

Deep learning school challenge

Group "Quattro"

Pavish Subramani, David Rumler, Svenja Hövelmann,

Markus Uehlein, Tobias Held



BERGISCHE
UNIVERSITÄT
WUPPERTAL



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



Kiel University
Christian-Albrechts-Universität zu Kiel

TU NACHWUCHSRING

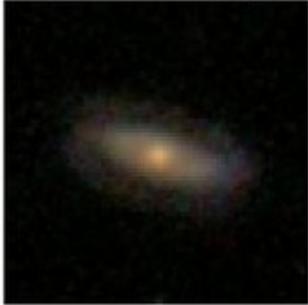
R
TU
P Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

Data set

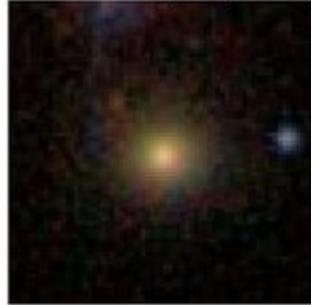
Total number : 21785

Different classes : 10

Disk, Face-on, No Spiral



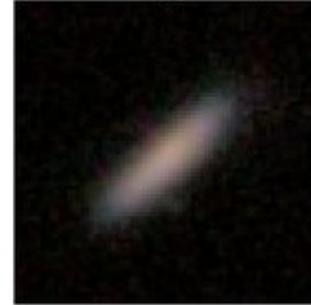
Smooth, Completely round



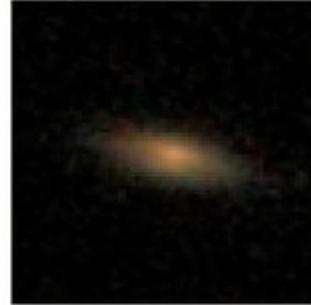
Smooth, in-between round



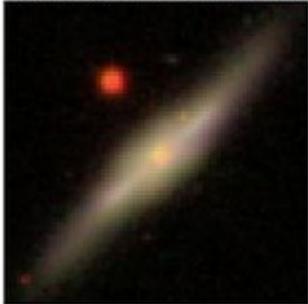
Smooth, Cigar shaped



Disk, Edge-on, Rounded Bulge



Disk, Edge-on, Boxy Bulge



Disk, Edge-on, No Bulge



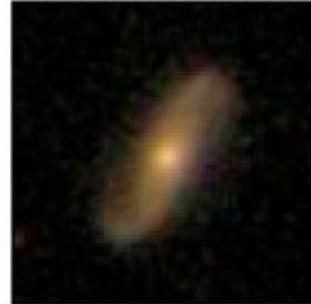
Disk, Face-on, Tight Spiral



Disk, Face-on, Medium Spiral



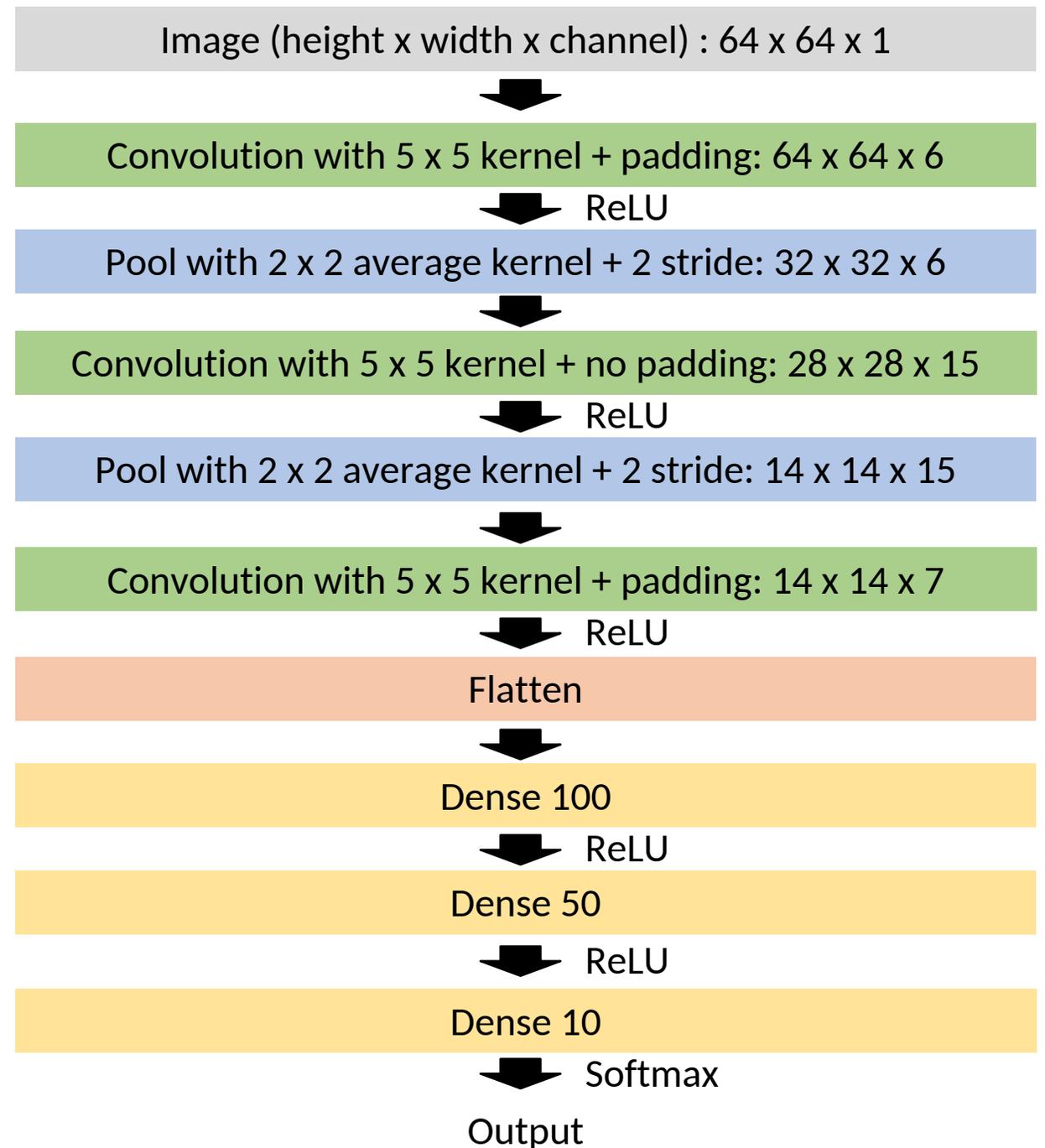
Disk, Face-on, Loose Spiral



Galaxy10 Dataset: Henry Leung/Jo Bovy 2018, Data Source: SDSS/Galaxy Zoo

Neural network

- Keras
 - Learning rate : 0.0003
 - batch_size : 10
 - epoch : 30
 - Algorithm : Adam
 - Loss : binary_crossentropy
 - Metrics : accuracy
- Split data set
 - Train data : 4/5
 - Test data : 1/5

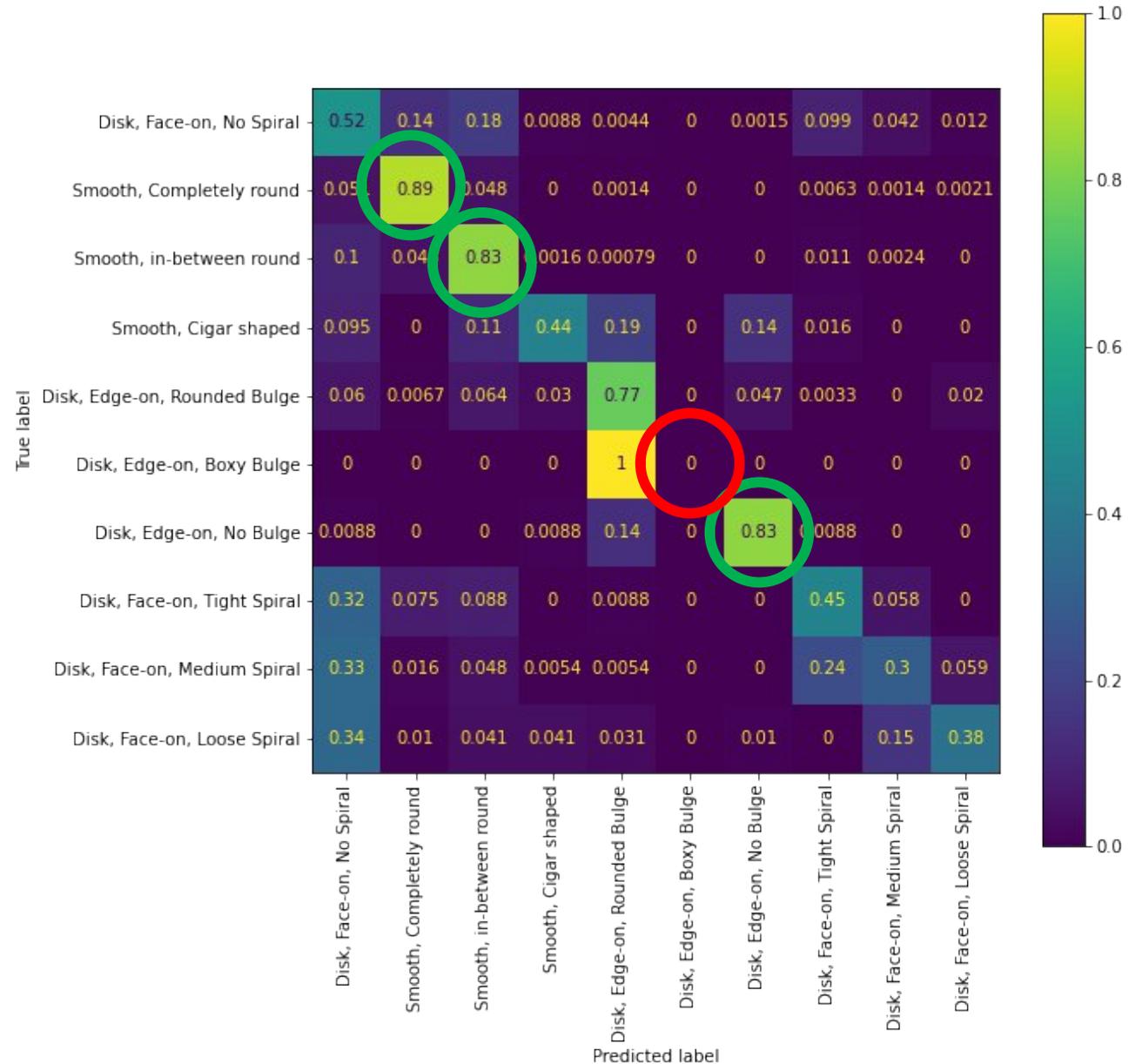


Neural network

- Number of training data: 17428
- Number of test data : 4357

- Accuracy : 0.738
- Loss : 0.296

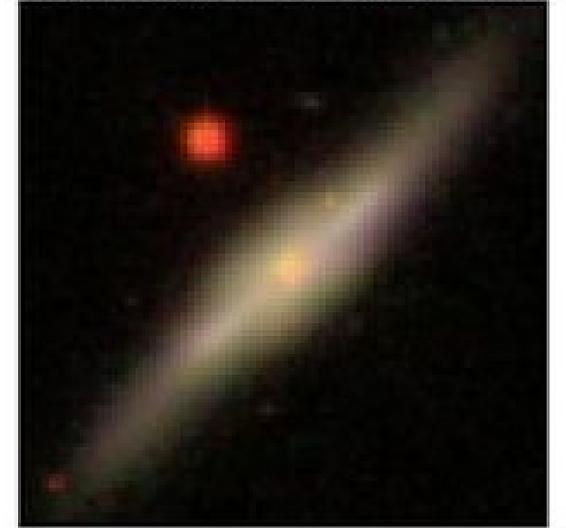
- Networks works fine for test data with a high number of training data
- Network fails for low number of training data



Disk, Edge-on, Boxy Bulge

- What makes this class stand out?
 - Lowest number of train images with only 17
 - Similar shape as Disk, Edge-on, Rounded Bulge
- How can this be addressed?
 - Optimise neural network by:
 - Equalise number of data sets per class in the training sets by copying data sets
 - Rotate and mirror copied data sets to “create” new data sets

Disk, Edge-on, Boxy Bulge



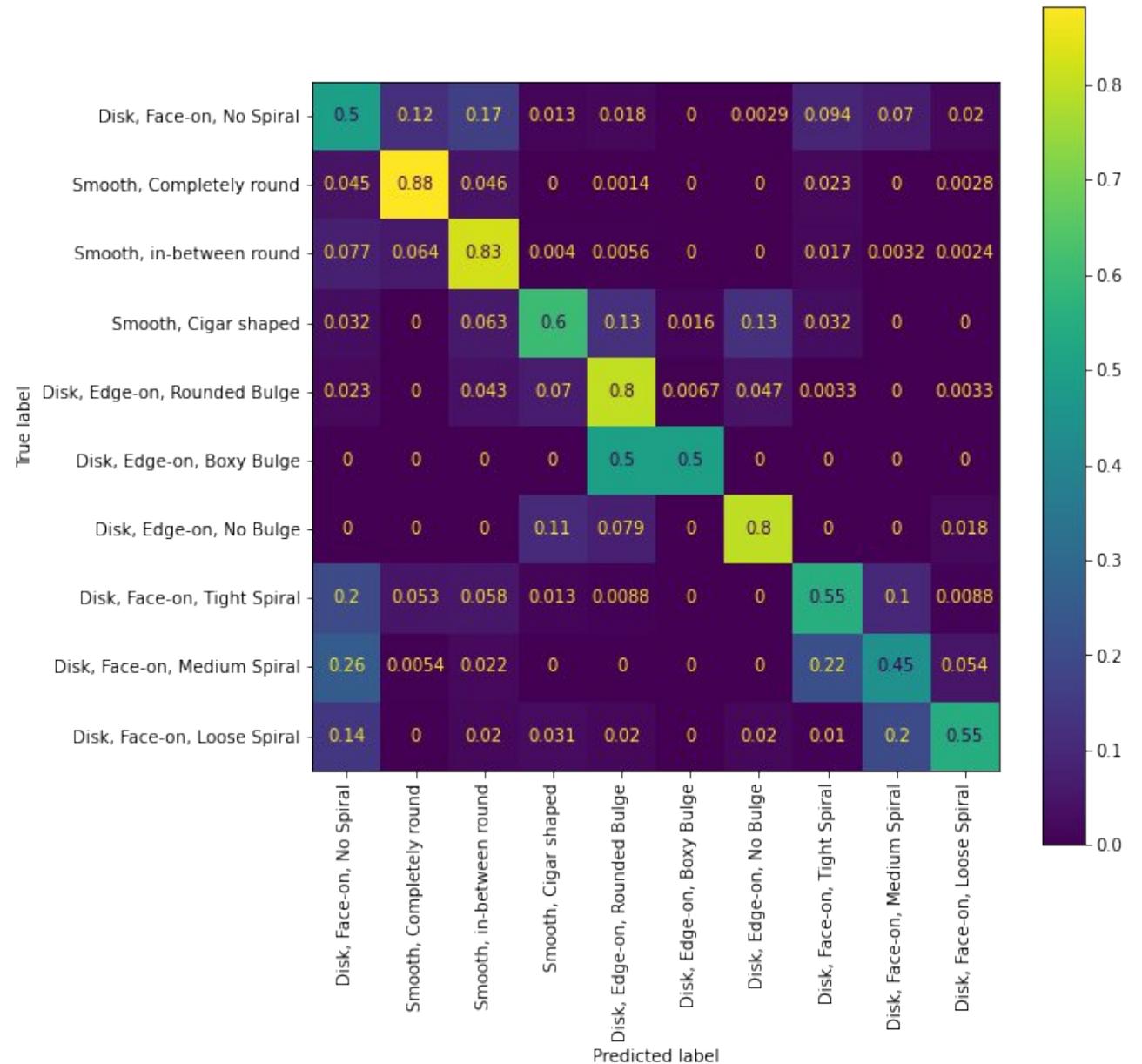
Disk, Edge-on, Rounded Bulge



Neural network with more training data

- Number of training data: 55649
- Number of test data : 4357
- Static learning rate

- Accuracy : 0.750
- Loss : 0.157
- loss: 0.0438 - accuracy: 0.9182 - val_loss: 0.1571 - val_accuracy: 0.7501



Performing images

Best

vs

Worst

Smooth, Completely round



Smooth, in-between round



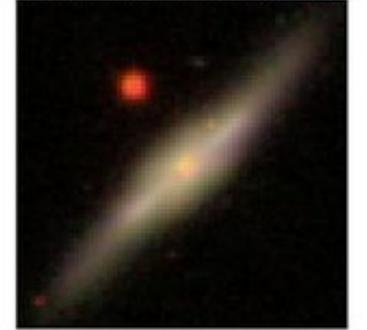
Disk, Edge-on, No Bulge



Disk, Face-on, Tight Spiral



Disk, Edge-on, Boxy Bulge



Reasons:

- Largest data set (Smooth)
- Sharp features (Edge-on, no bulge)

Reasons:

- Small data sets
- Classification of subclasses not precise for especially for Disk, Face-on classes

Appendix

Neural network with more training data

- Number of training data: 55649
- Number of test data : 4357
- Dynamic learning rate

- Accuracy : 0.725
- Loss : 0.395

- loss: 0.0137 - accuracy: 0.9818 - val_loss: 0.3953 - val_accuracy: 0.7246

