

Deep Learning for Emotion Recognition in Cartoons

Under the guidance of
Ms. Preeti Sehrawat
Assistant Professor

Shubham Kumar (00815007721)
Abhishek Sharma (00115007721)
Arjun Sharma (00415007721)

Aim

**MEASURE HOW ACCURATE THE PROGRAM IS ABLE TO
IDENTIFY AN EMOTION FROM A GIVEN CARTOON VIDEO.**

Several thin, white, parallel diagonal lines are positioned in the bottom right corner of the slide, extending from the right edge towards the center.

REQUIREMENTS



- Choose a cartoon: Choice was *Tom & Jerry*².
- Lots of various emotions in each episode.
- Segment faces from the cartoon.
- Build a dataset of emotions for each main character (Tom & Jerry)
- Train the network on a labeled dataset.

The logo for the cartoon "Tom and Jerry" is displayed in a bold, red, stylized font. The word "TOM" is on the left, "and" is in a smaller, cursive font in the middle, and "JERRY" is on the right. The entire logo is set against a blue background with white diagonal lines on the right side.

² *Tom & Jerry* © Warner Bros. Entertainment, Inc

DATASET GATHERING

Tom & Jerry



About 4,260,000 results

FILTER



Tom and Jerry, 88 Episode - Pet Peeve (1954)

Jonni Valentayn 1.4M views • 1 week ago

The short three-minute fragment from series is the 88th one-reel animated **Tom and Jerry** short, directed by William Hanna and ...



Tom and Jerry, 89 Episode - Touché, Pussy Cat! (1954)

Jonni Valentayn 387K views • 5 days ago

The short three-minute fragment from series is a 1954 one-reel animated **Tom and Jerry** short, directed by William Hanna and ...

New



Tom and Jerry, 87 Episode - Downhearted Duckling (1954)

Jonni Valentayn 1.6M views • 3 weeks ago

The short three-minute fragment from series is the 87th one reel animated **Tom and Jerry** cartoon released. It was created in 1953, ...



Jonni Valentayn

829,643 subscribers • 99 videos

Tom and Jerry - is an American animated series of short films crea...

SUBSCRIBE 829K



Tom and Jerry, 88 Episode - Pet Peeve (1954)

1.4M views • 1 week ago



Tom and Jerry, 87 Episode - Downhearted Duckling

1.6M views • 3 weeks ago



Tom and Jerry, 86 Episode - Neapolitan Mouse (1954)

2.1M views • 1 month ago



Tom and Jerry, 85 Episode - Mice Follies (1954)

1M views • 1 month ago



Tom and Jerry, 84 Episode - The Boating Party (1954)

2.1M views • 1 month ago



Tom and Jerry, 13 Episode - The Zoot Cat (1944)

168K views • 1 month ago



Tom and Jerry, 11 Episode - The Yankee Doodle Mouse

296K views • 1 month ago



Tom and Jerry, 10 Episode - The Lonesome Mouse (1943)

270K views • 1 month ago



Tom and Jerry, 9 Episode - Sufferin' Cats! (1943)

190K views • 1 month ago



Tom and Jerry, 8 Episode - The Cat and the Hat (1943)

587K views • 1 month ago



Tom and Jerry, 6 Episode - Puss n' Toots (1942)



Tom and Jerry, 5 Episode - Dog Trouble (1942)



Tom and Jerry, 4 Episode - Fraidy Cat (1942)



Tom and Jerry, 3 Episode - The Night Before Christmas



Tom and Jerry, 2 Episode - The Two Mouseketeers

SEGMENTATION



HAAR CASCADES



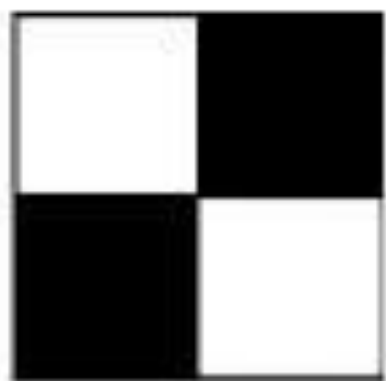
HAAR CASCADES



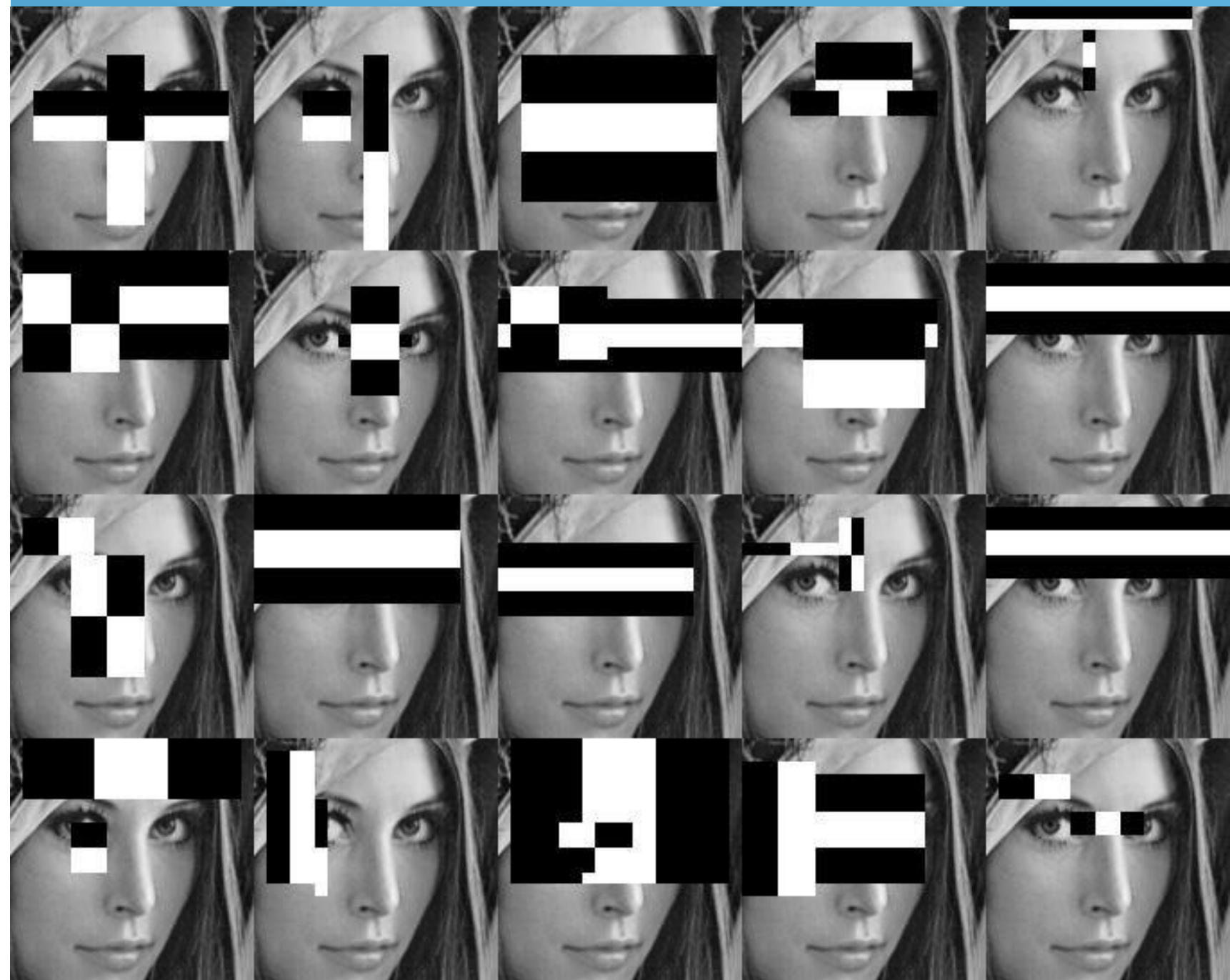
(a) Edge Features



(b) Line Features

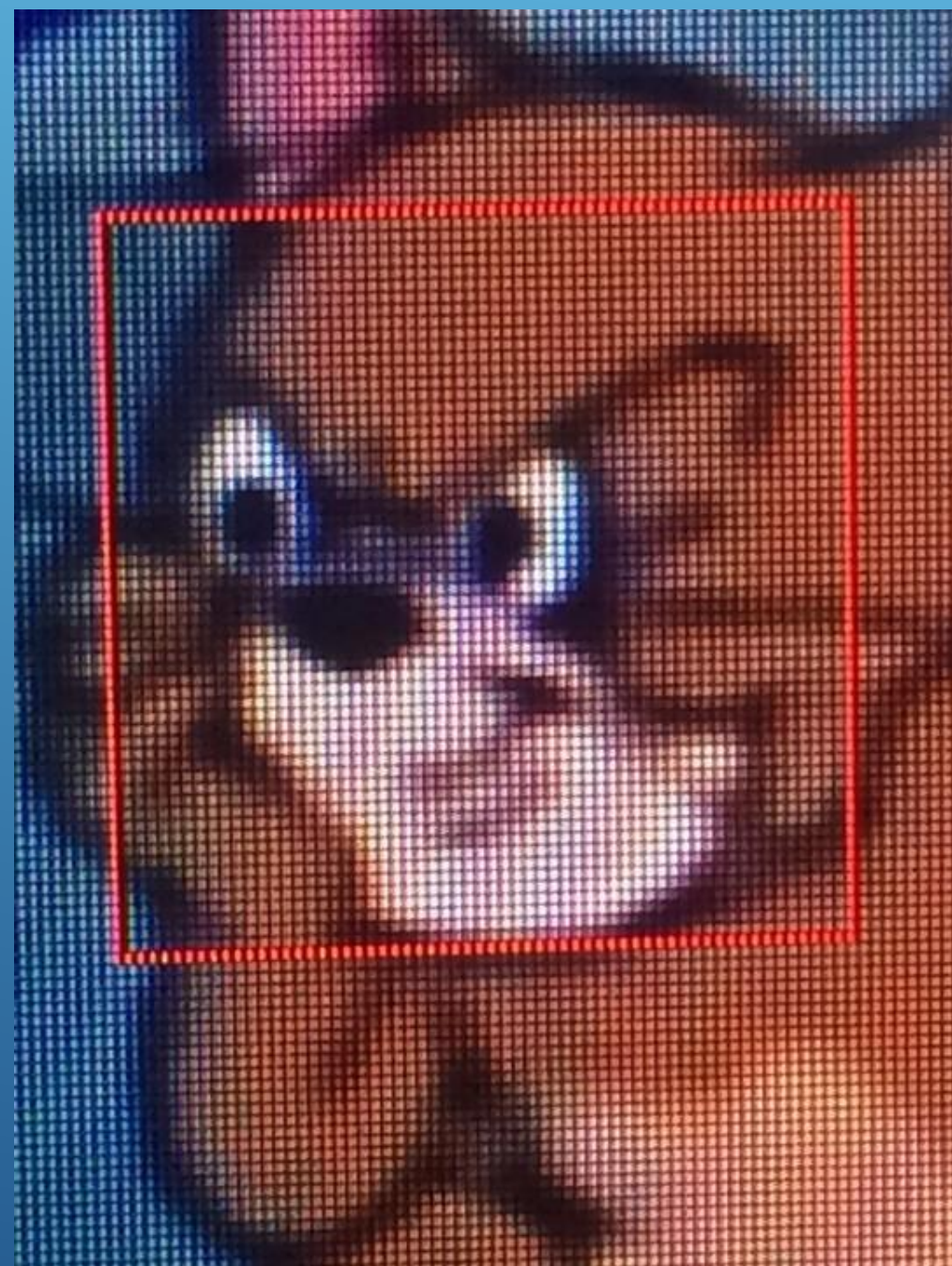
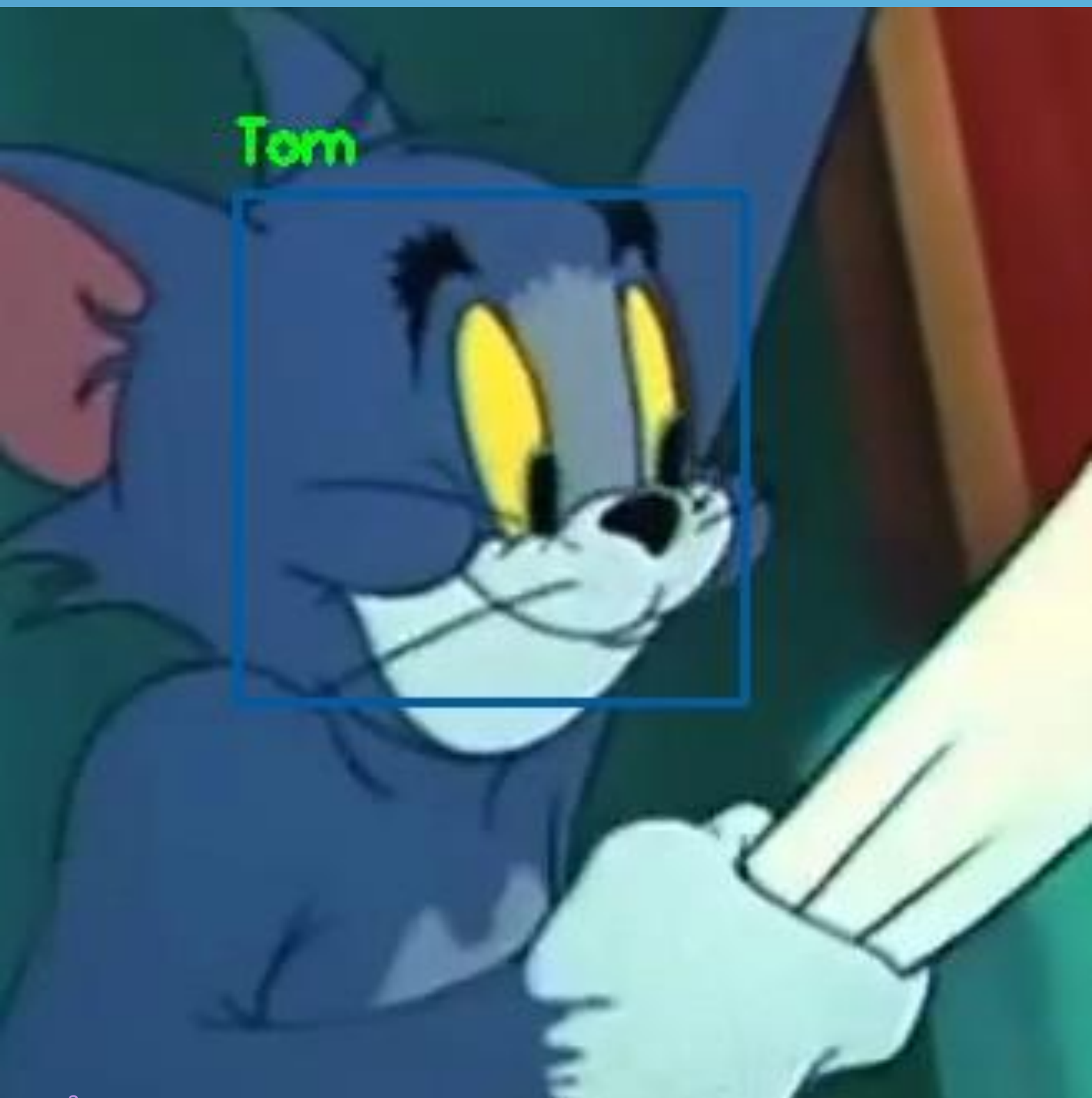


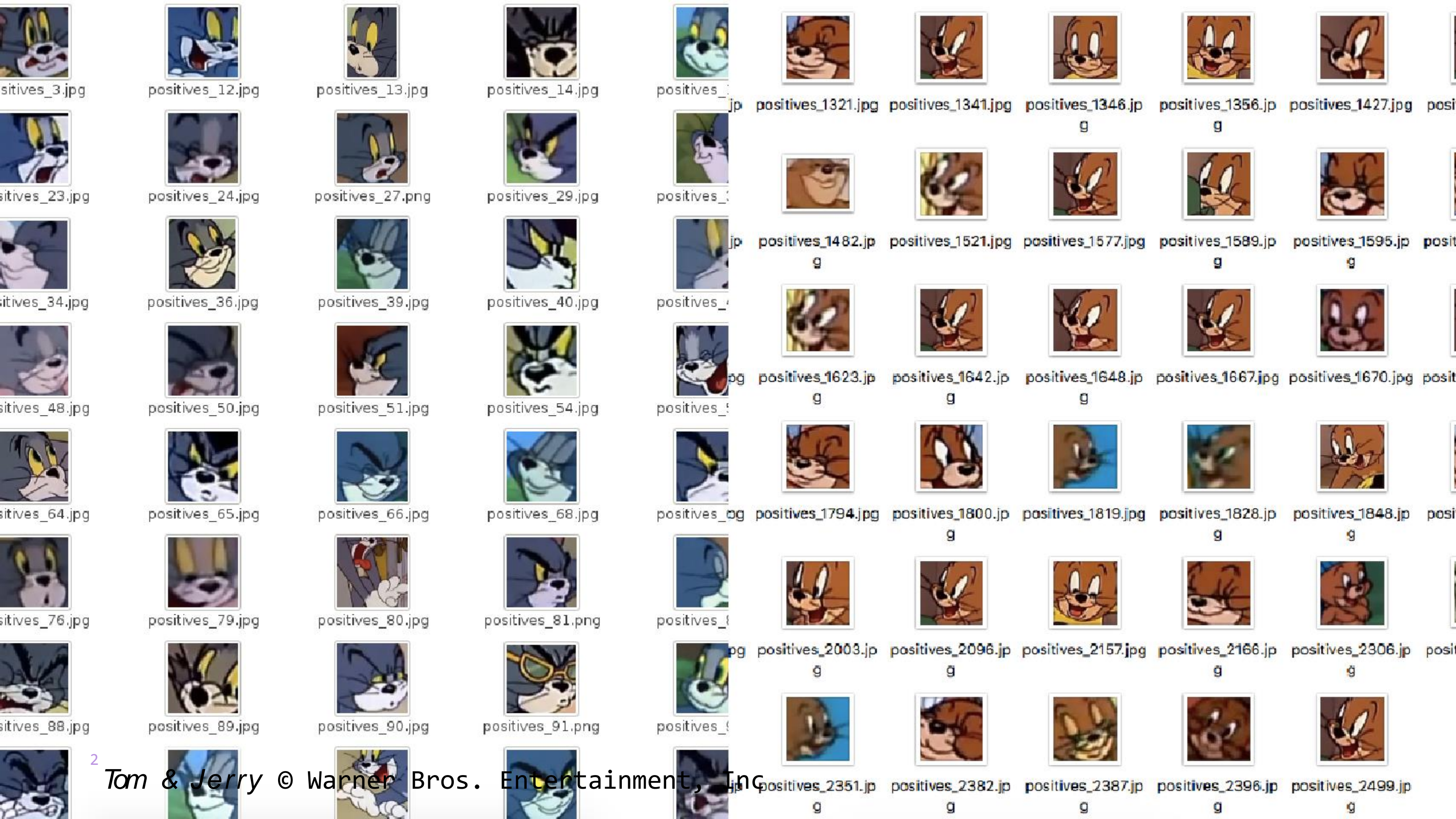
(c) Four-rectangle features



Haar Cascades

- Created custom Haar cascade for both *Tam & Jerry*.
 - There were none made online to detect cartoon faces, only human ones.
 - Depending on the window size it does detect other character faces in the cartoon.





DATASET STATS

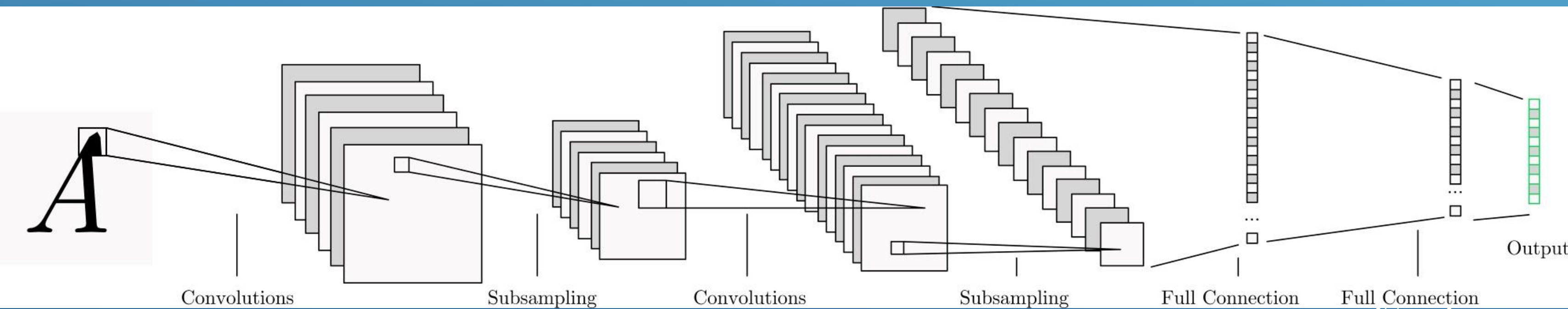


Dataset Stats

- In total about **159,035** images segmented.
- For about ~64 episodes. (*Tom & Jerry has over 100*)
- Selected around 400 images for each character & emotion, (angry, happy, surprise) for training and testing.

CONVOLUTIONAL NEURAL NETWORK

Several thin, white, parallel diagonal lines are positioned in the bottom right corner of the image, extending from the right edge towards the center.



Convolutional Nerual Network

- In recent years CNN's produced great results in image & object recognition.
- The CNN is used for this project to learn **features** (*eg. smile angles, eyebrows*).
- Framework for DL used was Keras + TensorFlow backend. (Keras also works with Theano)

Convolutional Nerual Network

- No pre-trained network.
- Inception-V3 predicts Tom & Jerry as 'comic books'.
- Images resized to 60x60 with 3 channels. (RGB)
- 3x3 convolution & 2x2 max pooling with a input image of size 60x60x3.

Convolutional Neural Network

- 3x3 Convolution & 2x2 Max Pooling. (ReLU activation)
- 3x3 Convolution & 9x9 Max Pooling. (ReLU activation)
- Fully connected layer of 512 neurons.
- Final output layer of 6 neurons for each emotion.

ALGORITHM USED



Algorithm Used

The optimization algorithms that are tested and evaluated in the artefact are the following:

- Stochastic Gradient Descent (SGD)
- Adagrad
- Adadelata
- RMSprop
- Adam

RESULTS



Results

- Split dataset into 80% training, 20% testing.
- Trained network for 50 epochs on one GPU (Nvidia).
- Tested 5 optimisers for 5 runs:
 - Adadelata, Adagrad, Adam, RMSprop & Stochastic Gradient Descent (SGD)
- Hyperparameters (Layer size, Max pooling size...)

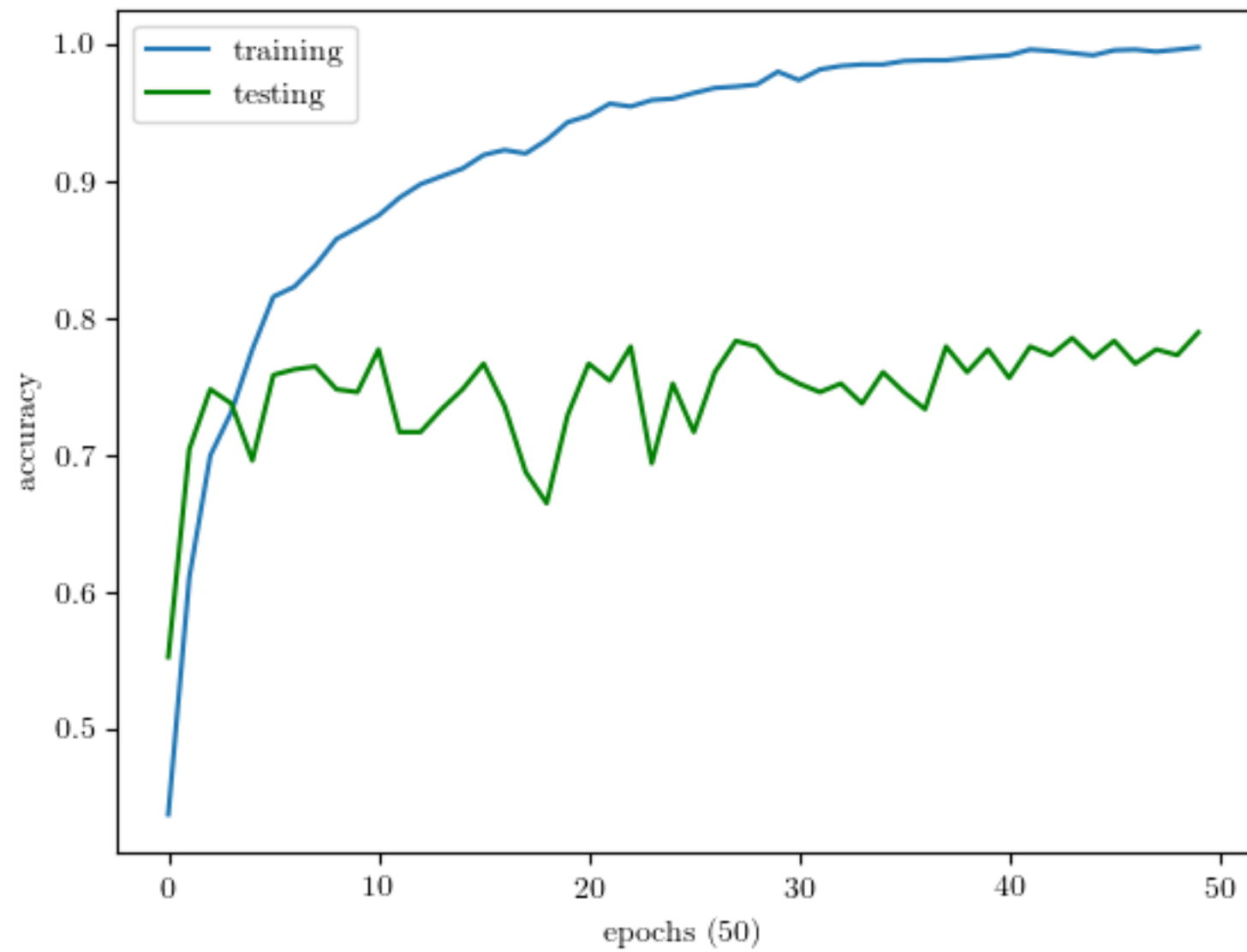
Results

- The network removes around 20-50% of neurons from the network when training. (Dropout)
- Prevents overfitting the network.
- Rmsprop overfits the network.
- Adadgrad tends to underfit the network slightly.

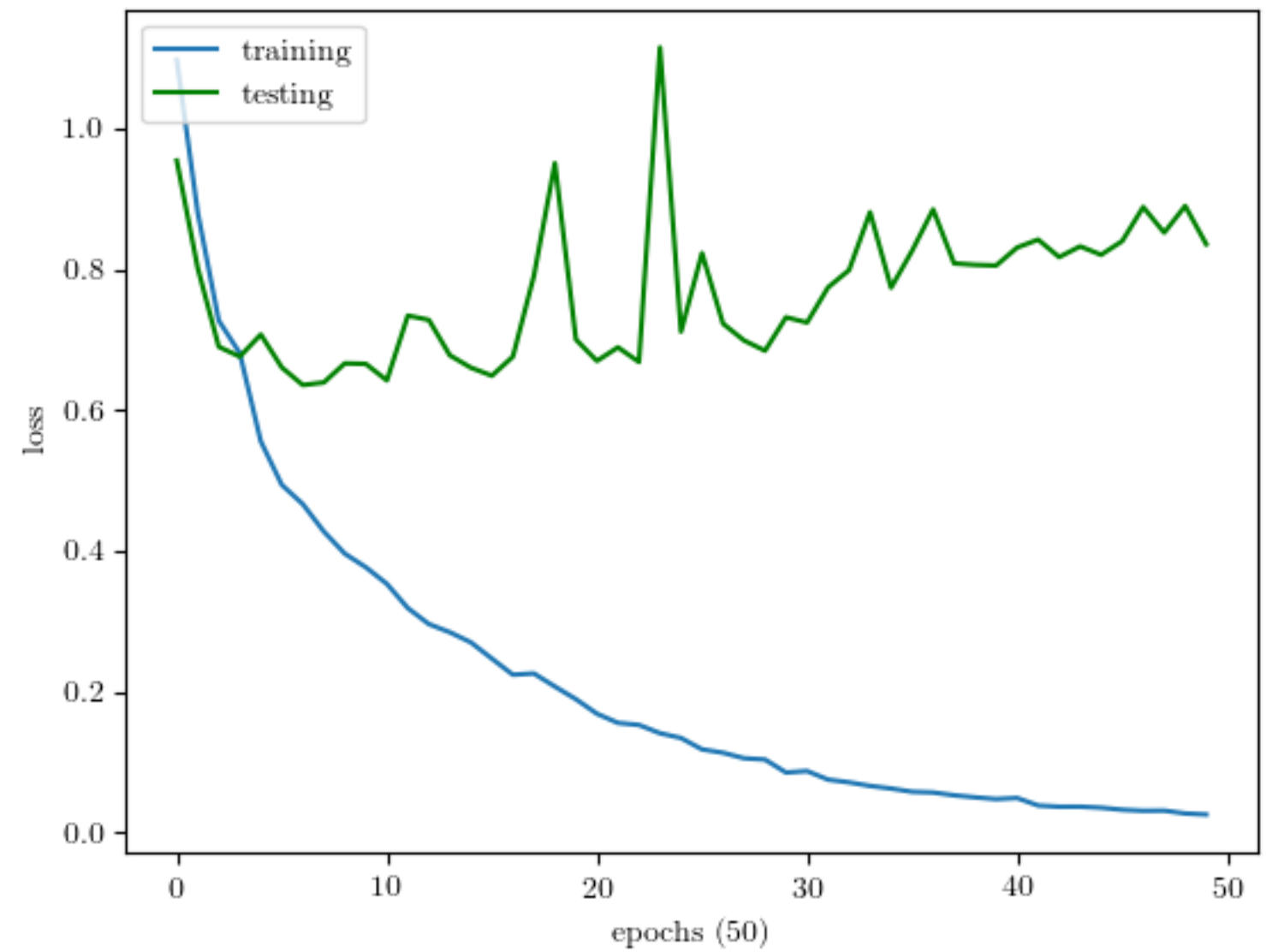
Results

- Adadelata & SGD optimiser works well with slight overfitting.
- Adam has comparable performance to SGD but underfits in some test runs.
- Adadelata was the best, but SGD was better for 3 test runs. (Both achieved ~80% accuracy)

Learning curve for model accuracy



Learning curve for model loss



Parameter	Value
1st Layer	3x3 Convolution
2nd Layer	2x2 Maxpooling
Dropout	20%
Neurons in 1st FC Layer	512
Neurons in 2nd FC Layer	6
Metric	Categorical Cross-Entropy
Epochs	50

Table 1.1: Hyperparameter Table, Run 1.

Algorithm	Model Loss	Model Accuracy
Adadelata	1.3%	0.8%
Adagrad	10.7%	0.3%
Adam	1.3%	0.8%
RMSProp	10.7%	0.3%
SGD	0.8%	0.8%

Table 1.2: Results Table, Run 1

Run	Algorithm	Model Loss	Model Accuracy
1	SGD	0.8%	0.8%
2	SGD	0.9%	0.8%
3	SGD	1.0%	0.8%
4	Adadelata	1.0%	0.7%
5	Adadelata	0.6%	0.8%





Table 1.3: Algorithms with the best loss accuracy out of all 5 runs

~80% MODEL ACCURACY


~90% MODEL LOSS

CLASSIFICATION RESULTS

		
Happy 86.8% Surprise 9.5% Angry 3.6%	Happy 97.9% Surprise 1.5% Angry 0.6%	Happy 91.0% Surprise 5.5% Angry 3.5%
		
Happy 52.7% Angry 34.0% Surprise 13.3%	Happy 55.8% Surprise 33.2% Angry 10.8%	Happy 67.7% Angry 24.7% Surprise 7.8%
		
Happy 50.5% Angry 37.1% Surprise 14.7%	Angry 48.5% Happy 36.2% Surprise 15.3%	Angry 55.7% Happy 29.6% Surprise 14.7%
Adadelta		

		
Surprise 99.4% Happy 0.5% Angry 0.1%	Angry 100.0% Happy 0.0% Surprise 0.0%	Surprise 99.9% Happy 0.1% Angry 0.0%
		
Happy 97.1% Angry 2.8% Surprise 0.1%	Happy 55.1% Angry 44.7% Surprise 0.2%	Angry 76.6% Happy 23.4% Surprise 0.0%
		
Surprise 99.3% Angry 0.5% Happy 0.2%	Happy 99.9% Surprise 0.1% Angry 0.0%	Surprise 88.3% Angry 8.8% Happy 2.9%
Stochastic Gradient Descent (SGD)		

Potential Applications

- Animators
 - Automatic reference dataset.
 - Drawing -> Results of cartoons with similar emotions.
 - Automatic subtitles.
 - Recommendation Systems (Movies: which character is the happiest?)
- 
- A series of white lines of varying lengths and slopes are drawn in the bottom right corner of the slide, creating a modern, abstract graphic element.

References

- Abadi, M. et al. (2016) TensorFlow - A system for large-scale machine learning. In: CoRR cs.DC
- Bishop, C.M. (1995) Neural Networks for Pattern Recognition. New York, NY, USA: Oxford University Press, Inc. isbn: 0198538642.
- Gajjarla, V. and Gupta, A. (n.d) Emotion Detection and Sentiment Analysis of Images. In: cc.gatech.edu.

THANKYOU

