**Colab link:** <https://colab.research.google.com/drive/128U-q6Yh9FJy_v1PmYzN8ABEaTZX3aqG?usp=sharing>

**Github repo link:** <https://github.com/MYCOURSEWORK/groupproject.git>

**Introduction**

Different quantities of image data will be generated by the next generations of astronomy surveys, including the Large Canonical gospels Survey Network, the Broadening Infrared Research Telescope, as well as Euclid. The ability to recognize, classify, and evaluate images requires effective, standardized, and reliable approaches. Deep learning is considered the best machine learning arrives that is used in image recognition. Data science has been used in different types of fields in order to improve the efficiency of some specific technologies for providing the best recommendations. Here there are various effective technologies are used for making image detection so that the early detection of fake and real images can be possibly made.

**Technical details:**

CNN model has been developed to identify the fake image in APOD image collection.

The APOD image dataset has been gathered from effective resources and this dataset has two types of features, one is the training dataset and another is the test dataset. At first, various libraries are imported and after that, by mounting the drive the image dataset can be read in the google collab platform for further implementation. the images can be perfectly analysed by merging the data for making a training model of the dataset. The train and the test splitting have occurred in order to apply the deep learning module in a perfect way. Each pixel is given a subclass at the final sparse coding stage that can be employed to mask resources and create slashes from real images (Fluke and Jacobs, 2020). Because improve training, the "smaller version masking" function stores mask forms at gridding locations.

Created a sequential model to train the dataset, the model created with batch normalization and dropout layers. Dropout is a regularization technique for neural networks where, during training, some number of layer outputs are randomly ignored or "dropped out." This helps to prevent complex co-adaptations on training data.

Convolutional neural network (CNN) model created using the Keras Sequential API. The model starts by adding a 2D convolutional layer with 64 filters of size 3x3 and using "same" padding, which means that the input and output dimensions of the layer are the same. The input shape for this layer is specified as the shape of the x\_train data, with the first dimension removed. The next layer applies a rectified linear unit (ReLU) activation function, followed by a batch normalization layer that normalizes the data along the first axis.

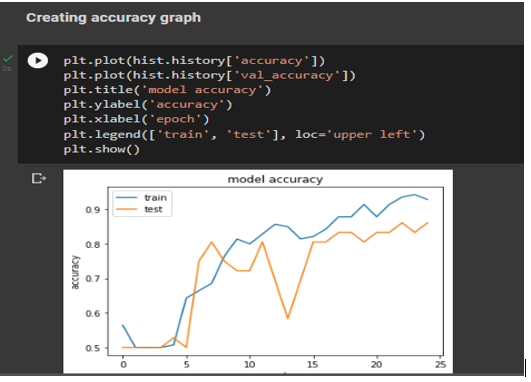
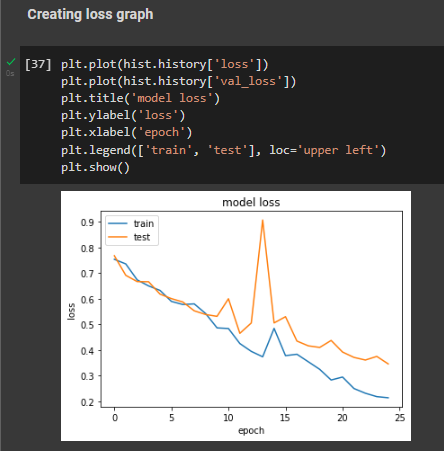
The next two layers are identical to the first, consisting of a 2D convolutional layer with 64 filters of size 3x3 and a ReLU activation, followed by a max pooling layer with a pool size of 2x2. The model then flattens the output from the max pooling layer and applies a batch normalization layer.

The next layer is a dense (fully connected) layer with 512 neurons and a ReLU activation, followed by a batch normalization layer and a dropout layer with a rate of 0.5. The final layer is a dense layer with 2 neurons and a sigmoid activation, which is used for binary classification.

Model has been compiled with different combination of optimizer (Adam and SGD) and loss functions (categorical\_crossentropy, binary\_crossentropy). When train the model with Adam optimizer and binary\_crossentropy, it giving better result. Model is giving training accuracy around 93 % and difference in validation and test loss is less compared to other combinations of optimizer and loss functions.

PFB screenshot with binary\_crossentropy and Adam:



****

In this graph accuracy for test set is low. It shows we need to train the model with more data to reduce the overfitting.

The model performance has been analysed using TensorFlow profiler

# 

# References

Burke, C.J., Aleo, P.D., Chen, Y.C., Liu, X., Peterson, J.R., Sembroski, G.H. and Lin, J.Y.Y., 2019. Deblending and classifying astronomical sources with Mask R-CNN deep learning. *Monthly Notices of the Royal Astronomical Society*, *490*(3), pp.3952-3965.

Fluke, C.J. and Jacobs, C., 2020. Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(2), p.e1349.

Paillassa, M., Bertin, E. and Bouy, H., 2020. MAXIMASK and MAXITRACK: Two new tools for identifying contaminants in astronomical images using convolutional neural networks. *Astronomy & Astrophysics*, *634*, p.A48.

Shandonay, A., Morgan, R., Bechtol, K., Bom, C.R., Nord, B., Garcia, A., Henghes, B., Herner, K., Tabbutt, M., Palmese, A. and Santana-Silva, L., 2022. Expediting DECam Multimessenger Counterpart Searches with Convolutional Neural Networks. *The Astrophysical Journal*, *925*(1), p.44.