International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 6, Issue 4, April 2017, ISSN: 2278 -7798

Automatic Colorization Of Gray-scale Images using Deep Learning

Arshiya Sayyed
Dept. of Computer Engg.
Cummins College of Engg.,
Pune, India

Apeksha Rahangdale
Dept. of Computer Engg.
Cummins College of Engg.,
Pune, India

Rutuja Hasurkar
Dept. of Computer Engg.
Cummins College of Engg.,
Pune, India

Kshitija Hande
Dept. of Computer Engg.
Cummins College of Engg.,
Pune, India

Abstract— **Automatic colorization of gray-scale images using deep learning is a technique to colorize gray-scale images without involvement of a human. Conventional techniques used for colorizing images need human intervention, which is time-consuming. The project deals with deep learning techniques to automatically colorize gray scale images. The proposed technique uses deep convolutional neural networks and has a number of advantages. The technique will reduce manual work, speed up the process and improve the accuracy. Automatic colorization techniques using ConvNet finds applications in various domains such as astronomy, electron microscopy, and archaeology. Conventional approach to achieve colorization included regression-based model, graph cut algorithm etc. Proposed model is a classification based technique but uses regression model as the base line model. Designed system consists of training and testing phases. Feature extraction and pixel-mapping from the input coloured image results in training of the system. In the testing phase the system is provided with gray scale input images to check the accuracy of colorization of these images. This technique can be used to eliminate the need of expensive image transferring equipments for astronomical images and to speed up the process of conversion of legacy images to modern coloured images, thus reducing manual effort needed by utilizing deep learning techniques.**

Keywords—Deep Learning,Automatic Colorization,Gray Scale Images,Convolutional Neural Networks.

I. Introduction

The proposed technique focuses on reducing human intervention required to colorize grayscale images with the help of deep learning techniques. The grayscale image will pass through different layers of neural network and be colored automatically. This project is divided into four modules which includes process of encoding and decoding the input images, Intermediate resulting images from the specific color-space from the training image dataset, eventually providing the colored images.

Figure 1: Sample Gray Scale and equivalent color image

II. Need of colorization techniques

Today we are in midst of a digital revolution, we are moving towards better quality of images and higher resolutions. But even in today's day and age there are few applications that still capture images in gray scale, and the techniques used to convert them to colored images for different purposes are conventional and require a lot of human effort. Usually photo editing softwares like Photoshop and the likes are used to color the images, which is done by a human. Although these techniques are easier, they have a number of drawbacks such as if there are huge amount of images to be colored then the process is time consuming. Furthermore, with increasing advancements in science and technology using manual techniques to colorize images isn't practically sound. Hence we propose a technique wherein neural networks and artificial intelligence are used to colorize gray scale images automatically. Automation has spread across every field in recent years, hence even for image colorization there is a pressing need for an automatic technique. These techniques are particularly useful in fields like archaeology, astronomy and electron microscopy.

III. Background Techniques

In order to enable the system to automatically colorize gray scale images we need to give it the thinking capability similar to a human being and to do this we program the system with Artificial Intelligence. AI is a comprehensive field which includes a number of subfields as shown in Figure 2.

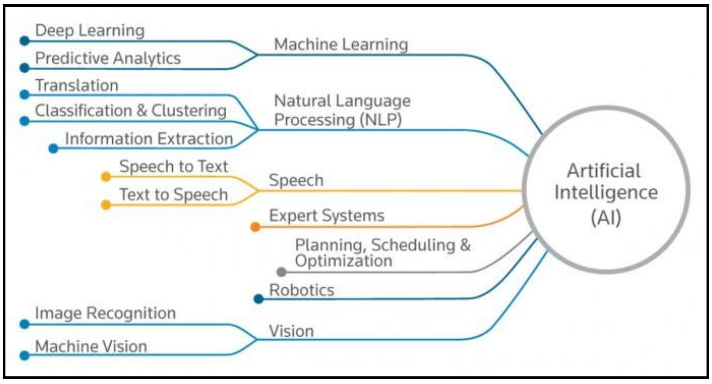

Figure 2 : Subfields included under AI

Our system makes use of deep learning technique, which is a part of Machine learning , which in turn is a subfield of Artificial Intelligence.

Deep learning is a technique which makes use of neurons connected to form neural network to train the network to perform a specific task. Using deep learning we can train our system to efficiently colorize gray scale images.
Artificial Neural Networks (ANN) are made up of artificial neurons which are modelled on biological neurons in our brains.
Figure 4 shows the structure of a biological neuron.

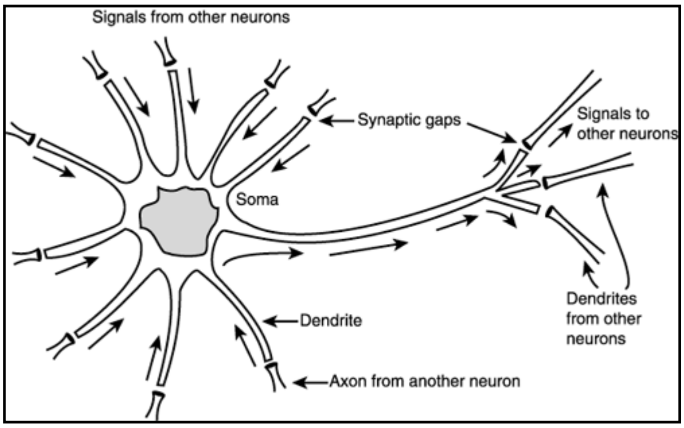

Figure 4 : Structure of a Biological Neuron

The brain is a collection of about 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive, process and transmit information. When one of those neurons fire, a positive or negative charge is received by one of the dendrites. The strengths of all the received charges are added together through the processes of summation. If the aggregate input is greater than the axon hillock's threshold value, then the neuron *fires*, and an output signal is transmitted down the axon. Dendrites bring in the signal from other neurons and axons transmit the signal to other neurons.
This biological neuron is used to create an artificial neuron as shown in Figure 5

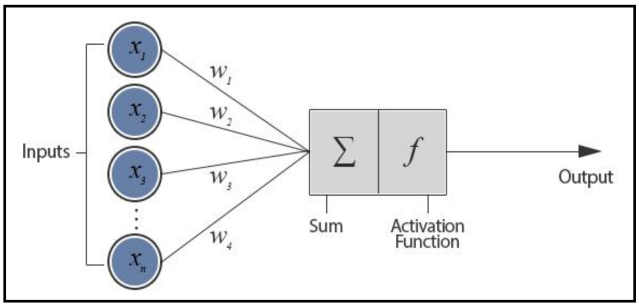

Figure 5 : Artificial Neuron/Perceptron

Artificial neurons are also called as Perceptron. As shown in the diagram above a typical perceptron will have many inputs and these inputs are all individually weighted. The perceptron weights can either amplify or deamplify the original input signal. For example, if the input is 1 and the input's weight is 0.2 the input will be decreased to 0.2. These weighted signals are then added together and passed into the activation function. The activation function is used to convert the input into a more useful output. There are many different types of activation function , we are using Rectified Linear Unit (ReLU) as the activation function.
Many such neurons are connected together to form artificial neural network, there are different types of neural networks, we use convolutional neural network (ConvNet) for our system because ConvNet architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture.
Figure 6 and 7 shows the structure a regular neural network and a ConvNet respectively

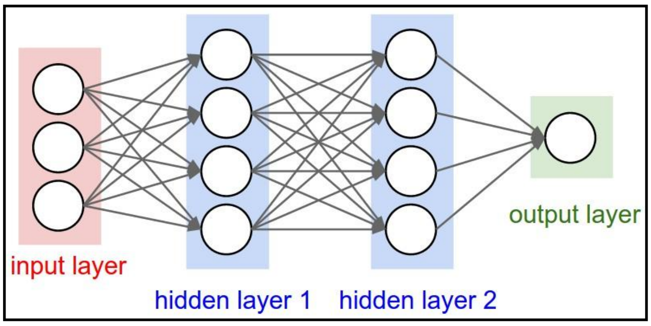

Figure 6 : Basic Neural Network

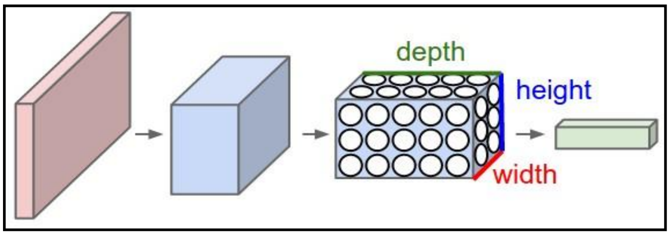

Figure 7 : Convolutional Neural Network

A ConvNet arranges its neurons in three dimensions (width, height, depth), as visualized in one of the layers. Every layer of a ConvNet transforms the 3D input volume to a 3D output volume of neuron activations. In this example, the red input layer holds the image, so its width and height would be the dimensions of the image, and the depth would be 3 (Red, Green, Blue channels).

IV. LITERATURE REVIEW

In the paper titled 'Image Colorization with Deep Convolutional Neural Networks' by the authors Jeff Hwang , You Zhou [1] have presented a convolutional neural network based system that faithfully colorizes the black and white photographic images without human assistance with the use of statistical learning driven approach. The authors build a learning pipeline that comprises a neural network and an image pre-processing front-end. A regression based model is used as the baseline which comprises of a 'summarizing', encoding process followed by a 'creating', decoding process. The model makes use of the MIT CVCL open country dataset.
R. Dahl in his paper titled 'Automatic Colorization' [2], relies on several ImageNet trained layers from VGG16, integrating

them with an auto-encoder like system with residual connections that merge intermediate outputs produced by the encoding portion of the network comprising the VGG16 layers with those later produced by the decoding portion of the network. Thus Dahl's system reports much larger decreases in training loss on each training iteration. In terms of results, Dahl's system performs extremely well in realistically colorizing foliage, skies, and skin. But, it is however, noticed that in numerous cases the images generated by the system are predominantly sepia toned and muted in color.

V. SYSTEM ARCHITECTURE

System design and flow is explained as follows:

A. Block Diagram

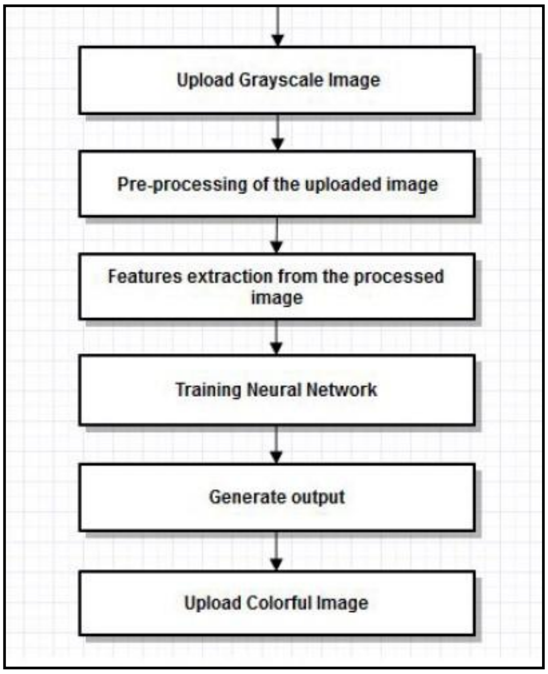

Figure 9 : Block Diagram for colorization process

B. System Flow

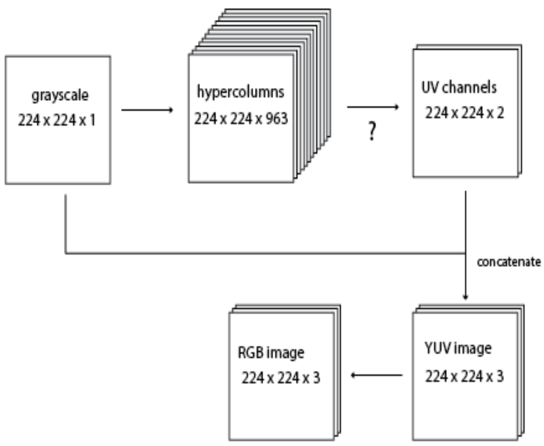

Figure 10 : System Flow through different modules

C. Functions used

1. **Activation Function**:

   Rectified linear unit for accelerating training convergence

$$\boxed{f(x)=\max(0,x)} \quad (1)$$

2. **Loss Function:**

   Derived loss function from Huber penalty function(Best of L1 and L2 norms)

$$L(u) = \begin{cases} u^2, & |u|<M \\ M(2|u|-M), & |u|>M \end{cases} \quad (2)$$

Where , u is residual
M is threshold

3. **Evaluation Function**

   To measure closeness of generated image and actual image

$$\text{Sat. diff.} = \frac{\left|\sum_{(i,j)}^{(N,N)} S_{p_{ij}} - \sum_{(i,j)}^{(N,N)} S_{a_{ij}}\right|}{\sum_{(i,j)}^{(N,N)} S_{a_{ij}}} \quad (3)$$

a. Encoding process
Involves operations to be performed on colored images i.e. converting them into different color space for eg. Lab color space for training purpose.

b. Decoding process
This is the reverse process of the above explained encoding process in which two channel Lab color space image will be converted back to the RGB three channel color space.

c. Discretization of Color Spaces
RGB two channel color spaces are discretized into 50bins forming the pixel maps and feature extraction matrices. Intermediate color space used is Lab –L for luminescence, a and b for two channels.

d. Convolutional Layers
2D Convolutional neural networks (ConvNet) are designed and programmed to form a layer of these networks through which gray-scale images will be processed.

e. Concatenation process

This process basically concatenates the result of the hidden layers in the network to produce the required output.

VI.Implementation

System is implemented in two phases:

Training phase

Provide a set of colored images to train the system. Convert the colored image to gray scale images for testing phase requirements.Input the gray scale image from intermediate phase to the system and system checks the deviation of output image and accordingly decides whether to finalize the image or adjust weights and repeat the process.

Testing phase

Provide other grayscale converted images from the data set for which system has not been trained and check for saturation among the actual colored image and generated colored image

A. Training phase

i. Provide a set of colored images to train the system.

ii. Convert the colored image to gray scale images for testing phase requirements.

iii. Input the gray scale image from intermediate phase to the system and system checks the deviation of output image and accordingly decides whether to finalize the image or adjust weights and repeat the process.

B.Testing phase

Provide other grayscale converted images from the data set for which system has not been trained and check for saturation among the actual colored image and generated colored image

IV. Tools And libraries

Image colorization using deep learning requires machine learning tools to build the system. These tools are described as follows:

A.Tensorflow :

Tensorflow is an open source tool, software library for machine intelligence. The flexible architecture allows you to deploy computation to one or more CPUs or GPUs. TensorFlow was developed for the purposes of conducting machine learning and deep neural networks research. Example- CIFAR-10 model , Multiple GPU cards.

Tensorflow library provides multiple functions to be performed in GPU based parallel platforms to enable simultaneous operations on multiple datasets at the same time. Approximately 500 images are given to the system for training purpose and feature extraction. Tensorflow in python enables programmers to write efficient codes for multiple image datasets.

B.Theano :

Theano is a numerical computation library for Python. In Theano, computations are expressed using a Numpy like syntax and compiled to run efficiently on either CPU or GPU architectures. Theano is an open source project primarily developed by machine learning researchers.

C.Lasagne

Lasagne is a lightweight library to build and train neural networks in Theano. Its main features are:

- Supports feed-forward networks such as Convolutional Neural Networks (CNNs), recurrent networks including Long Short-Term Memory (LSTM), and any combination thereof
- Allows architectures of multiple inputs and multiple outputs, including auxiliary classifiers
- Many optimization methods including Nesterov momentum, RMSprop and ADAM
- Freely definable cost function and no need to derive gradients due to Theano's symbolic differentiation.
- Transparent support of CPUs and GPUs due to Theano's expression compiler.

V. Dataset

Use of http://image-net.org/ (open source site for image datasets) is intended to be used for training the model. Below are a few sample images we shall be using –

Figure 11 : Sample Images from Image-net.org (Roman Buildings)

Image-net.org provides open source images from a number of different categories, the image sets are also downloadable thorough a URL file which can be later used to extract the images. We have considered mostly old architecture style images as they suit our application.

VI. Applications

We have shortlisted three major applications as follows

A. Archaeology

Images captured few decades ago when there were no color cameras are still in gray scale format and the process to convert them to color requires manual effort. Hence this application is chosen so that the process of colorization can be automized.

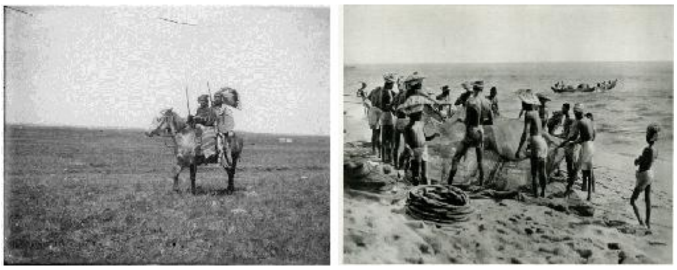

Figure 12 : Sample images for archaeology

B. Astronomy

Hubble telescope doesn't use color film. Its cameras record light from the universe with special electronic detectors.

These detectors produce images of the cosmos not in color, but in shades of black and white.

The matrix corresponding to the image is sent back to earth station, different filters are applied (red, green and blue) and the equivalent matrices are sent back to the earth station where these are combined to form a color image. This process is tedious and incures communication cost ,hence our system makes this process of colorization automatic and more effective.

Figure 13 : Sample images for astronomy

C. Electron Microscopy

Electron microscope is used to observe objects that cannot be seen with naked eye, such as viruses and bacterias.

The electron microscope functions by passing an electron beam and focusing it through electromagnets, this beam is either passed through the object or reflected off the surface of the object on fluorescent plate, this electron beam does not transfer color information, hence the images captured thereby are gray scale. Colors are essential in electron microscopy as it represents vital information. Hence our system would make this colorization process simpler.

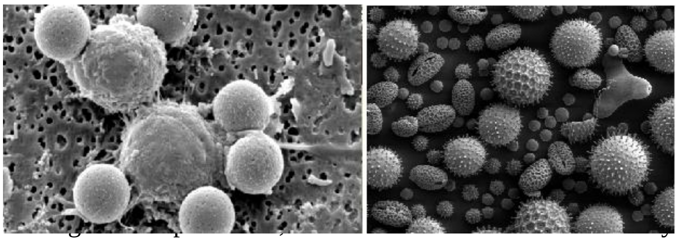

Figure 14 : Sample images for electron microscopy

empirically shown that formulating the task as a classification problem can yield colorized images that are arguably much more aesthetically-pleasing than those generated by a baseline regression based model, and thus shows much promise for further development. Our work therefore lays a solid foundation for future work. Moving forward, we have identified several avenues for improving our current system. To address the issue of colour inconsistency, we can consider incorporating segmentation to enforce uniformity in colour within segments. We can also utilize post-processing schemes such as total variation minimization and conditional random fields to achieve a similar end. Finally, redesigning the system around an adversarial network may yield improved results, since instead of focusing on minimizing the cross-entropy loss on a per pixel basis, the system would learn to generate pictures that compare well with real-world images. Based on the quality of results we have produced, the network we have designed and built would be a prime candidate for being the generator in such an adversarial network.

ACKNOWLEDGMENT

Our sincere thanks to the internal guide Prof. Hitendra Khairnar, Department of Computer Engineering , CCOEW, Pune for his continuous efforts and worthwhile guidance. Our gratitude to Image-Net for providing with the image datasets is duly appreciated

REFERENCES

[1] “Image Colorization with Deep Convolutional Neural Networks” by Jeff Hwang and You Zhou.

[2] R. Dahl. “Automatic colorization.” http://tinyclouds.org/colorize/, 2016.

[3] Charpiat, G., Bezrukov, I., Altun, Y., Hofmann, M., Sch¨olkopf, B. “Machine learning methods for automatic image colorization.”

[4] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: “Semantic image segmentation with deep convolutional nets and fully connected crfs” In: ICLR (2015)

[5] X. Glorot and Y. Bengio. “Understanding the difficulty of training deep feedforward neural networks.” In International Figure 2 Sample images from the MIT CVCL Open Country conference on artificial intelligence and statistics, pages 249–256, 2010.

[6] Cheng, Z., Yang, Q., Sheng, B.: “Deep colorization” In: ICCV (2015)

[7] Doersch, C., Gupta, A., Efros, A.A.: “Unsupervised visual representation learning by context prediction” In: Proceedings of the IEEE International Conference on Computer Vision. pp. 1422–1430 (2015)

[8] B. Hariharan, P. Arbel´aez, R. Girshick, and J. Malik. “ Hypercolumns for object segmentation and fine-grained localization” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 447–456, 2015.

[9] K. He, X. Zhang, S. Ren, and J. Sun. “ Deep residual learning for image recognition” arXiv preprint arXiv:1512.03385, 2015.

[10] S. Ioffe and C. Szegedy. “ Batch normalization: Accelerating deep network training by reducing internal covariate shift” arXiv preprint arXiv:1502.03167, 2015.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. “ Imagenet large scale visual recognition challenge” International Journal of Computer Vision, 115(3):211–252, 2015.

[12] “Lasagne” https://github.com/Lasagne, 2015

[13] https://www.tensorflow.org/

[14] https://developer.nvidia.com/cuda-downloads

[15] K. Simonyan and A. Zisserman. “ Very deep convolutional networks for large-scale image recognition” arXiv preprint arXiv:1409.1556, 2014.

[16] I. Sutskever, J. Martens, G. Dahl, and G. Hinton. “ On the importance of initialization and momentum in deep learning” In Proceedings of the 30th international conference on machine learning (ICML13), pages 1139–1147, 2013.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton. “ Imagenet classification with deep convolutional neural networks” In Advances in neural information processing systems, pages 1097–1105, 2012.

[18] M. J. Huiskes and M. S. Lew. “ The mir flickr retrieval evaluation” In Proceedings of the 1st ACM international conference on Multimedia information retrieval, pages 39–43. ACM, 2008.

[19] A. Olmos et al. “ A biologically inspired algorithm for the recovery of shading and reflectance images” Perception, 33(12):1463–1473, 2004.

A. Oliva and A. Torralba. “Modeling the shape of the scene: A holistic representation of the spatial envelope” International journal of computer vision, 42(3):145–175, 2001