



Analysis of GAN as a Tool for Data Augmentation

Yagyasha Rastogi

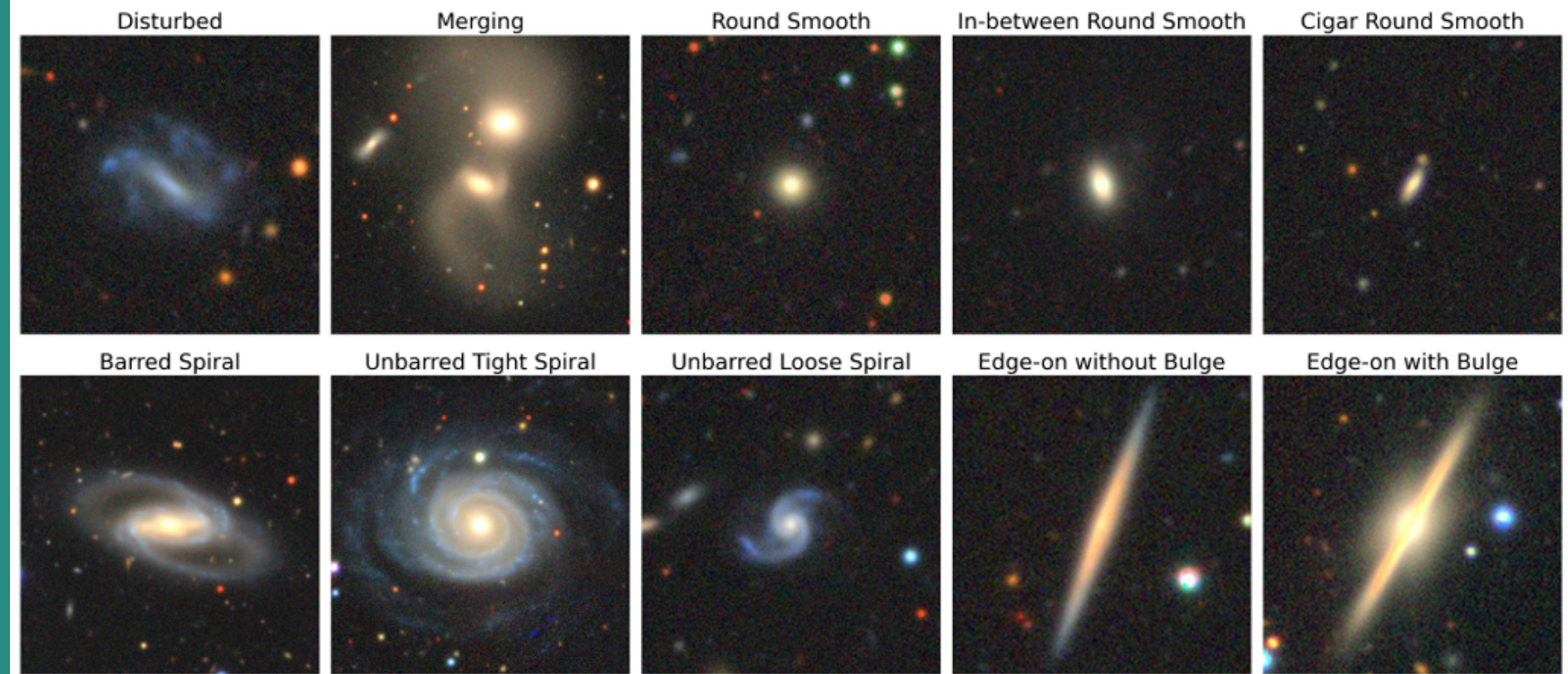
Data Augmentation

Sample Data

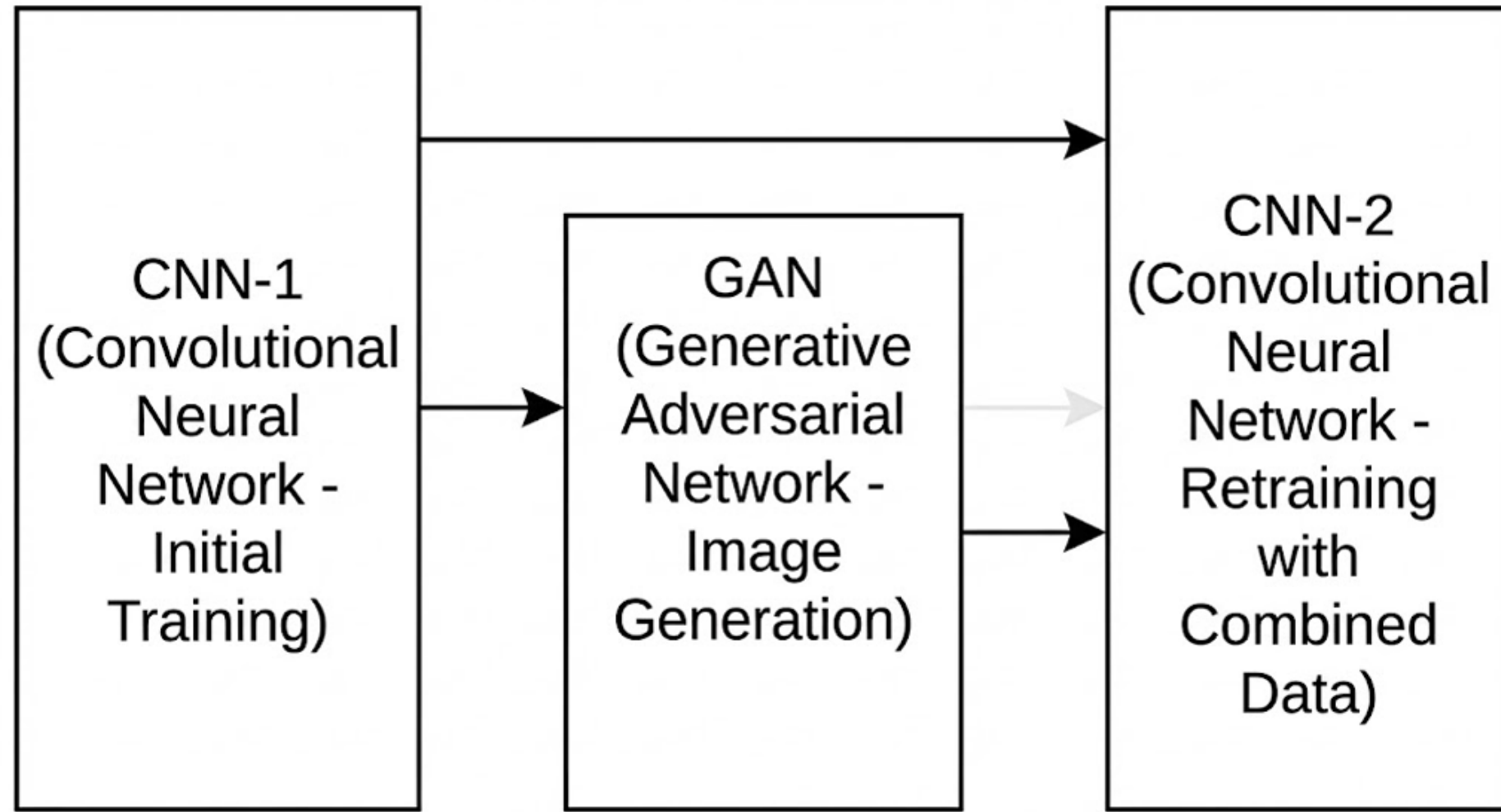


Galaxy10 DECals

Example images of each class from Galaxy10 DECals



CNN - GAN - CNN Sandwich



Should we trade-off time with quality?

Dilemma faced during GAN training:



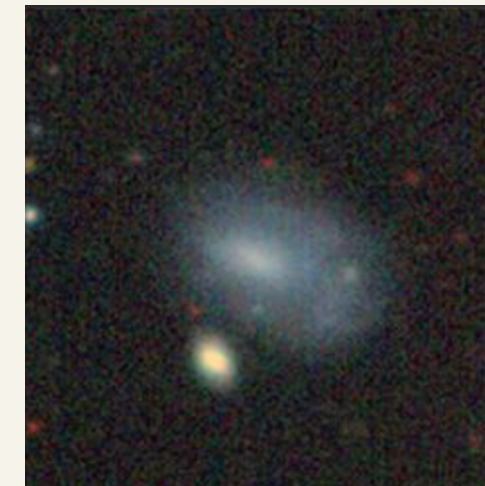
Train whole dataset at once...



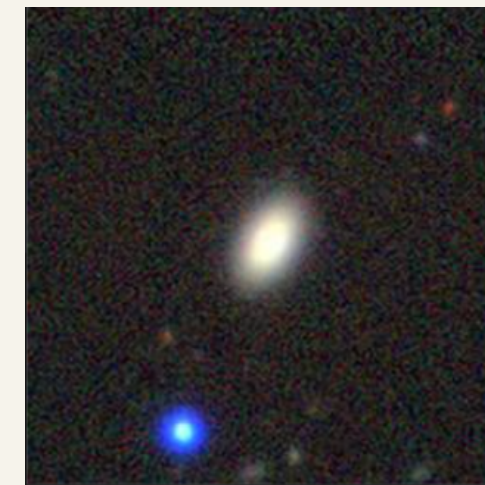
Train GAN separately for each class
(10 times!)

GAN Results

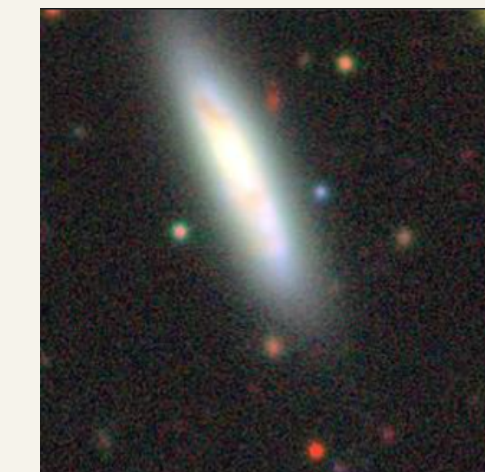
Galaxy10 DECaLS



Class 0

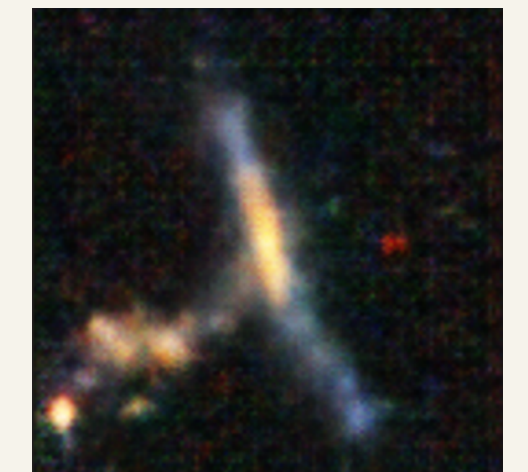
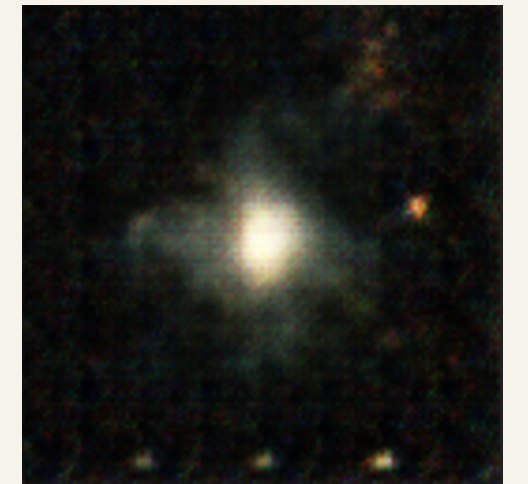


Class 3



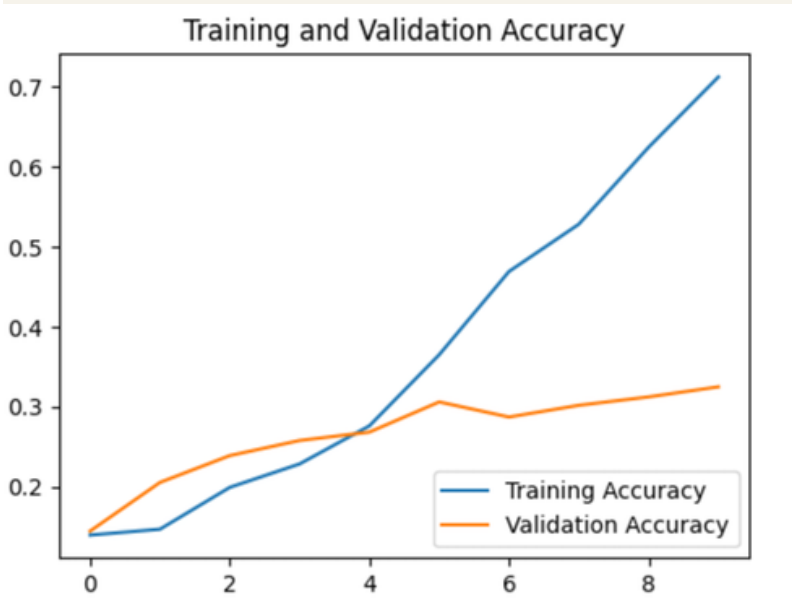
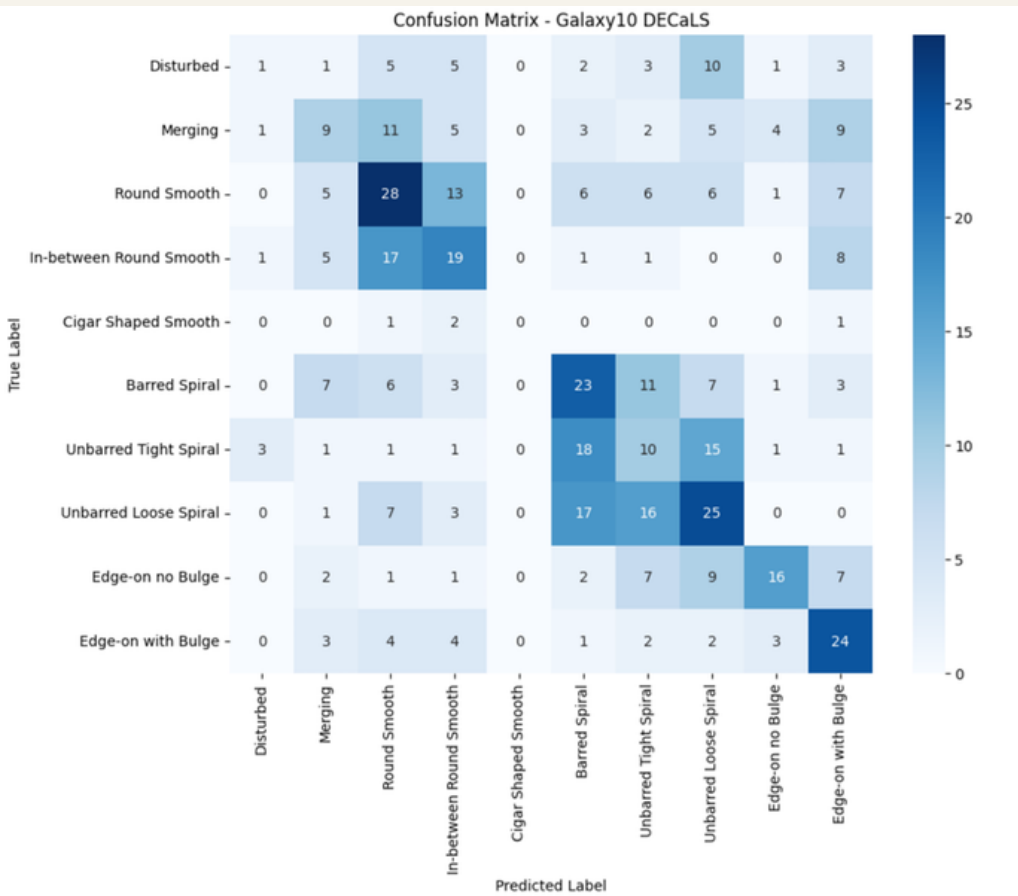
Class 8

GAN Generated



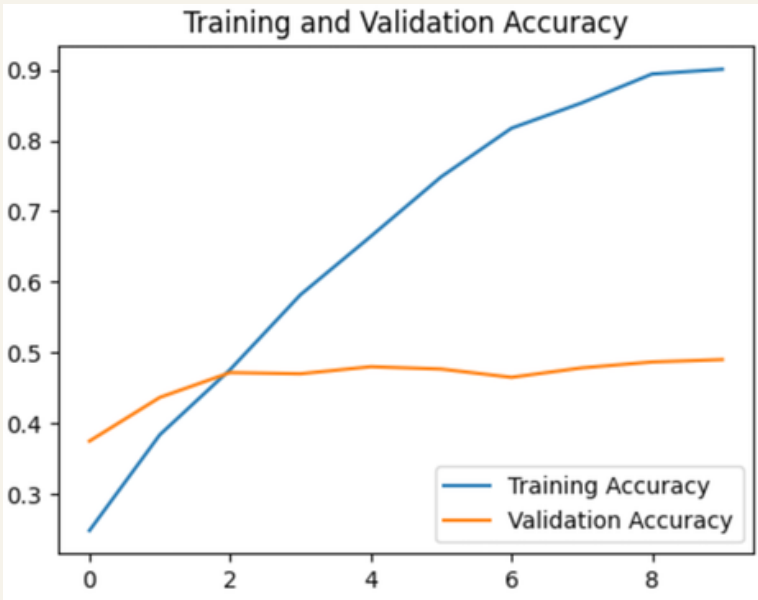
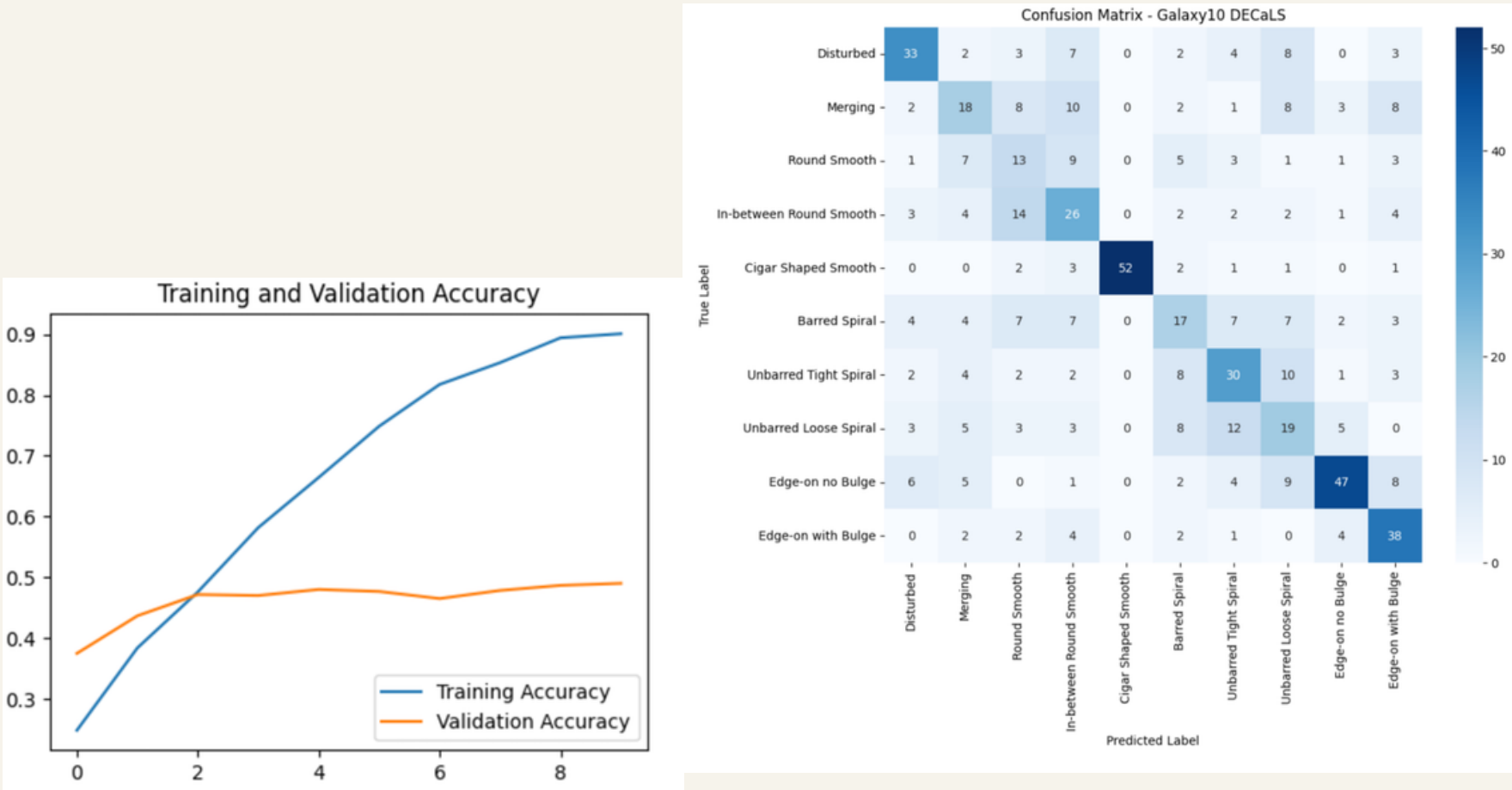
The accuracy of CNN increased from 71.01% to ~91% after use of GAN for data augmentation...

CNN-1
(Without Data Augmentation)



	precision	recall	f1-score	support
Disturbed	0.17	0.03	0.05	31
Merging	0.26	0.18	0.22	49
Round Smooth	0.35	0.39	0.37	72
In-between Round Smooth	0.34	0.37	0.35	52
Cigar Shaped Smooth	0.00	0.00	0.00	4
Barred Spiral	0.32	0.38	0.34	61
Unbarred Tight Spiral	0.17	0.20	0.18	51
Unbarred Loose Spiral	0.32	0.36	0.34	69
Edge-on no Bulge	0.59	0.36	0.44	45
Edge-on with Bulge	0.38	0.56	0.45	43
accuracy			0.32	477
macro avg	0.29	0.28	0.28	477
weighted avg	0.32	0.32	0.31	477

CNN-2
(With Data Augmentation)



	precision	recall	f1-score	support
Disturbed	0.61	0.53	0.57	62
Merging	0.35	0.30	0.32	60
Round Smooth	0.24	0.30	0.27	43
In-between Round Smooth	0.36	0.45	0.40	58
Cigar Shaped Smooth	1.00	0.84	0.91	62
Barred Spiral	0.34	0.29	0.31	58
Unbarred Tight Spiral	0.46	0.48	0.47	62
Unbarred Loose Spiral	0.29	0.33	0.31	58
Edge-on no Bulge	0.73	0.57	0.64	82
Edge-on with Bulge	0.54	0.72	0.61	53
accuracy			0.49	598
macro avg	0.49	0.48	0.48	598
weighted avg	0.51	0.49	0.50	598