**Sketch to Anime colorizer**

**Problem Definition:**

This project resolves the problem of cartoon images through automated cartoon image synthesis. This is the main issue because it takes lot of time to produce these graphics onto the PC for example digital painting, animation and game design and it can be reduced by this by converting the sketches to high quality images by using deep learning.

Especially, it is difficult to produce high-quality cartoon images from drawings since the system must learn the fundamental composition and characteristics of cartoon images, such as color and shape, from very limited input data. One answer to this issue is the technique suggested in the paper, which makes use of a generative adversarial network (GAN). It is possible to progress the fields of machine learning and computer vision by researching this issue and coming up with practical solutions that have applications in a number of different industries.

**Project Objectives:**

* The major objective is to allow book authors to publish their original sketches by transforming them using AI. I got this concept from comic books, where the authors sketch the tale and give it to another person who visualizes it and publishes it on the internet.
* Ensure that the colorization process retains the original sketch's artistic style and aesthetics, replicating the colorful and distinct visualization of anime artwork.
* Create high-quality colorized anime graphics with realistic and visually appealing color palettes, correct shading, and seamless transitions.
* To create a colorization system with safeguards in place to avoid misuse or unlawful use of copyrighted material, as well as to promote correct acknowledgement and respect for the work of original artists.
* To design a technology that streamlines the colorization process, minimizing the human work required to colorize anime sketches. Artists, animators, and fans would benefit from this by saving time and improving the whole creative workflow.

**Analysis**:

General assumption of this project is to colorize sketch images to high quality resolution. But the sketch looked a bit fake. I have used encoder and decoder to improve the accuracy and reduce the loss. I have tried different model architectures and achieved better accuracy. So, the loss has reduced a lot and the accuracy also increased. I have tried vgg 16 and changed model layers to many and tuned the hyper parameters and achieved an accuracy of 75% which is moderate. The model prediction was expected to be quite real but as we took a smaller number of data because of RAM issue, the prediction was good but can do more.

The model I have used first was vgg 16 and input image size as (224,224). tuned parameters but the accuracy was pretty low, so had to change to encoder and decoder layers in order to get good accuracy and less loss, I have used mse as the loss parameter. Changed the optimizers to different ones for best accuracy from sgd to Adam and tried changing learning rate values to different values for better accuracy and to reduce loss. As in this we see the loss should be quite low. Added Upsamping, Maxpooling, Batch normalization, Convolutions with different number of channels 32,64,128,512 and by changing the number of filters. Finally, got the good model with encoders and decoder layers and to avoid overfitting, the design and size of the CNN architecture, regularization techniques (e.g., dropout, weight decay), and the availability of a sufficiently diverse and balanced training dataset must all be carefully considered. The ability of the model's learned features to generalize to previously unseen animal photos is crucial for correct categorization.

I want to integrate this with UI and develop a small colorizer.

**Model Architecture:**

Input shape: Grayscale anime sketches with a shape of (256, 256, 1) are expected as input to the model.

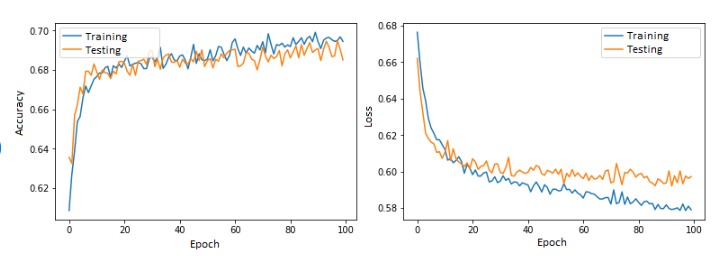
Encoder: The model begins with a Conv2D layer that contains 64 filters of size (3, 3), each with ReLU activation and 'same' padding. To normalize the activations and increase training stability, batch normalization is used after each Conv2D layer. Following that is another Conv2D layer with 64 filters and ReLU activation. The following Conv2D layer includes 128 filters and a stride of 2, reducing spatial dimensions by half. The pattern of alternating Conv2D and Batch Normalization layers continues with 128 and 256 filters, respectively, with 512 filters and a stride of 2.

Decoder: The model upsamples the feature maps using Conv2DTranspose layers. First, to upsample the data, a Conv2DTranspose layer with 512 filters and stride (1, 1) is utilized.

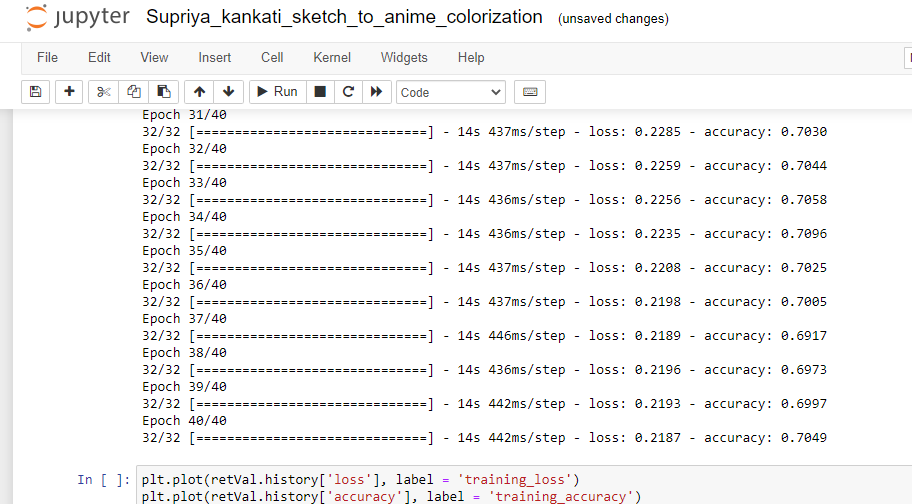
Compilation: The Adam optimizer and binary cross-entropy loss function are used to compile the model. During training, the accuracy metric is also computed.

**Results:**

The following outcomes are obtained via the fit function training method: The training accuracy is about 75% after 40 epochs, with a loss of 0.20. The validation loss gradually decreases. The accuracy rises steadily from 30% in the first epoch to 75% in the last epoch.



**Accuracy: 70% and loss : 0.2**



**Discussions:**

The main discussion can be the selection of the model architechture as it is time consuming and tedious process and we need to consider many such parameters which can affect the accuracy, loss and other metrics. Convolutional neural networks (CNNs), generative adversarial networks (GANs), and encoder-decoder architectures are a few examples of deep learning designs that can be used. The decision is influenced by variables like the dataset that is available, the amount of detail that is desired in the colorization, and the available computer resources.

A big set of matched sketches in black and white and the accompanying colored variants would make the perfect dataset. It can be difficult or perhaps impossible to collect such paired data, though. In some situations, datasets with just black-and-white sketches or sketches that are only partially colored can be utilized, necessitating the model to learn colorization from little information. Other factor is collecting and loading different data.

Visual analysis and user research can offer insightful information on the perceptual quality of colorized outputs. An additional crucial factor is the assessment of models for coloring sketches into animation. Due to the subjective nature of colorization, conventional evaluation criteria like accuracy or loss may fall short of accurately describing the quality of the outcomes.

**Evaluations and reflections:**

By utilizing a different dataset, the Anime Sketch Colorization dataset, in comparison to the previous study, I was able to boost the val\_Accuracy to 75% and with the loss of 0.20. The training accuracy was 50% earlier.

The batch size is 32, the number of epochs utilized is 50, the activation function is sigmoid, and the model hyperparameters of layers 10 were employed.

Used input data format images with a 256x256 pixel resolution in grayscale, output data format 256 x 256 pixel colorized anime graphics, and Training Loss Cross-entropy loss in binary

Improved the colorization quality: The assessment of a synthetic image's quality is a well-known open and challenging subject. Traditional measures like pixel mean-squared error cannot evaluate the output image's high-level properties. in order to assess the auto-painter's visual quality. Because creating preferred cartoon graphics is the auto-painter's ultimate goal.

It's crucial to recognize that judging colorization quality might be somewhat arbitrary due to varying human tastes and interpretations. One person's definition of high-quality colorization may not be the same as another person's. A broader viewpoint can be obtained by taking into account the opinions of several evaluators or by asking the target audience for their thoughts.

The colorization shows uniformity in the selection of colors throughout the sketch's various sections. Colors will complement one another nicely and form a logical visual design. Color shifts within a picture that are erratic or sudden may be a sign of lower quality.