# CPSC 393 Final Project

Sreya Vadlamudi, Sarah Fieck

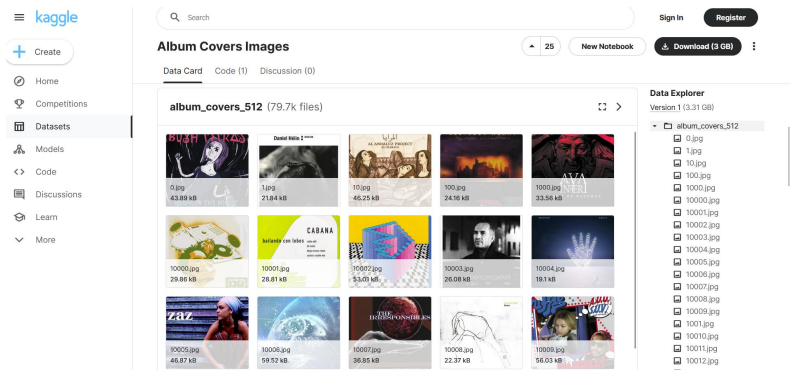
# Background

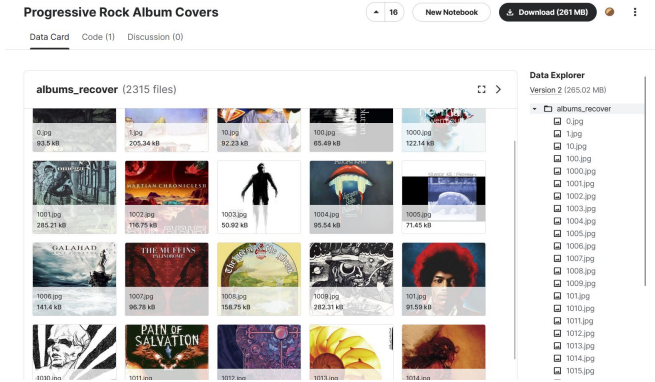
- Progressive Rock (aka Prog Rock) is a subdivision of rock music known for unique musical influences, longer song forms, and conceptual lyrics
- Has unique album cover art, abstract with elements of science fiction, fantasy, and irony

**Through this project, we wanted to see how machine learning can detect features of prog album covers in comparison to other album covers.**

# Data Sets

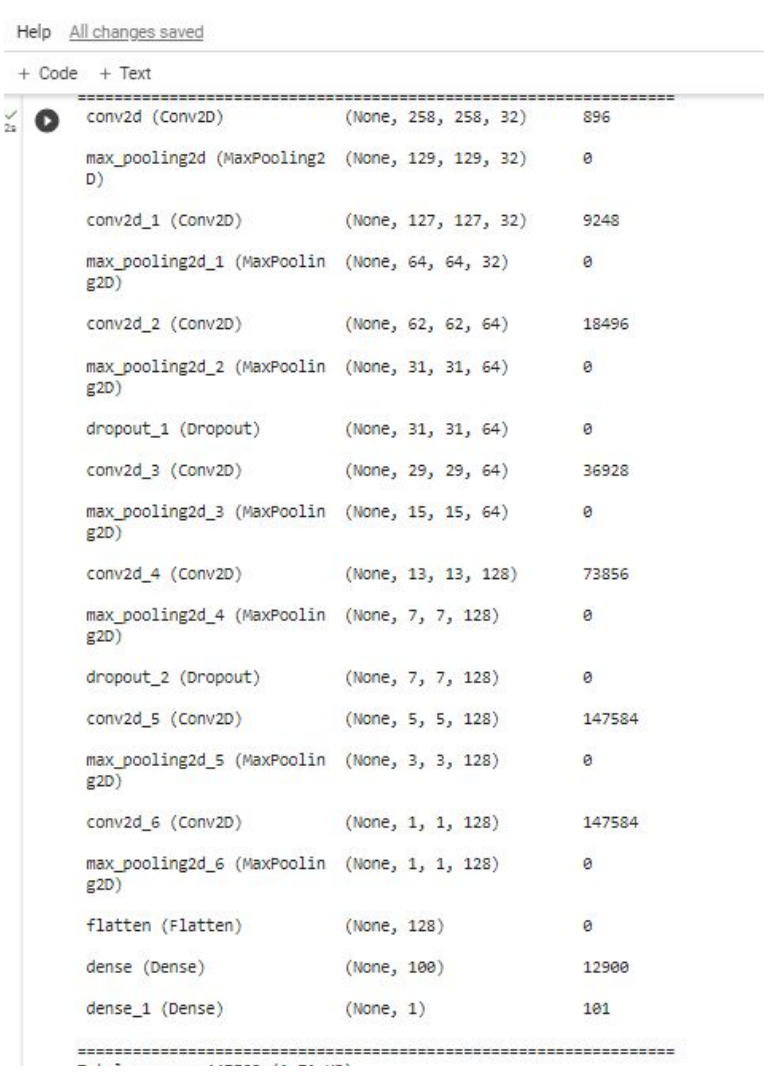
- Two datasets on Kaggle: One genreless, one prog
- Genreless: 80k of 512 x 512 album covers
- Prog Rock: 2,316 of 260 x 260 album covers
- Did some data cleaning to bring these two datasets together, training & testing file
- Scale the 512 x 512 to 260 x 260



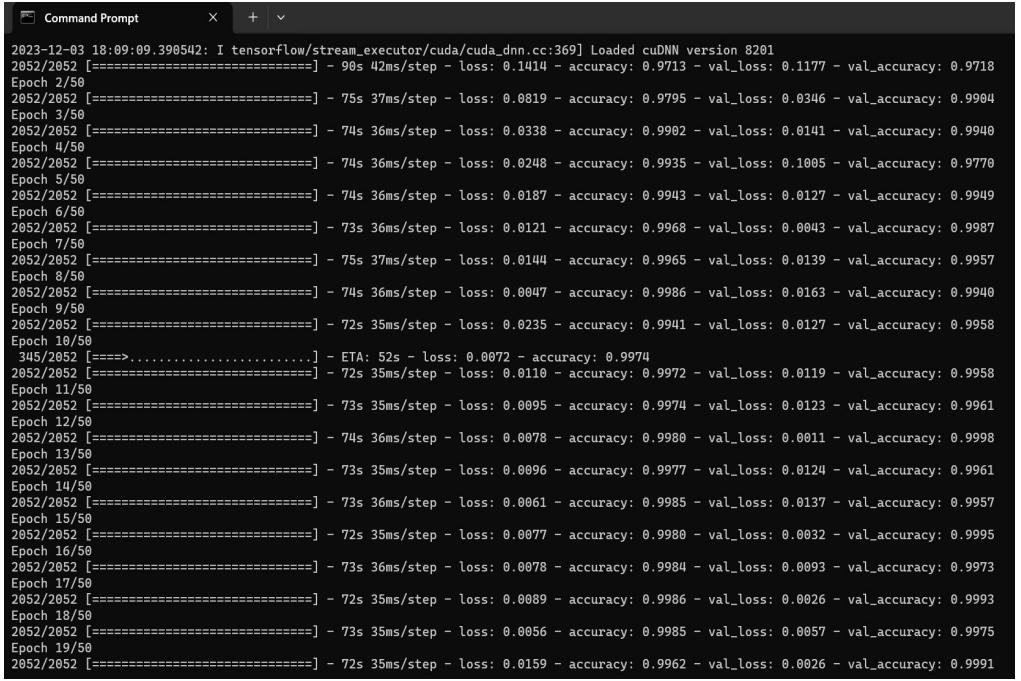
# Model & Architecture

- Categorical CNN used to detect features of progressive rock albums when found in the midst of other album covers of different genres
- Model contained seven MaxPooling2D layers for regularization to combat overfitting
- Dropout layers and Batch Normalization were also used for regularization as well
- We did run into an issue where our gradients were all 0 when trying to run our heat map. We fixed this by changing the architecture of our model and added more regularization
- Metrics:
    - Loss: binary_crossentropy
    - Accuracy was measured

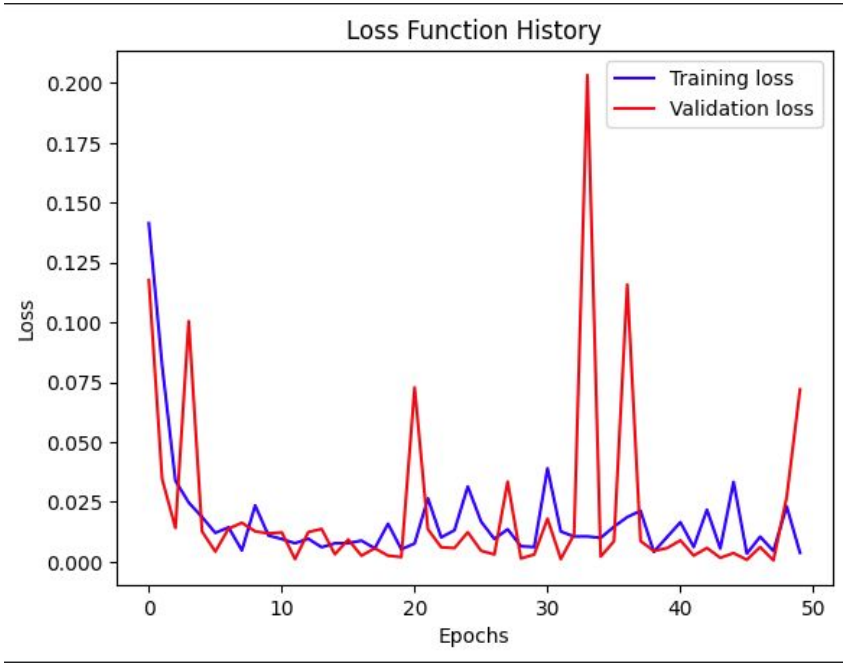


# Results

- Ran the model for fifty epochs on the server
- Overall, the model ran smoothly and was not overfit in regards to the loss values and accuracy.
- The model stayed around 99% for accuracy for most of the epochs and the loss .02 by the end.



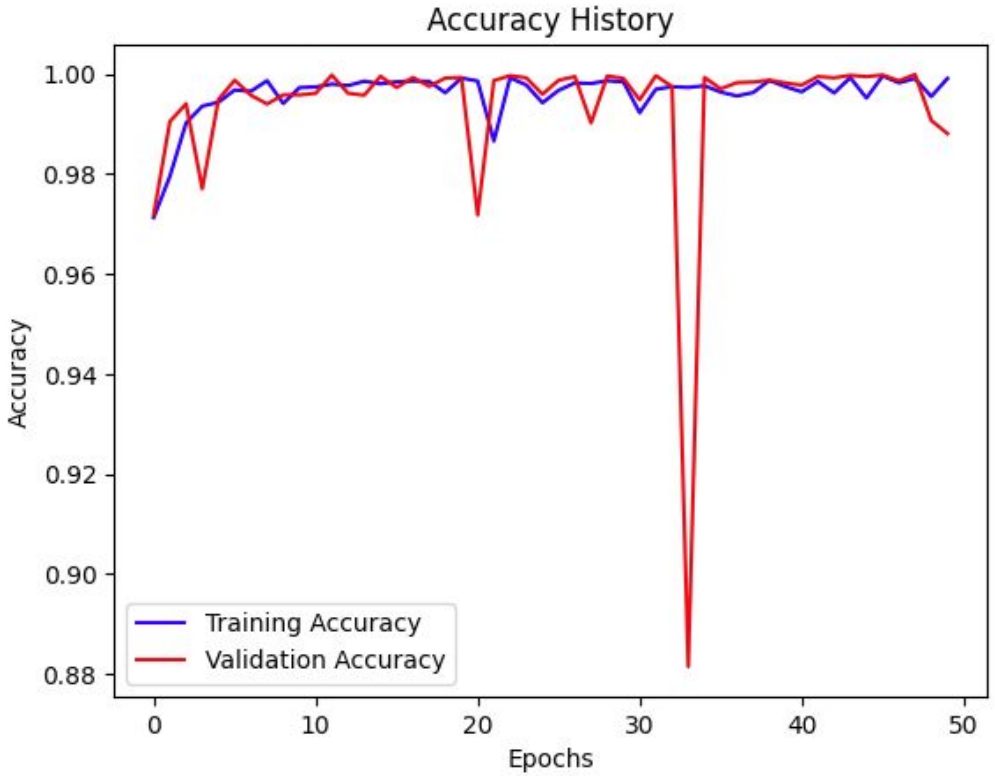
# Loss Function Metrics




- Used binary cross-entropy
- Training loss was fairly low, sticking around 0.02 towards later epochs
- Validation loss mostly mimicked this trend, with some abnormalities
  - The spikes?
  - Does not fully exhibit overfitting

# Accuracy Metrics

- Training accuracy is highly accurate, above 0.99
- Validation accuracy hugs most of training accuracy, with some dips



# Heatmap & Layer Activation

- We created a heatmap in order to see which features of the album covers were most important when comparing prog rock to other albums.
- We compared an image from the prog rock album to an image in the other album.
    - These are the two images(prog rock being on the left):


**Image 1:** Album is Images and Words by a progressive metal band, Dream Theater.

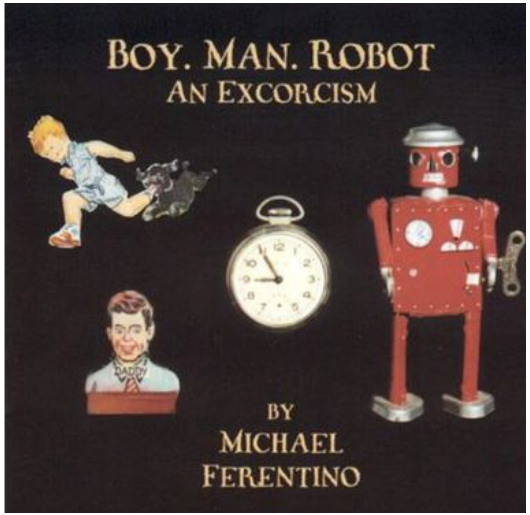
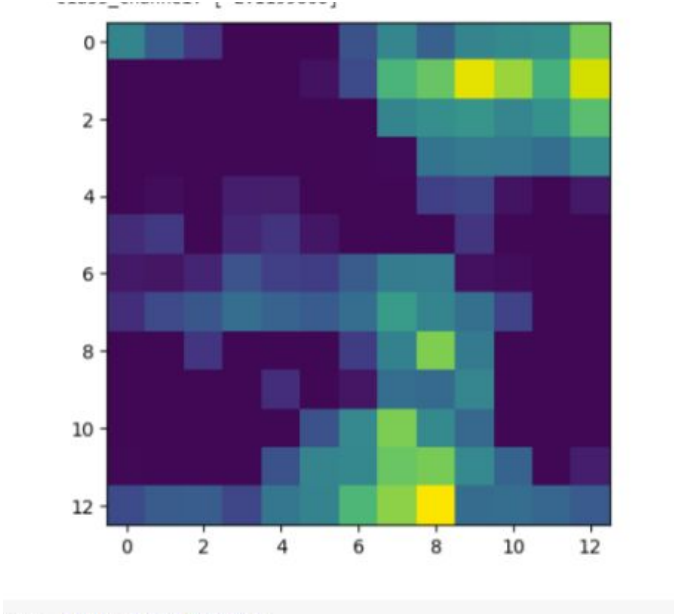


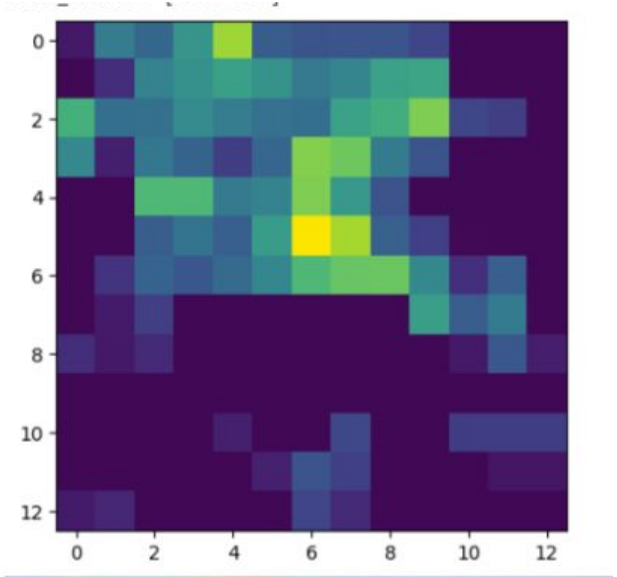

**Image 2:** Album is Boy. Man. Robot. An Exorcism by Michael Ferentino within the Genreless album dataset.

# Heatmaps for the Two Images

Prog Rock Heatmap

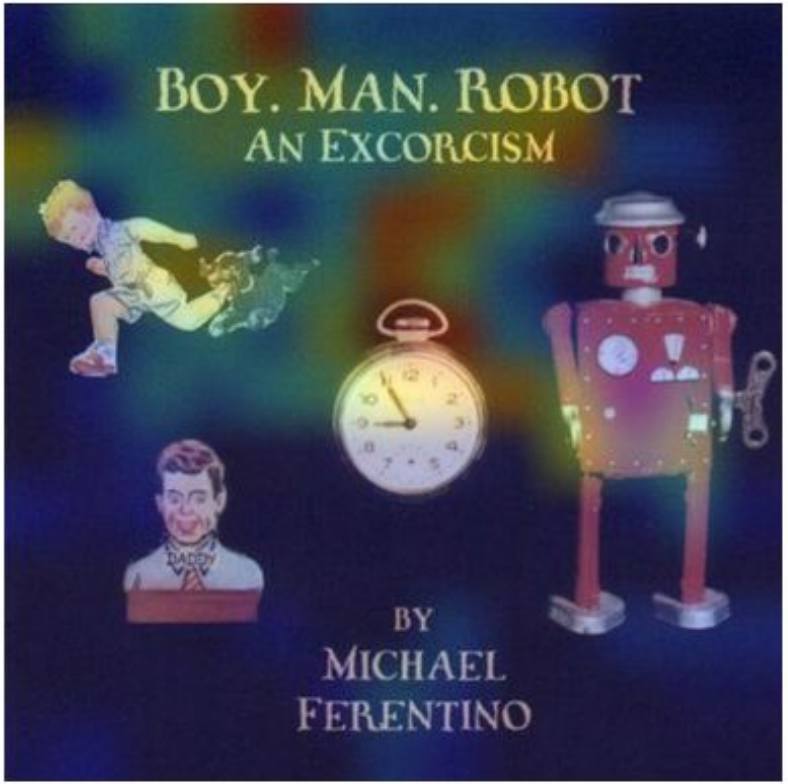



Genreless Heatmap

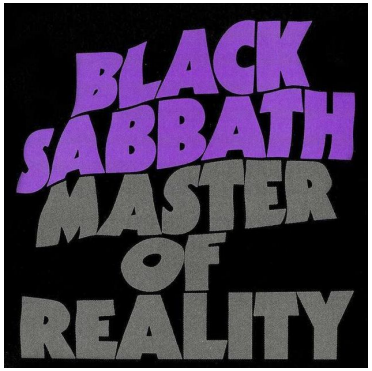


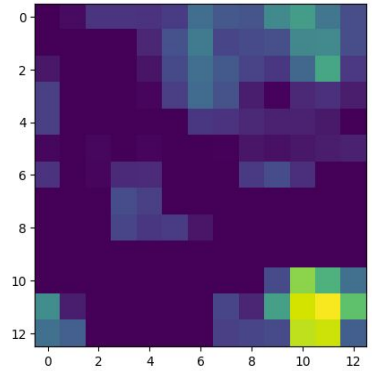
# GradCAM (superimposed visualization) - Image 1

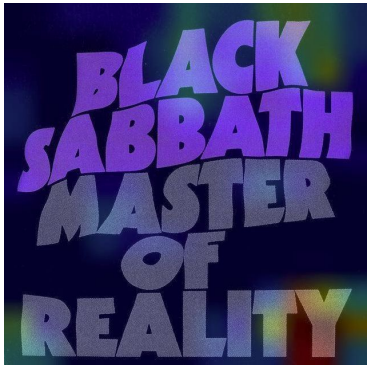
# GradCAM (superimposed visualization) - Image 2

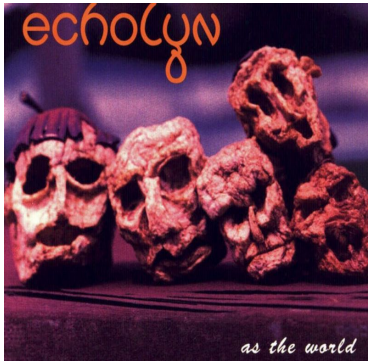


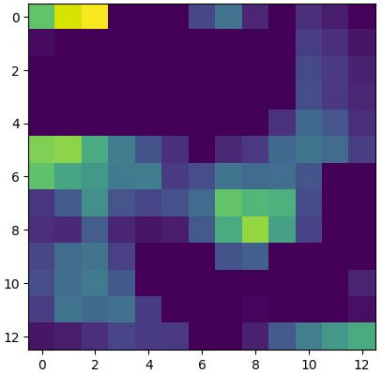
# Extra Prog

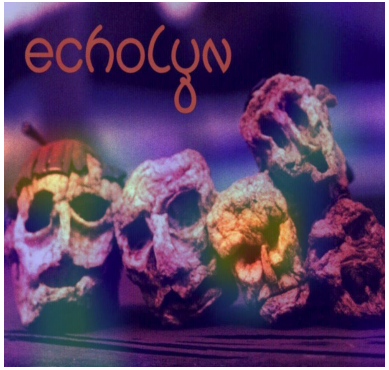

BLACK
SABBATH
MASTER
OF
REALITY


0
2
4
6
8
10
12
0
2
4
6
8
10
12


BLACK
SABBATH
MASTER
OF
REALITY


echolyn
as the world


0
2
4
6
8
10
12
0
2
4
6
8
10
12


echolyn

# Findings & Conclusions

- Metrics indicate we created an accurate CNN classification model that does not appear to be overfit
- Heat maps & GradCAM visualize activated aspects of the images to classify albums
- Prog rock albums identified by abstract patterns and symbols
    - Weird right corner?
- Other genre albums identify faces and words
- Questions of hidden overfitting from scaling?

# Why Deep Learning?

- Deep Learning is necessary since image classification with CNNs  is not something that can be done with simple models.
- Simple models cannot take in so much data with specific features
- Multiple layers and training required for this classification
- Depth of model is important in understanding the accuracy in detecting layered/deep architectures.
- Grad-CAM/heatmaps to detect most important features for specific details/features also use deep learning.

# Questions?