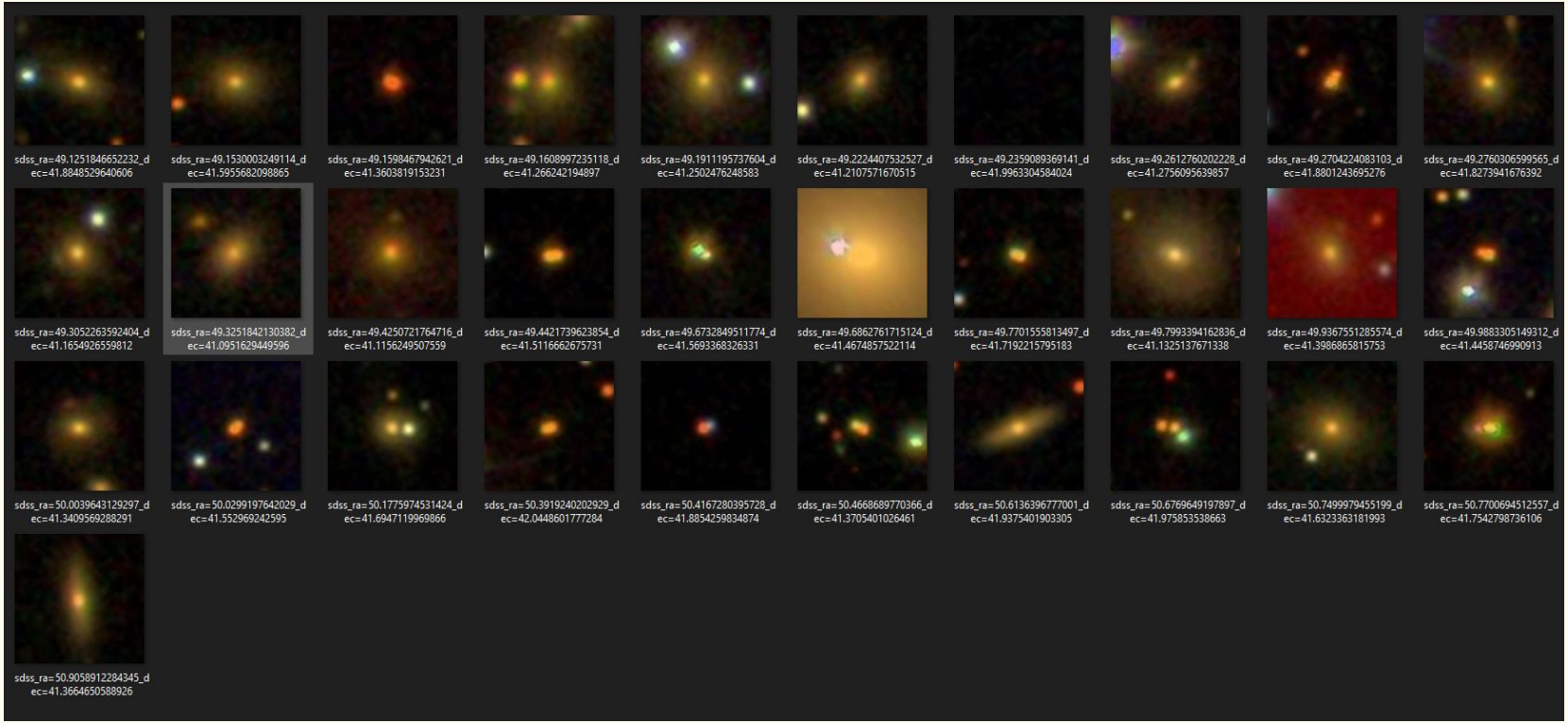
First Independent Test - Low Performance



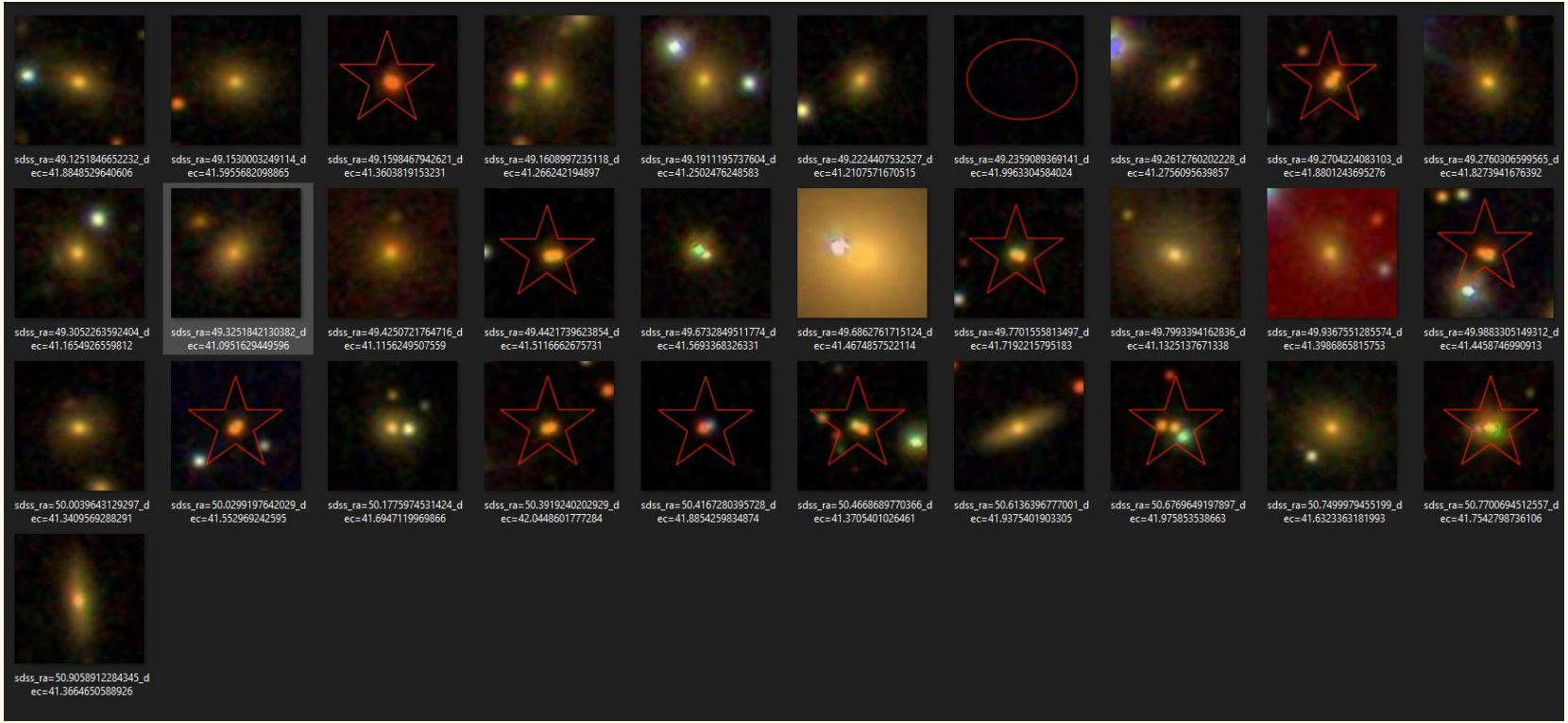
Other models with the same parameters:



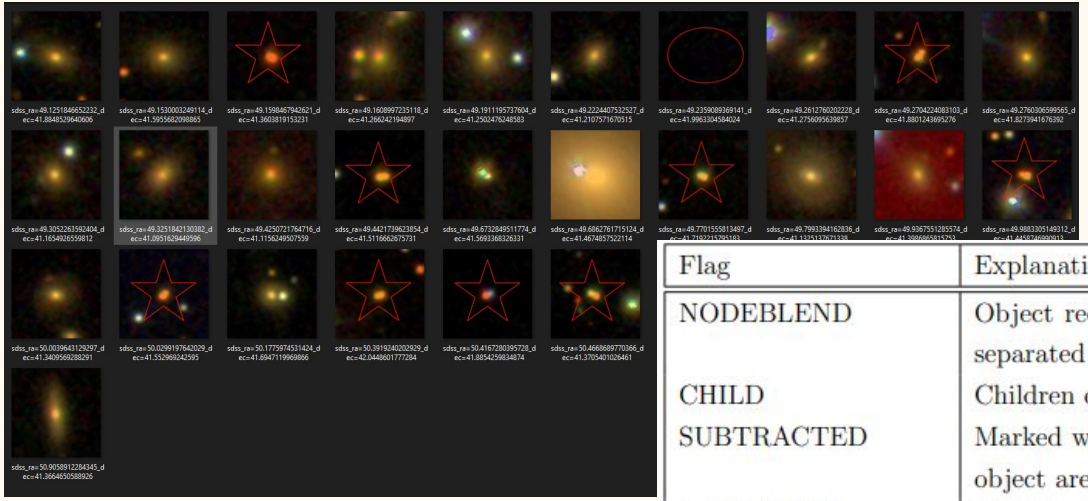
Stars Mislabeled as Galaxies



Stars Mislabeled as Galaxies

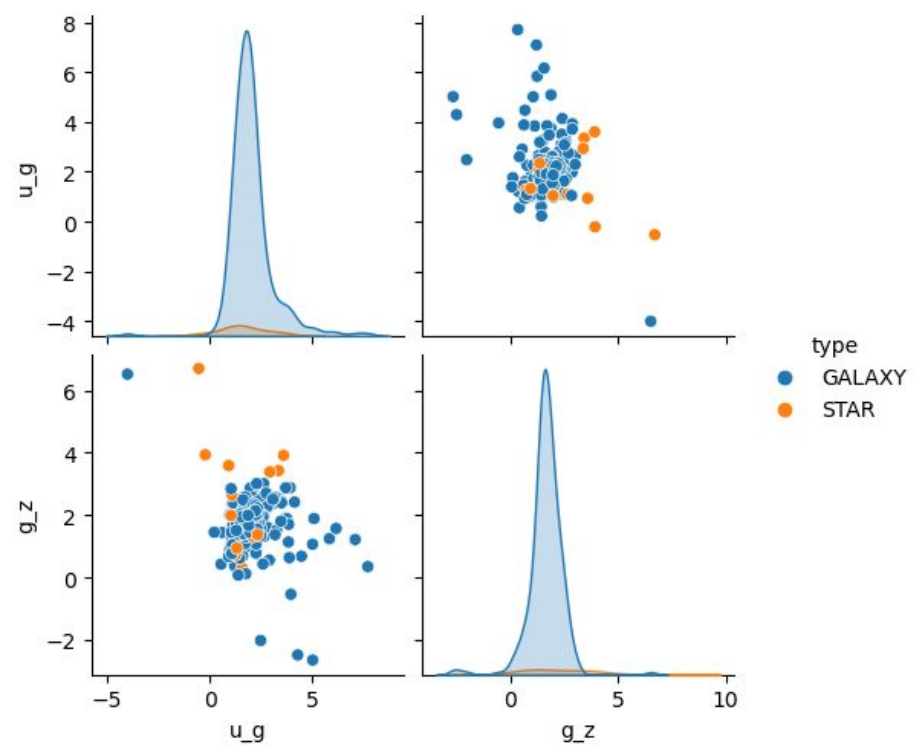


Stars Mislabeled as Galaxies - Flags?

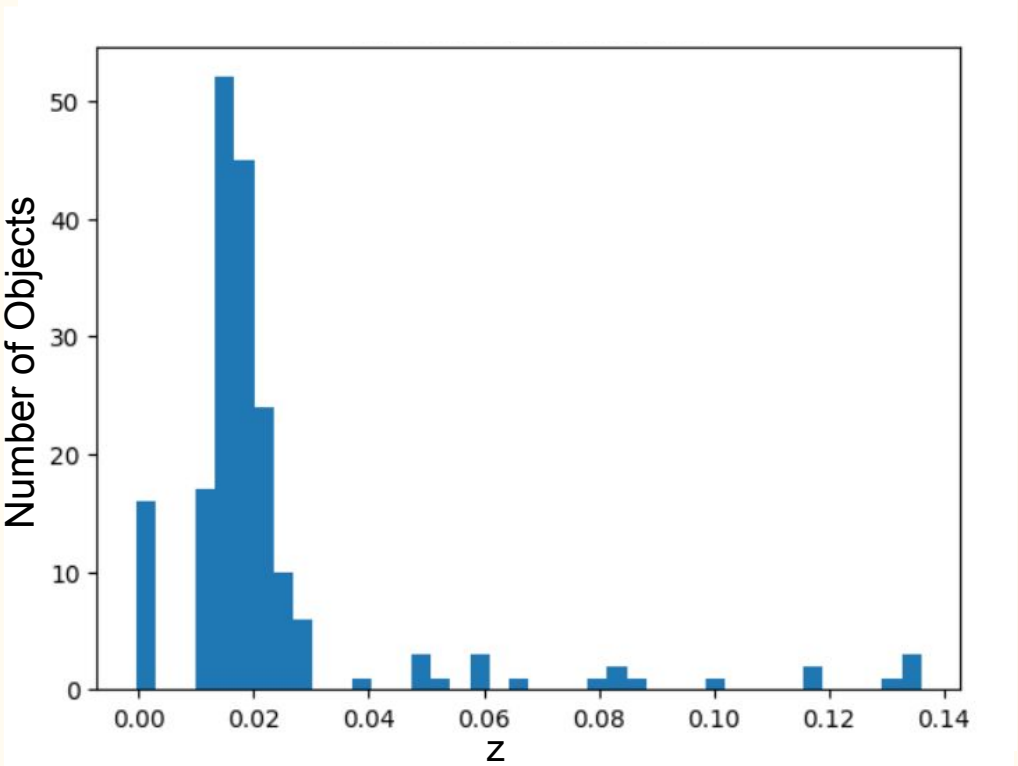


Flag	Explanation
NODEBLEND	Object recognized as composite and was unable to be separated (Deblended).
CHILD	Children of a deblended object.
SUBTRACTED	Marked when the extended wings around a bright star object are subtracted out.
SATURATED	Questionable photometry, i.e. one of the bands was found to have saturated pixels.
CR	Image in question contains a cosmic ray that was interpolated over.
NOPETRO	The object's Petrosian radius was unable to be determined.

Stars Mislabeled as Galaxies



Spectroscopic Redshift (z)

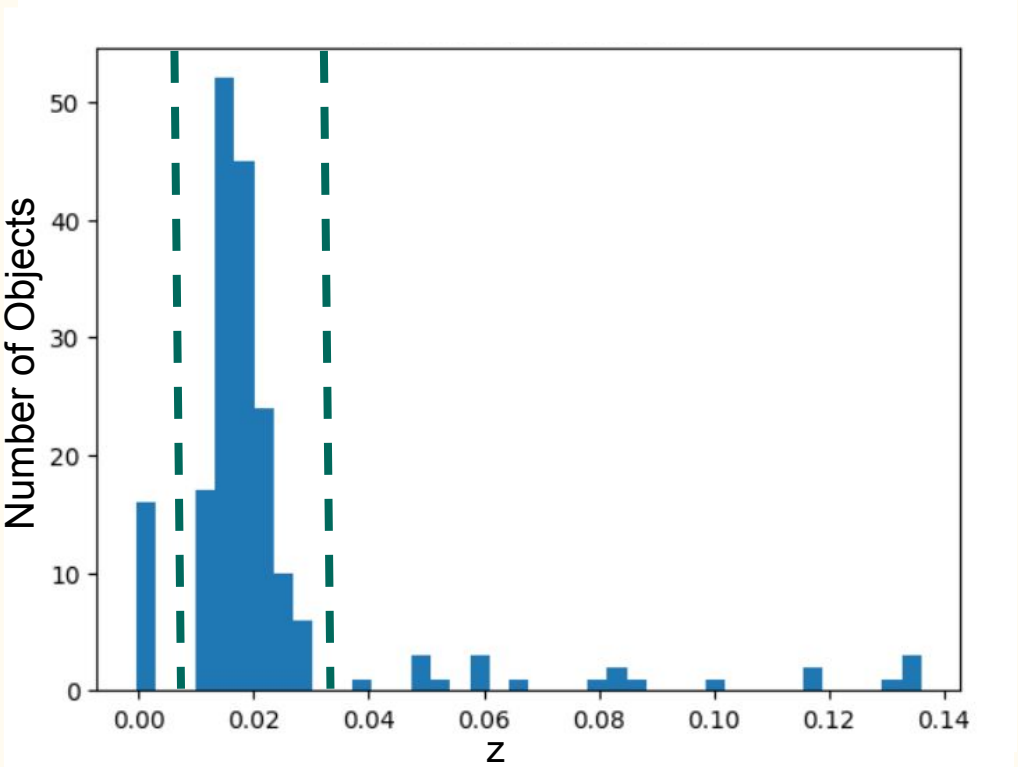


Cluster Galaxies:
$0.01 < z < 0.033$

Stars:
$z < 0.01$

Background Galaxies:
$z > 0.033$

Spectroscopic Redshift (z)

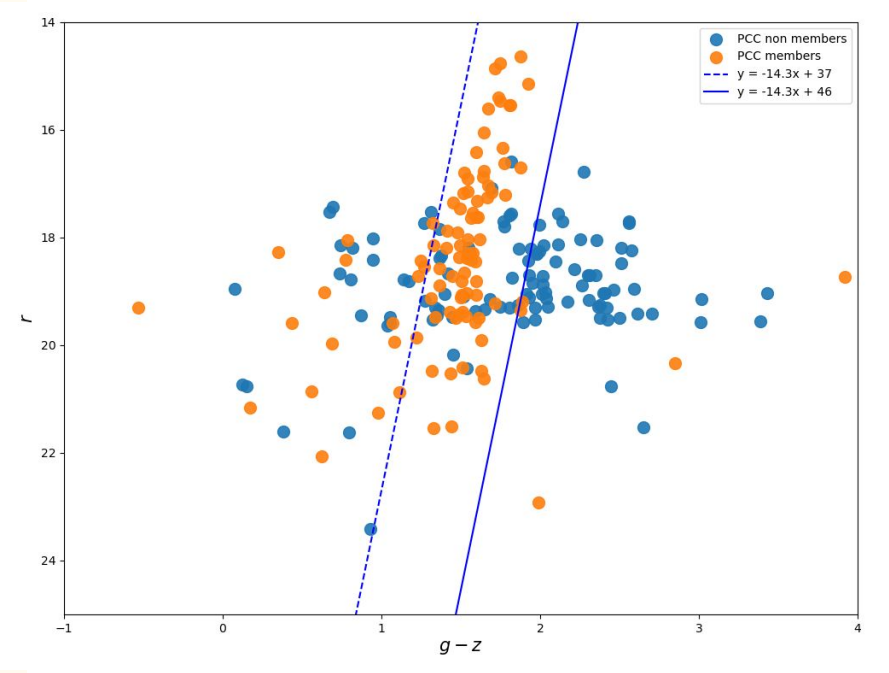


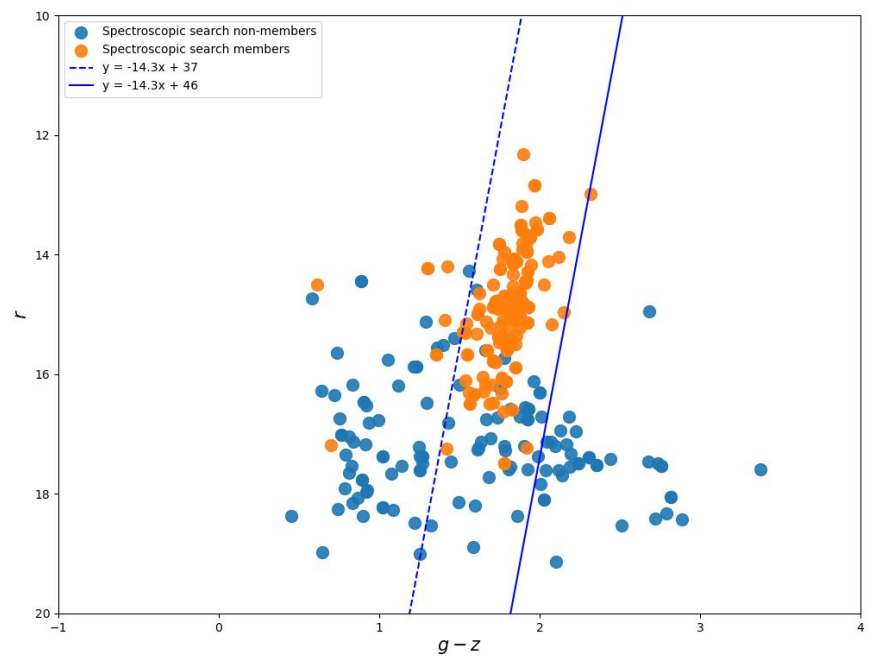
Cluster Galaxies:
$0.01 < z < 0.033$

Stars:
$z < 0.01$

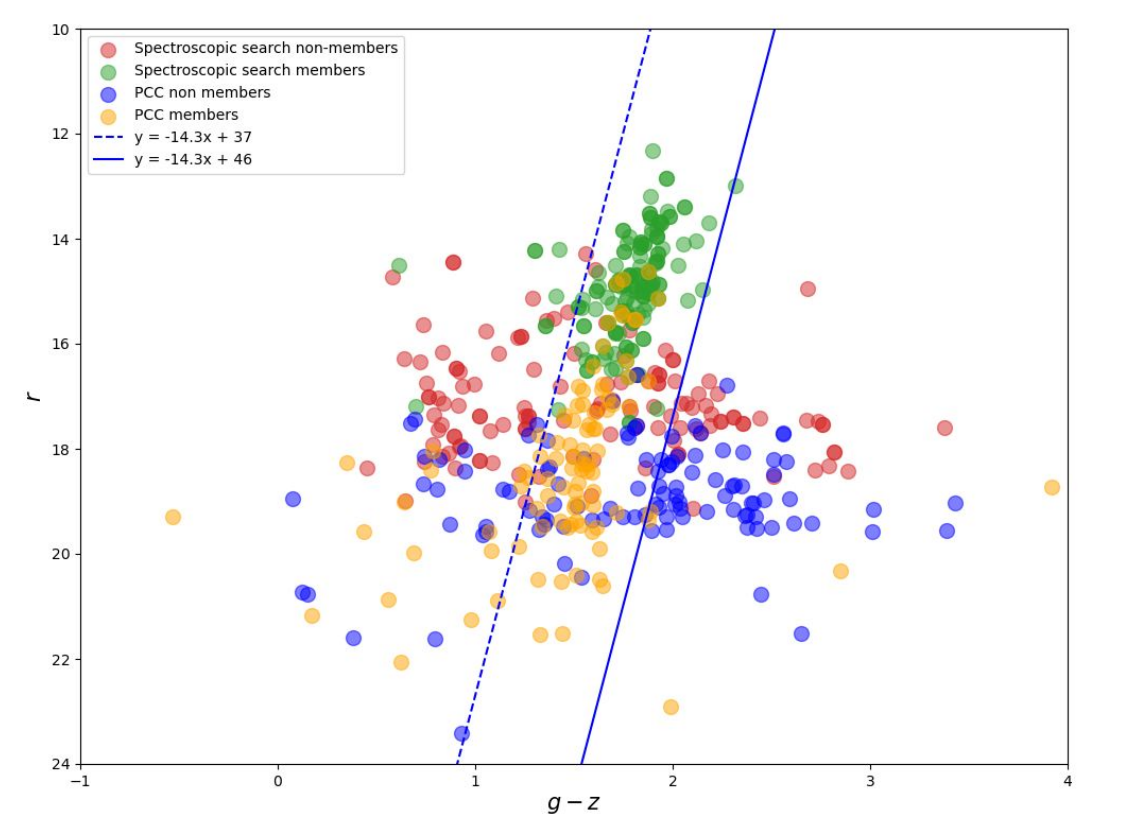
Background Galaxies:
$z > 0.033$

Adding the Radial “Spectroscopic Search” to Training Set

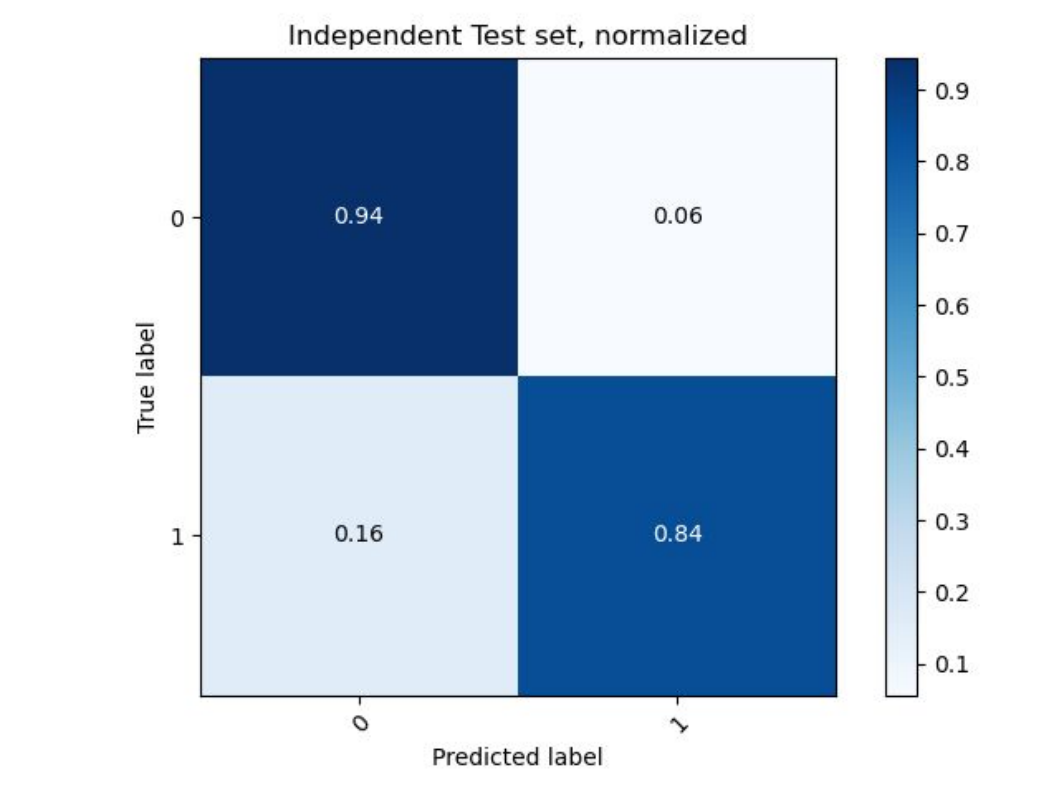




Adding the Radial “Spectroscopic Search” to Training Set



Independent Test Set Results



	precision	recall	f1-score	support
0	0.94	0.94	0.94	247
1	0.84	0.84	0.84	88
accuracy			0.92	335
macro avg	0.89	0.89	0.89	335
weighted avg	0.92	0.92	0.92	335

Performance in Different Color Regions

Blue Cloud - Indep. Test Set

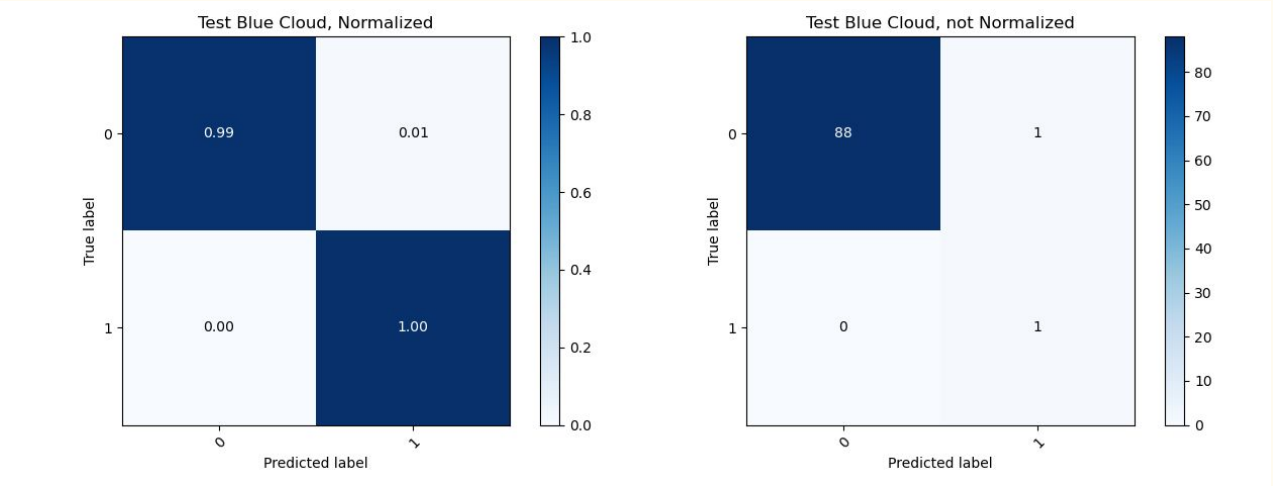
	precision	recall	f1-score	support
0	1.00	0.99	0.99	89
1	0.50	1.00	0.67	1
accuracy			0.99	90
macro avg	0.75	0.99	0.83	90
weighted avg	0.99	0.99	0.99	90

Red Sequence - Indep. Test Set

	precision	recall	f1-score	support
0	0.85	0.88	0.87	78
1	0.87	0.83	0.85	71
accuracy			0.86	149
macro avg	0.86	0.86	0.86	149
weighted avg	0.86	0.86	0.86	149

Red Shoulder - Indep. Test Set

	precision	recall	f1-score	support
0	0.97	0.95	0.96	80
1	0.78	0.88	0.82	16
accuracy			0.94	96
macro avg	0.88	0.91	0.89	96
weighted avg	0.94	0.94	0.94	96



Performance in Different Color Regions

Blue Cloud - Indep. Test Set

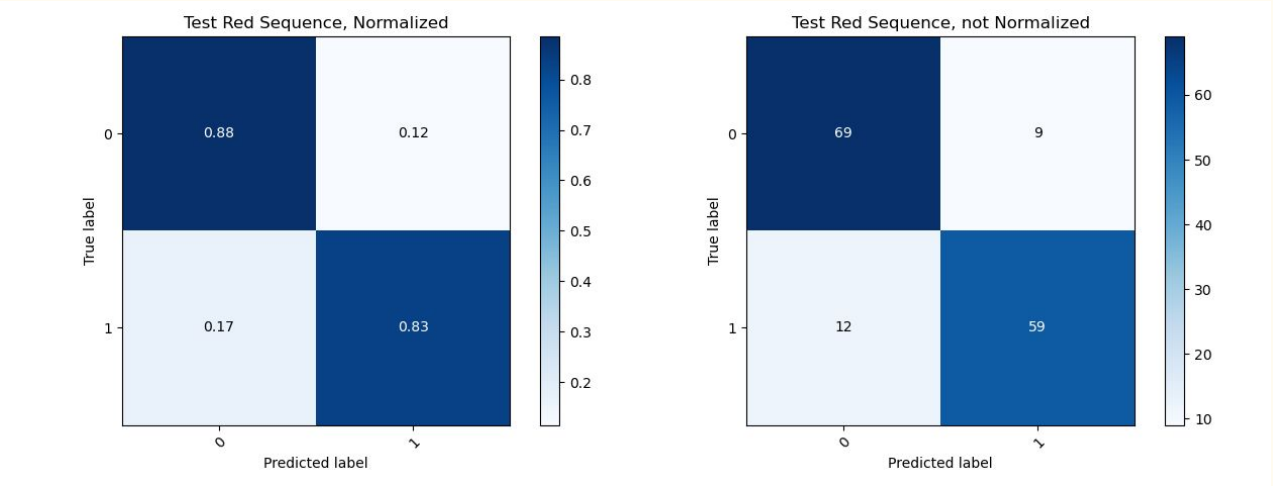
	precision	recall	f1-score	support
0	1.00	0.99	0.99	89
1	0.50	1.00	0.67	1
accuracy			0.99	90
macro avg	0.75	0.99	0.83	90
weighted avg	0.99	0.99	0.99	90

Red Sequence - Indep. Test Set

	precision	recall	f1-score	support
0	0.85	0.88	0.87	78
1	0.87	0.83	0.85	71
accuracy			0.86	149
macro avg	0.86	0.86	0.86	149
weighted avg	0.86	0.86	0.86	149

Red Shoulder - Indep. Test Set

	precision	recall	f1-score	support
0	0.97	0.95	0.96	80
1	0.78	0.88	0.82	16
accuracy			0.94	96
macro avg	0.88	0.91	0.89	96
weighted avg	0.94	0.94	0.94	96



Performance in Different Color Regions

Blue Cloud - Indep. Test Set

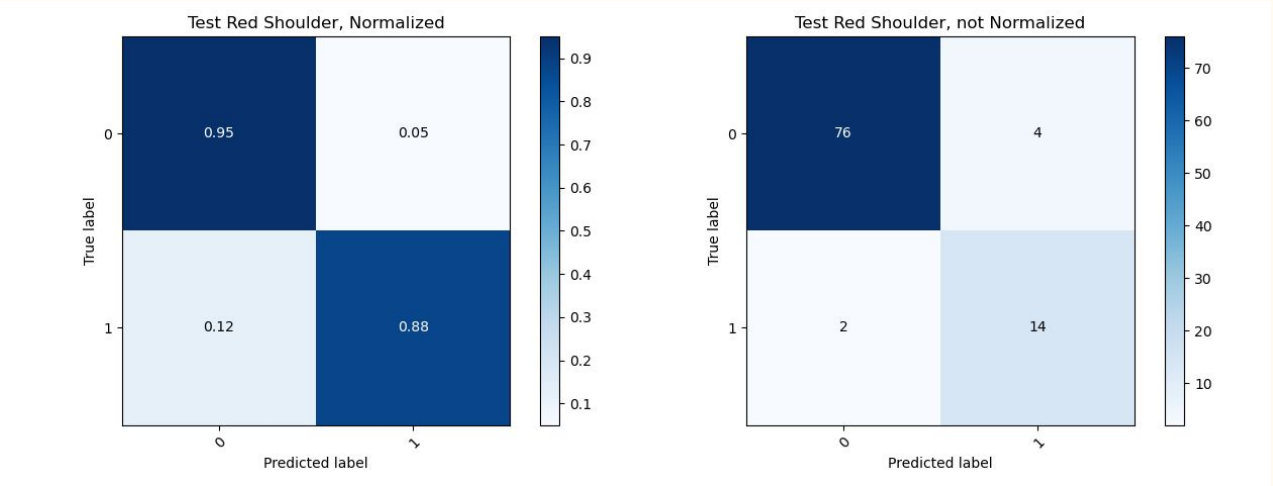
	precision	recall	f1-score	support
0	1.00	0.99	0.99	89
1	0.50	1.00	0.67	1
accuracy			0.99	90
macro avg	0.75	0.99	0.83	90
weighted avg	0.99	0.99	0.99	90

Red Sequence - Indep. Test Set

	precision	recall	f1-score	support
0	0.85	0.88	0.87	78
1	0.87	0.83	0.85	71
accuracy			0.86	149
macro avg	0.86	0.86	0.86	149
weighted avg	0.86	0.86	0.86	149

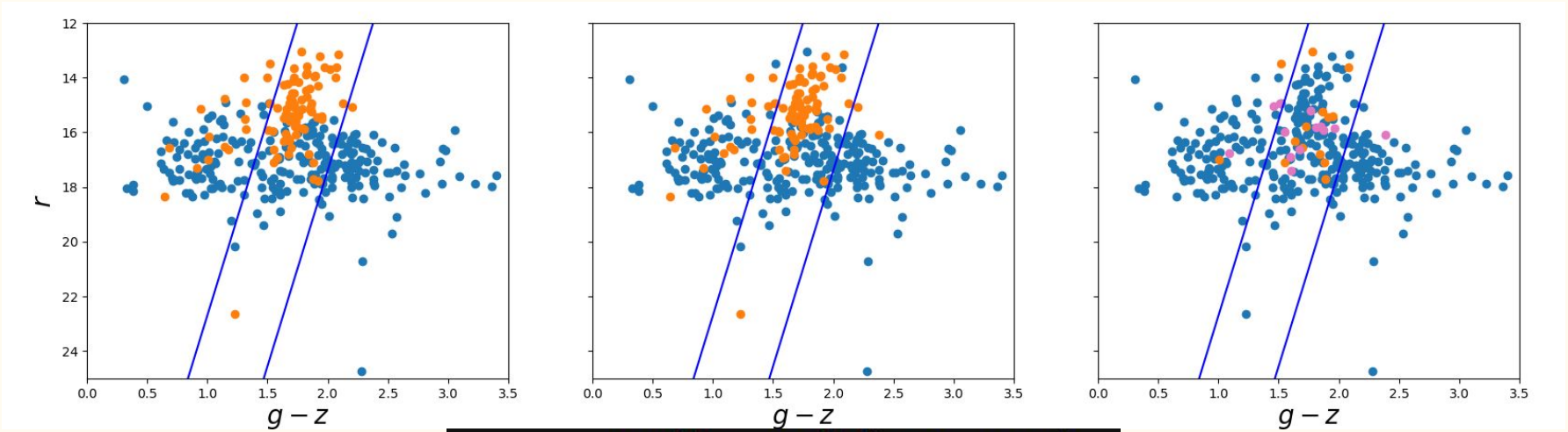
Red Shoulder - Indep. Test Set

	precision	recall	f1-score	support
0	0.97	0.95	0.96	80
1	0.78	0.88	0.82	16
accuracy			0.94	96
macro avg	0.88	0.91	0.89	96
weighted avg	0.94	0.94	0.94	96



Independent Test Set Results

True Labels



Predicted Labels

Differences

	precision	recall	f1-score	support
0	0.94	0.94	0.94	247
1	0.84	0.84	0.84	88
accuracy			0.92	335
macro avg	0.89	0.89	0.89	335
weighted avg	0.92	0.92	0.92	335

Concluding Remarks and Future Work

Add more images from blue cloud/red shoulder regions

Subaru Hyper Suprime-Cam imaging

European Sky Agency (ESA) Euclid imaging

ResNet-15 or potentially more efficient models

Transformers/visual attention maps

Different hyperparameters