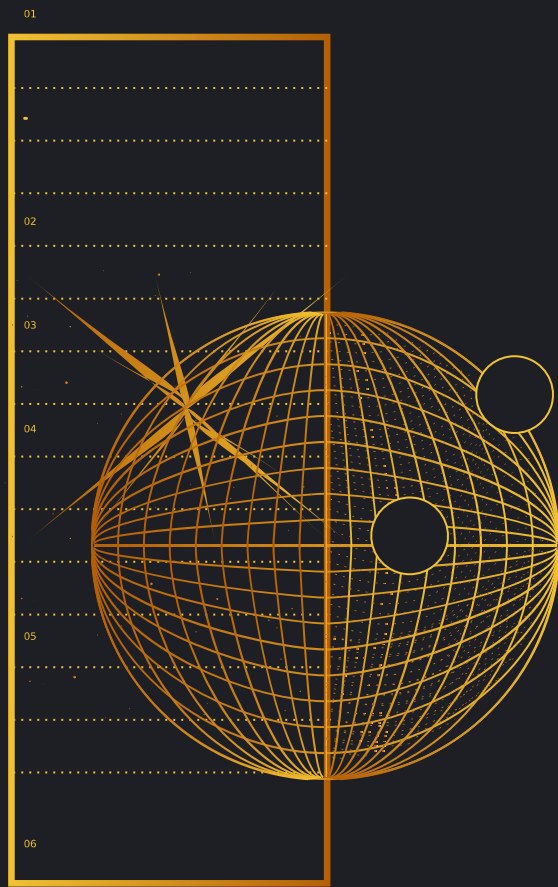




PHYS 449/773 Midterm Project Presentation GALAXY CLASSIFICATION



Group Members: Yusuf Ahmed,
Alexander Caires, Jordan Ducatel,
Ashley Ferreira, Guillaume Hewitt

01

02

03

04

05

06



TABLE OF CONTENTS

PROBLEM STATEMENT 01.

DATA 02.

ARCHITECTURES 03.

RESULTS 04.

MILESTONES & TIMELINE 05.



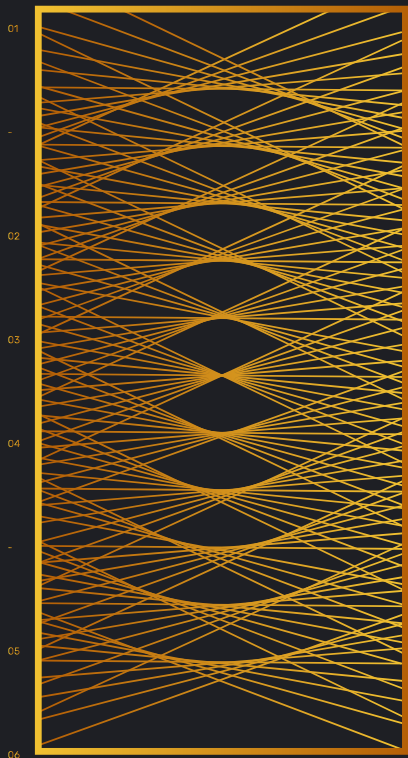
M. K. Cavanagh, K. Bekki and B. A. Groves,

“Morphological Classification of Galaxies with Deep Learning: Comparing 3-way and 4-way CNNs”

MNRAS, 506-1, 659-676 (2021).

DOI: <https://doi.org/10.1093/mnras/stab1552>





01.

PROBLEM STATEMENT



01

02

03

04

05

06

01

02

03

04

05

06

PROBLEM STATEMENT

- Classify galaxies by different types by examining their morphologies.
- In the past this was done through visual inspection.
- Worked in the past when a smaller set of galaxies were being observed.
- Today larger sets of galaxies must be observed so this method is not sufficient.
- Solution is to use machine learning methods to classify these galaxies,

01

02

03

04

05

06

01

02

03

04

05

06

Image Classification

- Galaxies are classified through deep learning, specifically image recognition and computer vision.
- Computer vision focuses on interpreting the visual world through machine learning techniques.
- It recognizes patterns in digital images and provides meaningful information from these patterns.
- Image recognition is a subset of computer vision.
- Image recognition will process the images and interpret them.
- Image recognition and computer vision is done through neural networks and for this project specifically Convolutional neural network(CNN) is used.

01

02

03

04

05

06

01

02

03

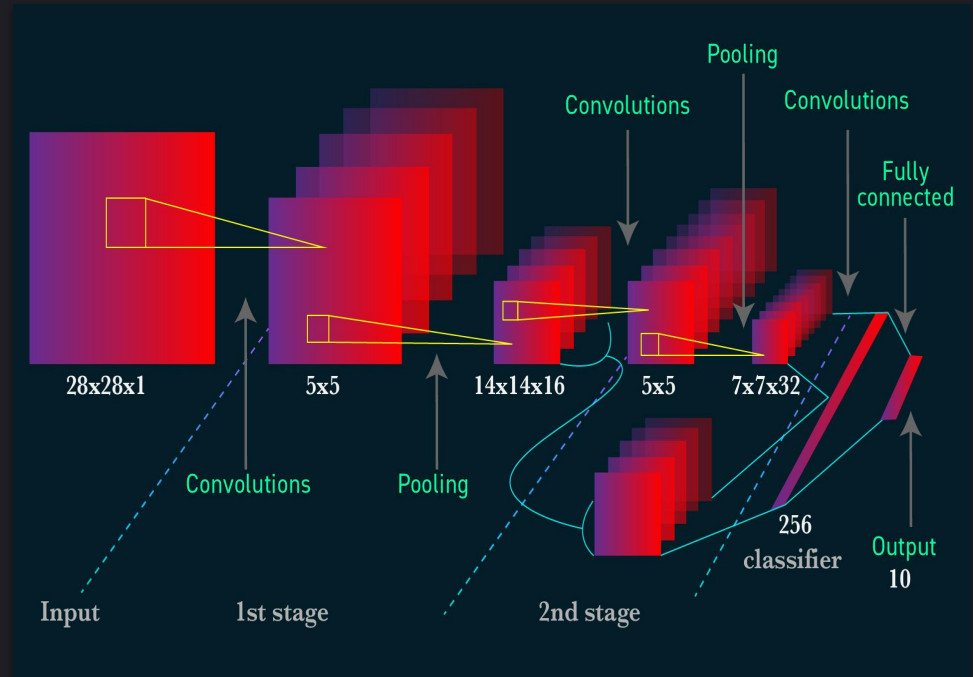
04

05

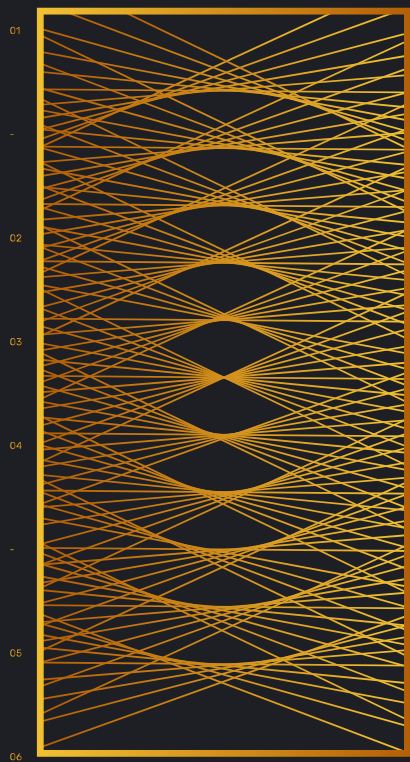
06

Why Convolutional Neural Network?

- Images have a high dimensionality.
- Each pixel in an image is considered a feature which is what contributes to high-dimensionality.
- A feature is a variable in the data set.
- The advantage of using a CNN is it does dimensionality reduction which makes training, testing, and validating data very quick.



Maladkar, K. (2018, January 25). *Overview of Convolutional Neural Network in Image Classification*. Analytics India Magazine.
<https://analyticsindiamag.com/convolutional-neural-network-image-classification-overview/>



02.

DATA



01

02

03

04

05

06

01

02

03

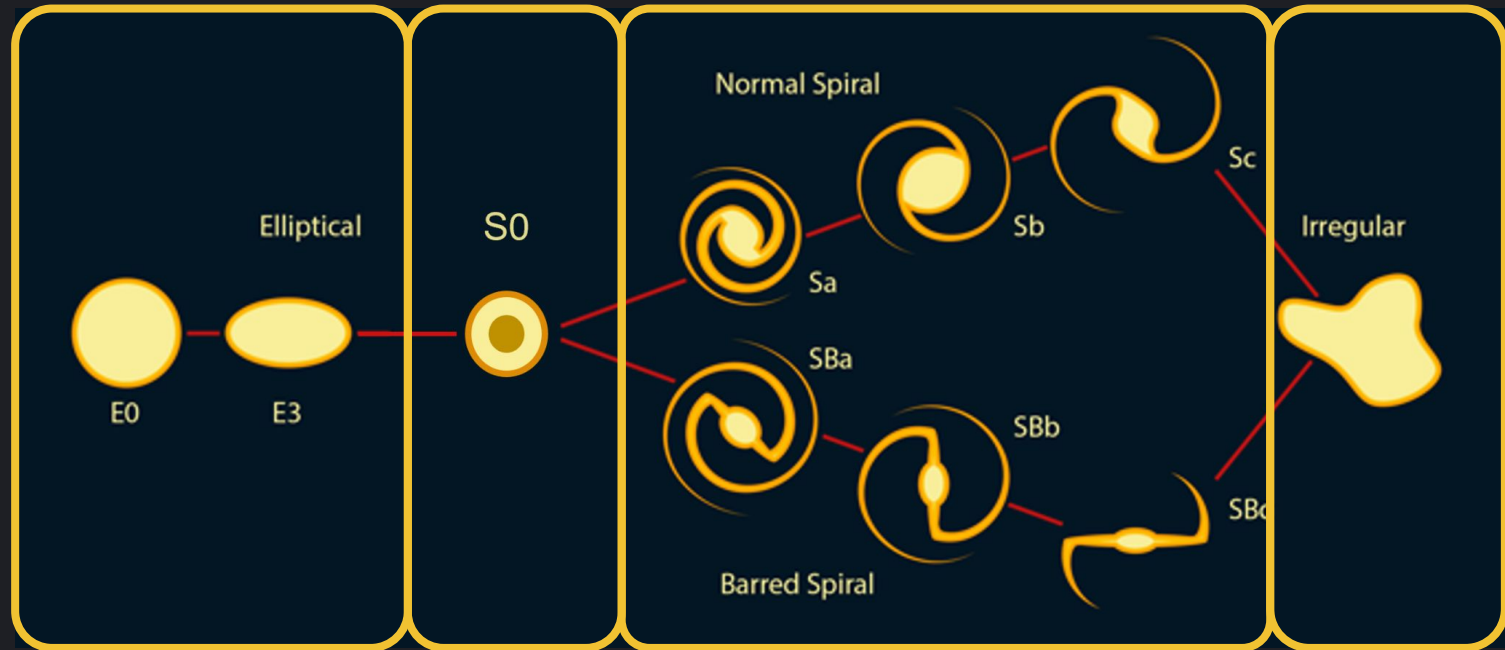
04

05

06

From visual CATALOGUE (2010) to CLASSES:

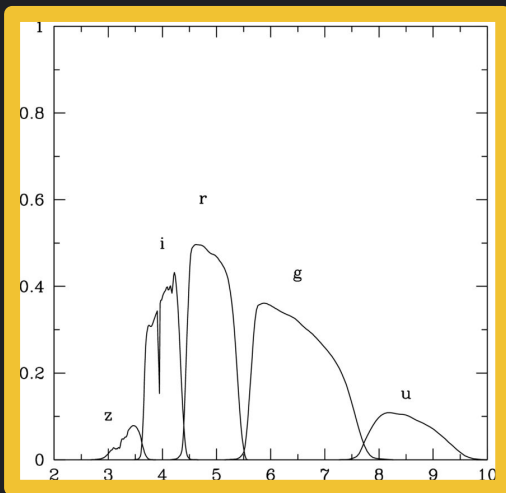
Nair *et al.* (2010) ApJS 186 427 DOI: 0.1088/0067-0049/186/2/427



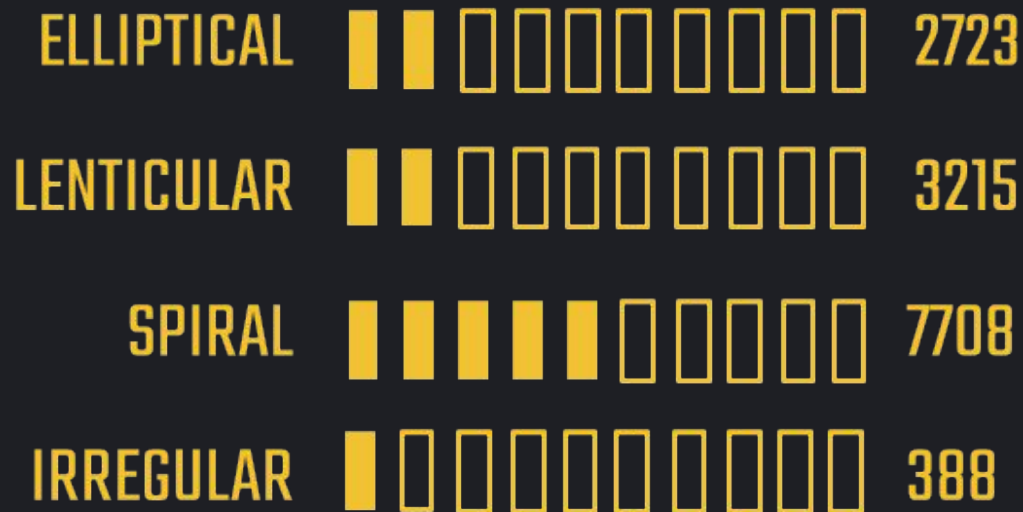


Sloan Digital Sky Survey (SDSS):

- 14,034 g-band Galaxy images



CLASSES



DATA AUGMENTATION



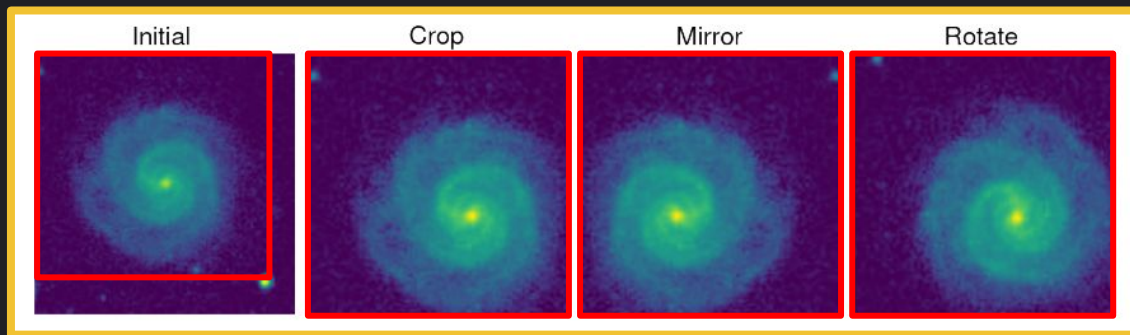
Crop image by each
corner and center.
X5 more images



Rotate image by 90 deg
X4 more images



Flip/mirror image on
y=x axis
X2 more image



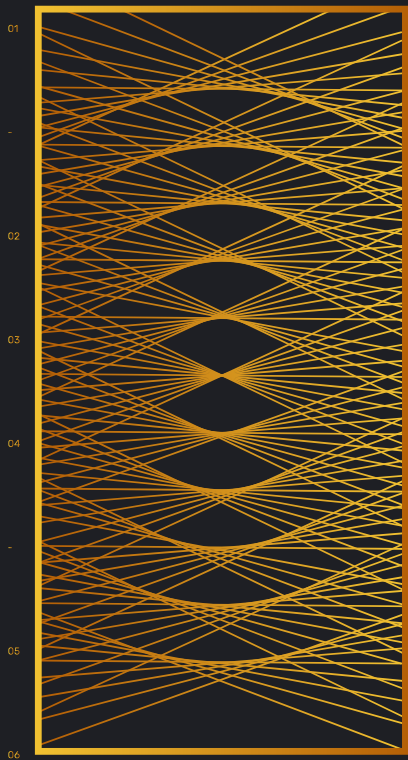
+ all images normalized



14,034 GALAXIES

561,360 Images

Classified into 3-way and 4-way classes



03.

Model Architectures

01

02

03

04

05

06



01

02

03

04

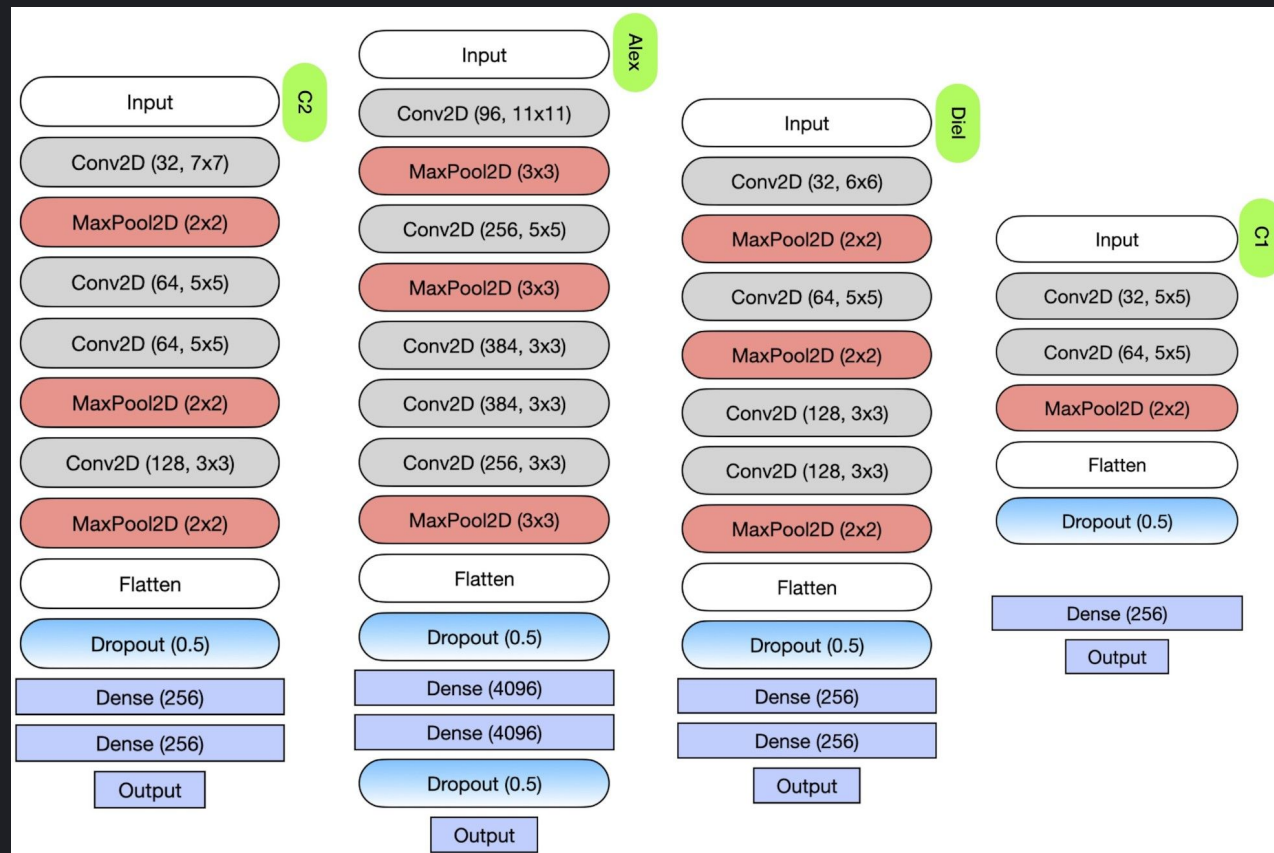
05

06

Their network, other networks, hyperparameters

C2 Model

- Batch normalisation
- ADAM
- Keras Tuner
- ReLu
- Softmax



C2

Input

Conv2D (32, 7x7)

MaxPool2D (2x2)

Conv2D (64, 5x5)

Conv2D (64, 5x5)

MaxPool2D (2x2)

Conv2D (128, 3x3)

MaxPool2D (2x2)

Flatten

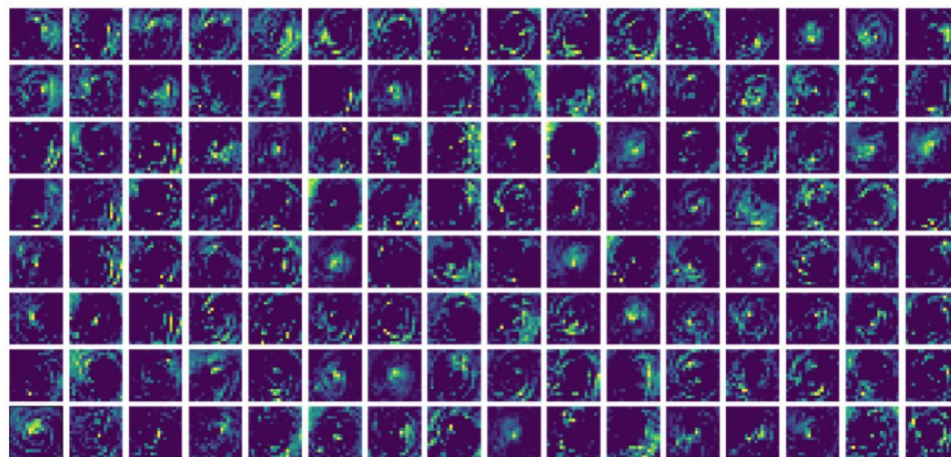
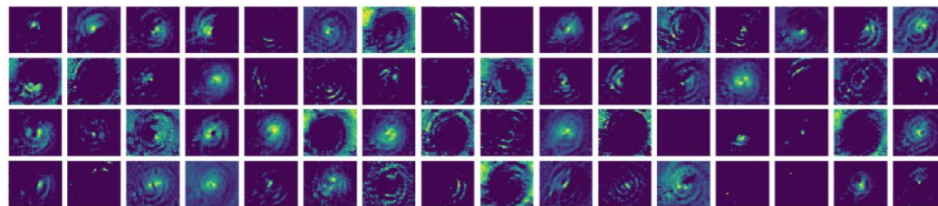
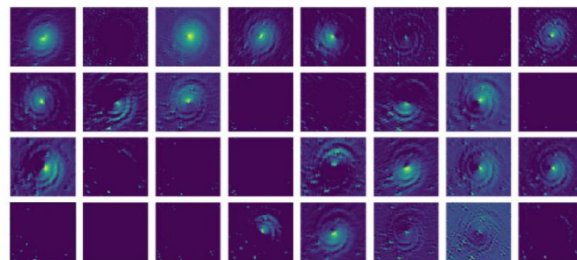
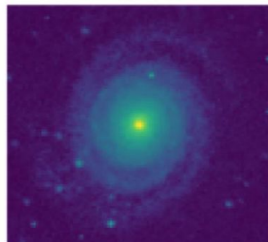
Dropout (0.5)

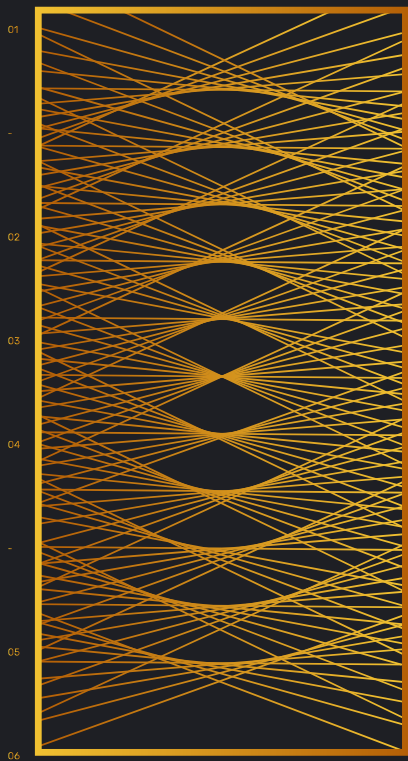
Dense (256)

Dense (256)

Output

950-52378-334





04.

RESULTS

01

02

03

04

05

06



01

02

03

04

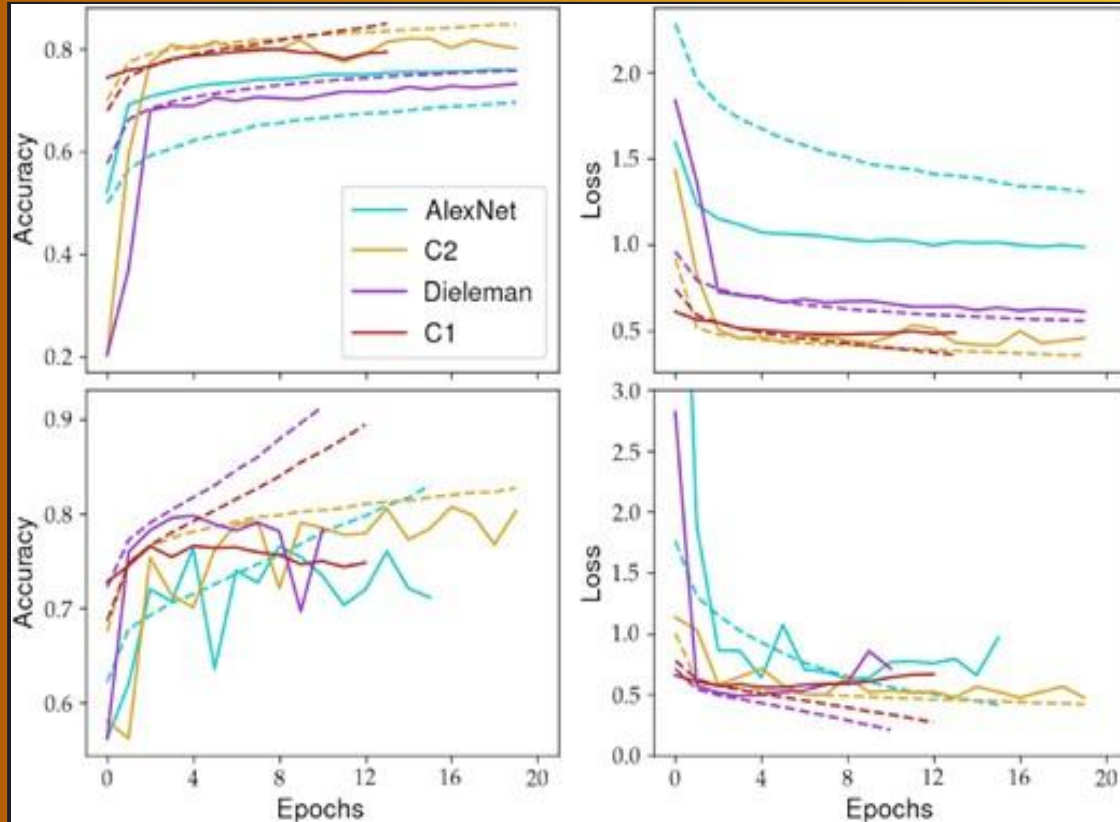
05

06

RESULTS

3-way

4-way



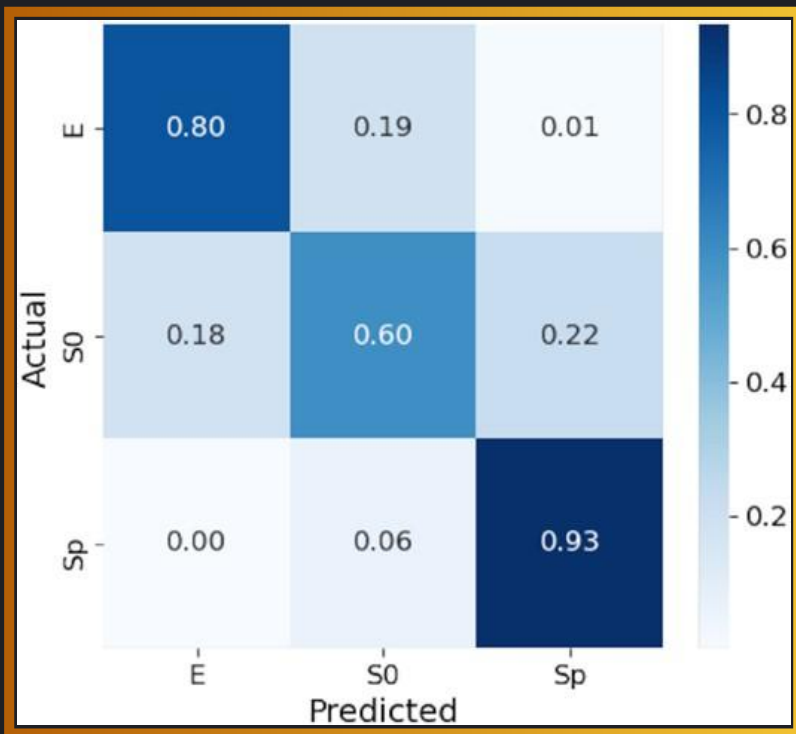
dashed = training
solid = validation

- Early stopping used to discourage overfitting and not waste compute resources
- Their network (C2) performs best on both loss and accuracy

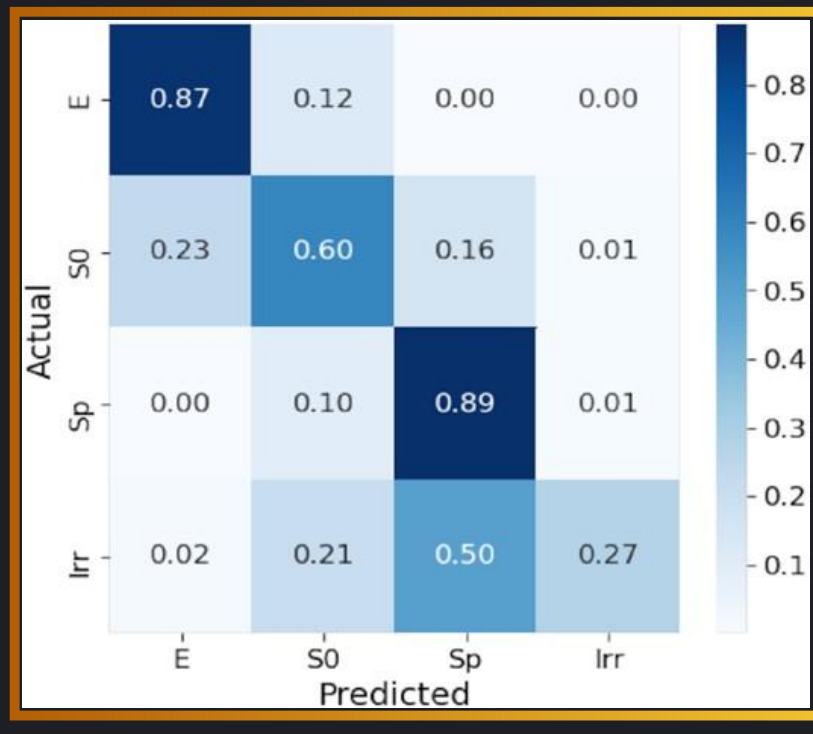
CONFUSION MATRICES

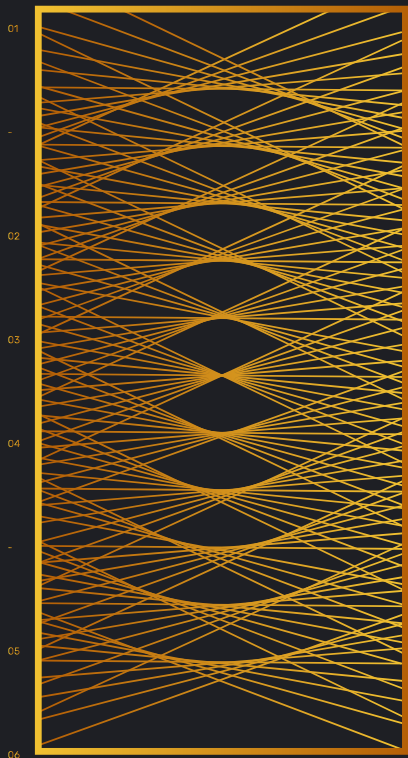
for C2, on test set

3-way classification



4-way classification





05. MILESTONES & TIMELINE



01

02

03

04

05

06

01

02

03

04

05

06

PREVIOUSLY COMPLETED MILESTONES



Group and Topic Discovery



Proposal Approval



Data Collection and Augmentation



CNN Skeleton Creation



ALEX



YUSUF



ASHLEY



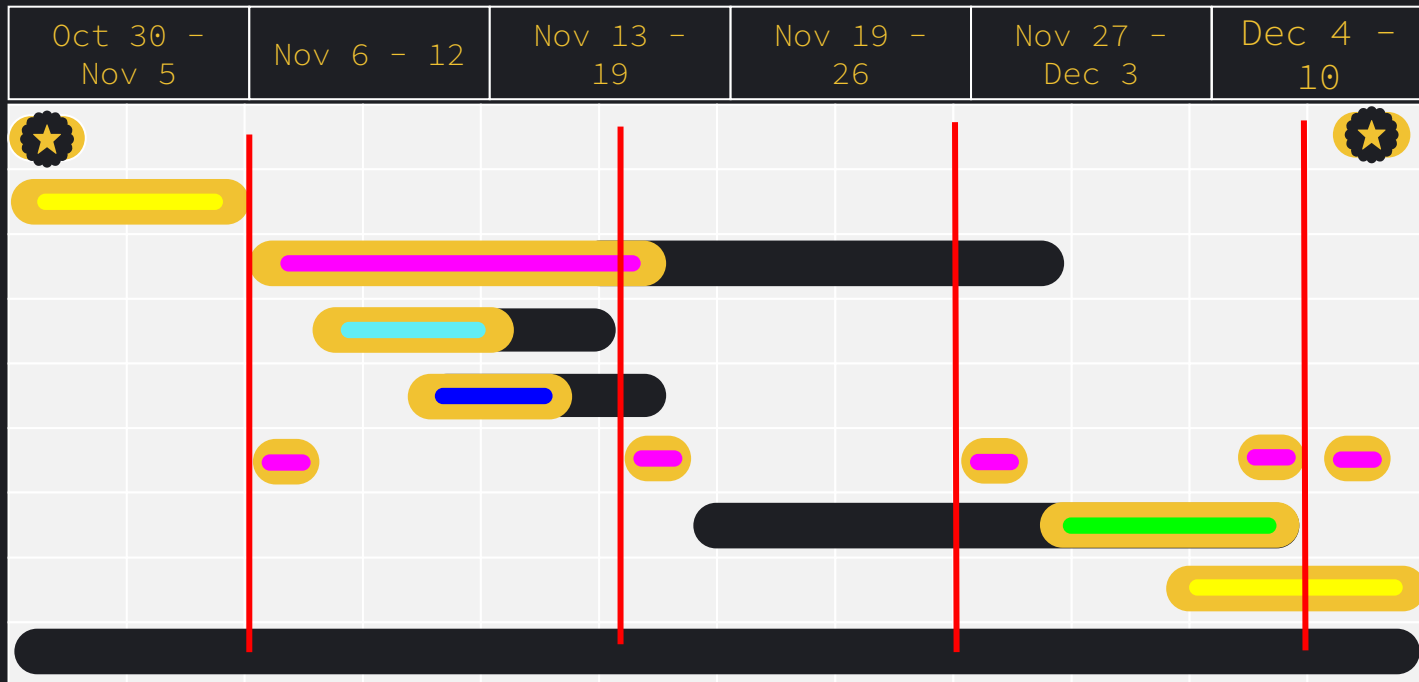
GUILLAUME



JORDAN



TIMELINE



COLLABORATION TOOLS



DISCORD

Meetings &
Chats



GITHUB

Version Control &
Code Development



DRIVE

Document
Sharing



COLAB

Shared Notebooks
& Extra Compute



WEIGHTS & BIASES

Experiment
Tracking

Cloud GPU access (paid account
with ~ NVIDIA V100 GPUs)
+ Alex local GPU as backup

The background features a dark navy blue color. In the top-left and bottom-left corners, there are yellow crop marks consisting of a circle with a crosshair. In the top-right and bottom-right corners, there are similar yellow crop marks. A large, flowing, wavy line made of many thin, parallel yellow lines starts from the top center and extends towards the right side of the frame. Another similar wavy line is located in the bottom-left corner, extending towards the center.

QUESTIONS & SUGGESTIONS?

Thank you for your time and feedback

GitHub Repo



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

