Astronomy Picture of the Day

Overview

The project will be based on the detection of real and fake images from the NASA image database. This will help detect actual and fake astronomical objects with the application of deep learning algorithms (Revi, et al., 2022). To execute the project, several operations such as image processing, preprocessing, and visualization have been done on the database followed by the application of a deep learning model to detect the real and fake images (Wang, et al., 2021).

Objectives

The objectives are stated below:

1. To upload the database to google drive and create a notebook in google colab
2. To select the libraries that will be used for the artefact.
3. To read the database and analyze the images to visualize the counts of real and fake images from the directories.
4. To extract features from the image database using the image augmentation process and prepare the training and validation data for detection.
5. To design the deep learning model (Convolutional Neural Network) and train it with the training data and validate it with validation data.
6. To evaluate the model and finally test the images to show the detection result.

Data Description

The database has been collected from the NASA image repository which contains images of real and fake astronomical objects (APOD, 2021). A brief description of the database is presented below:

Table 1 Description of Database

Directory Name	Subdirectory	Image Types	Image Counts	Image Category
Train	Real	PNG	5000	RGB
	Fake	PNG	2000	RGB
Valid	Real	PNG	1510	RGB
	Fake	PNG	673	RGB

So, the database contains 7000 images in two sub-directories for training the deep learning model and the validation directory contains 2183 images for validating and evaluating the deep learning model.

Image Analytics

Primarily, the analytics have been done on the images of the database and prepared for further execution such as the detection of real and fake images (Demirović, et al., 2018). The process of analytics and the relevant outcomes are presented below.

Image Processing

The images from the directories have been counted. The count of the images has been shown below from the train and validation directories:

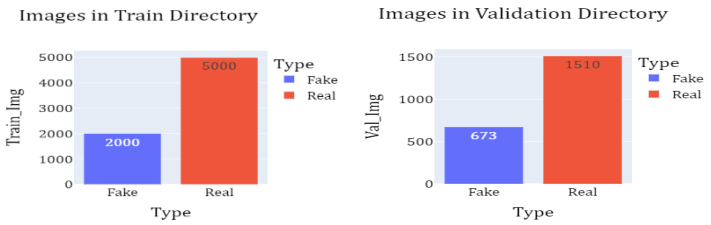

Figure 1 Image Counts

Image Visualization

Next, the images have been processed and two sample images have been visualized from the train directory of real and fake categories. Similarly, the images from real and fake subdirectories of the validation directory have been processed and visualised below:

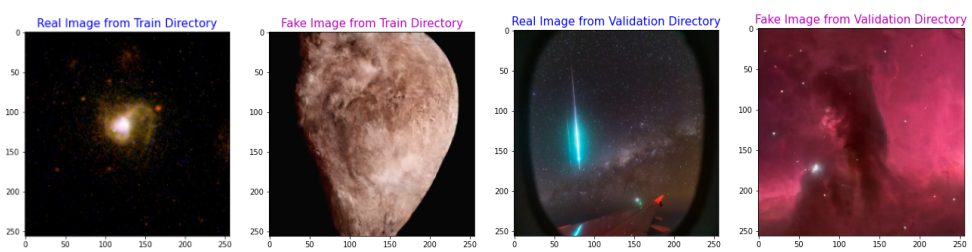

Figure 2 Sample Images from Train and Test Directories

Feature Extraction

Feature Extraction is a process through which the actual image will be transformed into an image array. As the raw images are not compatible to be processed in the algorithms, so this conversion is essential (Razazzadeh & Khalili, 2015). Feature extraction has been performed here with the application of image augmentation by using ImageDataGenerator. The parameters are discussed below:

Table 2 Parameters of Feature Extraction

Parameter	Value
Image Rescaling	1/255
Degree of Rotation	90°
Range of image Shearing	20%
Range of Image Zooming	30%
Range of Image Dimension Shifting	Height (20%)
	Width (20%)
Range of Image Flipping	Horizontal
	Vertical

Design of Model

Selected Model

Convolutional Neural Network has been chosen as a neural network model in this project to detect real and fake images. It is a configurable neural network where the layers and parameters can be added and tuned to make it work perfectly (Li, 2021). It has three sections. The first section is the Input Layer where the image array will be provided with a predefined resolution. The second section is the Hidden layer where the image arrays are processed and prepared for detection. The last section is the Output layer from where the result of detection will be obtained (Xin, et al., 2020). The generalised stricture of the Convolutional Neural Network is shown below (Jordan, 2017):

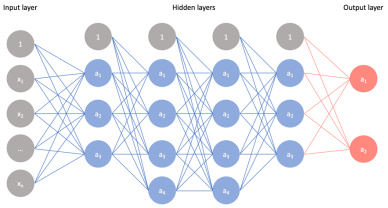

Figure 4 Structure of Convolutional Neural Network

Design of Convolutional Neural Network

1. Convolutional Layer: This layer (2D) has been added in five sequential sections of the model where the number of neurones has been increased by 2 times in each section starting from 16 neurones (For example Section 1 16 neurones, Section 1 32 neurones and so on).
2. Pooling Layer: This layer has been added just after Convolutional Layer in each section with the max pooling process.
3. Dropout Layer: This layer has been added just after the Pooling Layer in each section with 20% of the input drop which has been coming from the previous layer.
4. Dense Layer: This layer has been added at the end of the hidden layer section with 128 neurones and the output layer, contains 2 neurones.

Model Summary

The model summary has been discussed below with the generated parameters:

Table 3 Model Summary

Layer	Section	Parameters
Convolutional Layer	Section-1	448
	Section-2	4620
	Section-3	18496
	Section-4	73856
	Section-5	295168
Dense Layer	Section -6	295040
	Section -7	258

Now, after designing the complete model, the total parameters have been found by 687,906 where all parameters are trainable.

Model Compilation

The model has been compiled to make it prepared for training and detecting real and fake images. The compilation has been done using the Adam optimizer, Categorical Crossentropy loss function and Accuracy metric.

Image Detection

Model Training and Evaluation.

The model has been trained with training data (prepared using augmentation) and validated with validation data using a batch size of 32. In this training process, 100 epochs have been used after which the model has been evaluated and the final prediction has been done. The model has been evaluated in two phases. In the first phase, the evaluation has been done using training data to determine training accuracy and losses. In the next phase, the evaluation has been done using validation data to determine validation accuracy and losses. The accuracies and losses for training and validation are shown below:

Table 4 Result of Model Evaluation

Evaluation Using	Metrics	Value
Training Data	Loss	0.207
	Accuracy	90.42%
Validation Data	Loss	0.537
	Accuracy	83%

Image Detection Result

Finally, to check the detection result, the images have been taken for the test cases and the model has been applied to those. In this context, one image has been taken from the real category and another from the fake category. The result of detections are shown below:

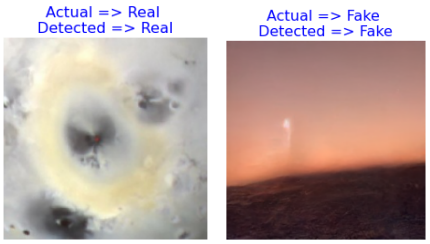

Figure 5 Detection Result

From the results of the detection, it can be seen that the Convolutional Neural Network is detecting both categories precisely.

Discussion

In this project, the detection of real and fake images of the astronomical body has been done from the APOD space image database. For this purpose, the images have been processed and visualised followed by the extraction of features. After that, the model for Convolutional Neural Network has been prepared and compiled. Using this compiled model, the detection of real and fake images has been conducted where the Convolutional Neural Network has gained 90.42% accuracy. To judge the effectiveness, the model has been applied to images and the outcomes have shown the precise detection of real and fake images.

]References

APOD, 2021. *Astronomy Picture of the Day.* [Online]
Available at: https://apod.nasa.gov/apod/ap211109.html
[Accessed 2022].

Demirović, D., Skejić, E. & Šerifović–Trbalić, A., 2018. Performance of Some Image Processing Algorithms in Tensorflow. *25th International Conference on Systems, Signals and Image Processing (IWSSIP),* pp. 1-5.

Jordan, J., 2017. *Convolutional neural networks.* [Online]
Available at: https://www.jeremyjordan.me/convolutional-neural-networks/

Li, H., 2021. Computer network connection enhancement optimization algorithm based on convolutional neural network. *International Conference on Networking, Communications and Information Technology (NetCIT),* pp. 1-4.

Razazzadeh, N. & Khalili, M., 2015. A high performance algorithm to diagnosis of skin lesions deterioration in dermatoscopic images using new feature extraction. *IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE),* pp. 1-4.

Revi, R., Isaac, M. M., Antony, R. & Wilscy, M., 2022. GAN-generated Fake Face Image Detection using Opponent Color Local Binary Pattern and Deep Learning Technique. *International Conference on Connected Systems & Intelligence (CSI),* pp. 1-5.

Wang, Y., Zarghami, V. & Cui, S., 2021. Fake Face Detection using Local Binary Pattern and Ensemble Modeling. *IEEE International Conference on Image Processing (ICIP),* pp. 1-6.

Xin, R., Zhang, J. & Shao, Y., 2020. Complex network classification with convolutional neural network. *Tsinghua Science and Technology ,* 25(4), pp. 447 - 457.