Colourisation of GrayScale Images

Aanchal Narendran
PES1UG19CS006

A Narendiran
PES1UG19CS001

Abdul Rahman
PES1UG19CS009

Abhishek D
PES1UG19CS020

***Abstract*—The project aims to colourise grayscale images using Image Processing and Deep Learning techniques. Colourising an image is one such problem that doesn't have a clear cut solution or approach. A wide range of authors have developed novel Convolution based approaches, Auto-Encoder based models and so on for the same. This project aims to enhance images using Filtering, Morphological and Segmentation techniques to better map a RGB value to a pixel.**

***Index Terms*—Image Processing, Deep Neural Networks, Image Colourisation**

I. Introduction

Image colourisation involves using a wide range of statistical models to estimate the red, green and blue values of pixels in a grayscale image. In the past decade, there has been a lot of interest in colourising grayscale images. Most photographic solutions between 1826 - 1980 have relied on capturing their images in Grayscale. Colourising these images is similar to art restoration. Although old, they are of importance and hold sentimental and historic value. CCTV solutions over the past decades have predominantly been in black and white. Effectively limiting the usage of these cameras in identifying vehicles and people of interest in legal cases. Although luminance and grayscale help isolate information regarding the objects in the image, colour helps improve the quality of information obtained from grayscale images. This project aims to develop a solution to aid in colourising such images.

Due to the complexity of colourising images, there is no standard solution for this problem. This project explores a wide range of models and image processing methods to enhance estimates of colour and improve on present colourising solutions. The present solution is to convert to grayscale a colourised dataset like the Animal 151 dataset and to explore and validate colourisation strategies on this data where the colourised version serves as a benchmark. Post colourisation, pre-processing methods such as filtering, smoothing, and so on will be used to bring the grayscale image to a usable state.

II. Literature Review

A. Collaborative Image and Object Level Features for Image Colourisation

The objective of the paper [1] is to introduce a novel approach named UCapsNet, which considers image and entity-level representations to understand image context better. The paper develops an unsupervised learning model that not only isolates the best RGB values at a location but also provides a probability distribution of potential RGB values at the location to isolate better colour schemes. The model developed extensively uses convolutions, capsule layers and basic image processing to predict the pixel value at a location. Convolutions allow the model to extract and identify features while Capsule layers enable isolation and understanding of object entities. Skip connections are extensively used to provide a bridge between image-level and entity-level features. The model is divided into a downsample and upsample phase. Image-level representations are learnt using convolution layers and a double block down operator, a sequence of Convolution, Batch Normalisation and ReLU operations. The entity-level representations are learnt using Primary Caps Down, flattens repeated Convolutions, and a routing agreement. The model developed was trained and tested on a variety of benchmark datasets such as ImageNet, COCOStuff, and Places205. Apart from quantitative evaluation with Peak Signal to Noise Ratio and Learned Perceptual Image Patch Similarity, the authors used a sample of 180 participants to understand which methodology generated better-colourised images.

Some of the assumptions in the paper are nested in its ideology that image-level and entity-level representations are enough to effectively model image colourisation. Moreover, the pre-processing strategy of the model involves the usage of black box layers that do not allow for the effective isolation of other factors impacting colourisation. Additionally, the paper mentions that each epoch on a GPU takes over 10 hours. This makes the model less deployable and less likely to be used commercially. However, the convolution architecture is a brilliant starting point for improved models, the pre-processing techniques could be made more image-centric.

B. Image Colorization Using Similar Images

The paper [2] leverages exemplar images to map the RGB values using superpixels. Exemplar images are coloured images that are semantically similar to the target image (the exemplar image has elements present in the target image). The author uses superpixels to map a region from the exemplar image to the target (grayscale) image. The paper addresses a multiple exemplar image model to colourise any grayscale image. Superpixel extraction happens via a geometric-flow based algorithm that computes compact superpixels with standard shape and size while preserving original image edges. A total of 172 features are extracted from the images. SURF and Gabor feature easily discriminate between superpixels.

The intensity and the standard deviation is easily computable and available but they give 4 parameters for discrimination. Whereas, SURF and Gabor give the model 128 and 40 features respectively to enhance discrimination. These features for a superpixel is calculated by aggregating the values of the pixels that form the superpixel. Fast cascade feature matching along with euclidean distance is used to compute the similarity between a superpixel from the exemplar image and the target image. A voting strategy is used amongst the neighbours of a superpixel to fix an invalid assignment. K-Means clustering helps isolate dense clusters and sparse clusters. Dense clusters have near-accurate colour mapping. The images obtained were visually evaluated by a panel of 30 volunteers.

The authors were able to back their superpixel assumption with sound results where 65% of a panel of 30 volunteers categorised their image as real, as opposed to the nearest benchmark at 53.2%. Due to superpixelation, the model struggles with handling areas with a sharp increase in intensity or around the edges. It leads to bleeding artefacts. The model is heavily reliant on the availability of exemplar images. However, the model is nested heavily in standard morphological processes making it easier to fix an issue in the image as opposed to black-box models.

C. End-to-End Conditional GAN-based Architectures for Image Colourisation

The existing convolutional neural network impacts the colourfulness of an image. This paper aims to address this problem by using Generative Adversarial Networks to enable better generalisation. Additionally, integration of instance and batch normalisation further help generalise. The image is trained and tested using the ImageNet dataset. Generative Adversarial Networks mimic the natural colour distribution and make the samples indistinguishable from natural images. Spectral Normalisation methods further improve generalisation without hampering the stability of training by regularising the weights. Multi-scale discriminators improve generation in small areas and enhance local details. CIELab space is used as it is more perceptually uniform and linear. The generator and discriminator are trained using min-max to obtain nash equilibrium. The generator captures the original distribution of data while training a realistic mapping. Discriminator helps distinguish between real images and colourised ones. The generator is modelled towards generating images that are harder for the discriminator to differentiate. The U-Net is used as a generator while Markovian PatchGANs are used as discriminators.

The author of the paper uses a pure Black-box model to colourise the image. This doesn't leave a lot of space for handling noise and other image artefacts. Noise in datasets has a huge role in successful feature extraction and segmentation. However, the author is successful in demonstrating that Generative Adversarial Networks perform better in terms of colourfulness.

III. Dataset and Preprocessing

The dataset "Animals 151" will be used in this project. It contains RGB pictures of about 151 animals, with each animal having anywhere between 30 -60 images. Each image is of the size 224x224 pixels 1. The pre-processing of this project has been done entirely using Matlab.

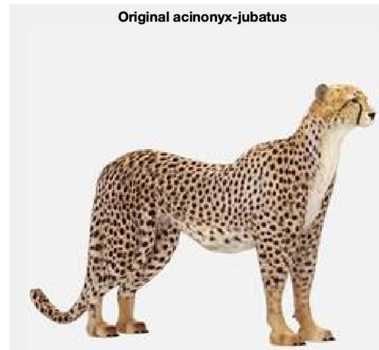

Fig. 1. Original Version of The Image

As a part of pre-processing, the first step is to convert the coloured image to greyscale 2. It is achieved via the rgb2gray function in Matlab. Due to the smaller size of the image, resizing tends to introduce a certain amount of unavoidable distortion. In an attempt to solve this artefact, unsharp masking was used 3. However, no perceivable impact was noted, visually. Finally, segmentation of the image using SuperPixels with the Simple Linear Iterative Clustering was done 5 4. Superpixels were extracted as a part of preprocessing due to the speedup achieved during colourisation [2].

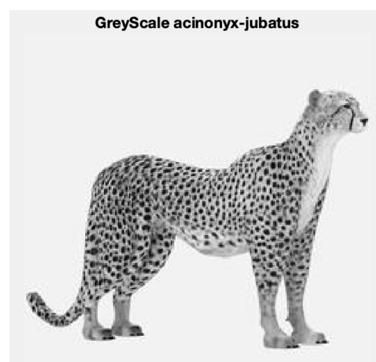

Fig. 2. GreyScale Version of The Image

IV. Conclusion

In conclusion, the literature review provided the team with certain directions to gauge the best pre-processing strategy and model development technique to use for image colourisation. Although, convolutional neural networks and other black-box models are heavily favoured in this setting. The paper aims to develop a strategy nestled in pure image processing and basic machine learning to develop colourised versions of grayscale images.

REFERENCES

[1] R. Pucci, C. Micheloni, N. Martinel, "Collaboration among Image and Object Level Features for Image Colourisation.", ArXiv, 2021

[2] R. K. Gupta, A. Y. Chia, D. Rajan, E. S. Ng, H. Zhiyong, "Image Colorization Using Similar Images", Proceedings of the 20th ACM international conference on Multimedia, 2012

[3] M. G. Blanch, M. Mrak, A. F. Smeaton, N. E. O'Connor, "End-to-End Conditional GAN-based Architectures for Image Colourisation", arXiv, 2019

Fig. 3. Sharpened version of the greyscale image

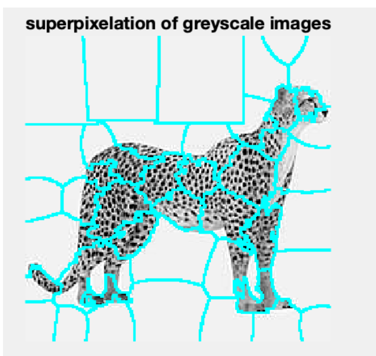

Fig. 4. Overlay of 50 Superpixels on greyscale image

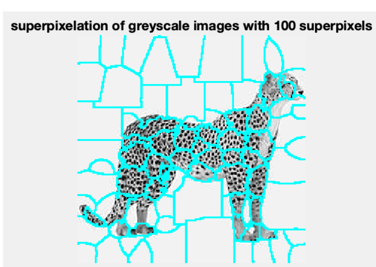

Fig. 5. Overlay of 100 superpixels on greyscale image