International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 2, Mar- Apr 2024
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

BLACK AND WHITE IMAGE COLORIZATION WITH OpenCV AND DEEP LEARNING

Dr. Bhargavi Peddi Reddy, Kallemollu Upendra, Perka Uday Kiran
Department of Computer Science, Vasavi College of Engineering, Hyderabad, India
bhargavi@staff.vce.ac.in, k.sunny.upendra@gmail.com,160220b044@gmail.com

Abstract:
This paper presents an innovative approach to automatic black-and-white image and video colorization using Convolutional Neural Networks (CNN). Our method focuses on predicting image color based on grayscale input using the Lab Color space. Leveraging the extensive ImageNet dataset, we curated a mini dataset of 39,604 images, dividing it into 80% for training and 20% for testing. Evaluation using Mean-squared error (MSE) and peak signal-to-noise ratio (PSNR) yielded promising results, with an average MSE of 51.36 and 31 PSNR.

Colorizing blackandwhite images can be a tedious and time-consuming task.Traditional methods, such as Photoshop editing, require extensive manual intervention and can take up to a month per image. In contrast, our approach integrates sophisticated image colorization techniques with cutting-edge deep learning principles. By harnessing the power of CNNs, we automate the process of feature extraction from training data, significantly reducing the time and effort required for colorization.
Our study also delves into the broader landscape of image colorization, where digital image processing meets deep learning. We explore various CNN and Generative Adversarial Network models, utilizing pre-trained architectures to improve feature extraction capabilities. This integration of deep learning with transfer learning enables us to streamline the colorization process and achieve superior results compared to manual methods.

Keywords:CNN, GAN, image colorization, deep learning, transfer learning, pre-trained models, Lab color space.

1 Introduction

In the realm of photography's evolution, the transition from black and white (B&W) to color imagery has been a transformative journey. Early attempts at colorizing B&W images involved painstaking manual processes using paints and brushes, often taking days or even weeks to achieve a semblance of reality. As technology advanced, software like Adobe Photoshop and GIMP facilitated digital colorization, albeit still requiring meticulous pixel-level adjustments and manual interventions.

Today, with a vast repository of photographs captured in color, the task of colorizing old B&W images has taken a new turn. Modern techniques leverage the wealth of information provided by color cameras, enabling the extraction of color schemes, object hues, lighting nuances, and intensity variations from a plethora of images. This wealth of data, combined with advancements in deep learning, has opened avenues for automating and modernizing the colorization process.

Our study delves into this intersection of historical imagery and cutting-edge technology.Employing Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs),we explore sophisticated models for image colorization. These models are trained on custom datasets meticulously curated to encompass diverse scenery, backgrounds, and artistic themes, ensuring the capture of generic yet nuanced color patterns.

Moreover, our research extends beyond mere implementation, focusing on quantitative and qualitative assessments of different CNN and GAN architectures. We demonstrate the efficacy of incorporating pre-trained models, streamlining training processes, and enhancing overall performance while reducing manual intervention.

Through this exploration, we aim to bridge the past and present of image colorization, offering insights into the transformative role of deep learning in revitalizing historical imagery and easing the burden on artists engaged in colorization tasks.

2 Literature Survey

Recent breakthroughs in deep learning have fueled remarkable progress in bringing color back to black and white photos. Researchers have made significant contributions using convolutional neural networks (CNNs) to achieve this. Early approaches focused on creating natural-looking colorizations, with works by Zhang et al. (2016) and Larsson et al. (2016) paving the way. Iizuka et al. (2016) further enhanced these techniques by incorporating knowledge about typical image features into the colorization process. The field took another leap forward with the introduction of generative adversarial networks (GANs) by Isola et al. (2017). GANs are powerful tools for image-to-image translation tasks, and colorization is a perfect example. Zhang et al. (2019) built on this by developing a system that allows users to guide the colorization process for even more control. Finally, Cheng et al. (2018) combined global and local image information to achieve cutting-edge colorization results. These advancements collectively demonstrate the immense potential of deep learning in transforming black and white photos into vibrant color representations of the past.

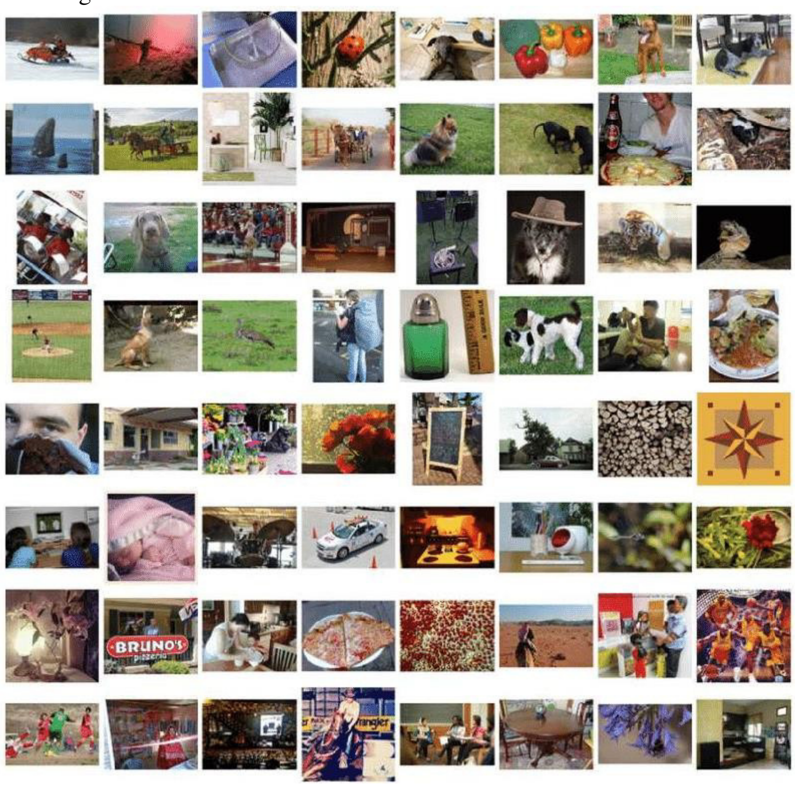
3 Design

A. Dataset

The dataset utilized in this research is a subset of the COCO (Common Objects in Context) dataset, which is a part of the larger ImageNet dataset, one of the most comprehensive resources in computer vision research. ImageNet comprises millions of labelled images across numerous categories, providing a rich source for training and evaluating machine learning models. The COCO dataset, derived from ImageNet, focuses specifically on images depicting common objects within their contextual surroundings, making it particularly suitable for tasks such as object detection, segmentation, and image captioning.

For this project on black and white image colorization, a subset of the COCO dataset consisting of 8000 images was selected. These images were meticulously curated to represent a diverse range of scenes, objects, and lighting conditions, ensuring a comprehensive and representative training set for the colorization model. By leveraging a subset of the COCO dataset, the project benefits from the extensive annotations and high-quality images present in ImageNet, enabling the development and evaluation of the colorization model with robust and varied data.

The dataset was divided into training and validation sets, with approximately 60% (4800 images) allocated for training and the remaining 40% reserved for validation purposes. This division ensured that the model was trained on enough data while also allowing for rigorous evaluation of its performance on unseen images. Leveraging a subset of the COCO dataset provides several advantages, including access to a wide variety of objects, scenes, and contexts, which are crucial for training a colorization model capable of handling diverse real-world scenarios.



B. Algorithms used

Convolutional Neural Networks (CNNs) are widely employed in image recognition tasks due to their ability to analyze two-dimensional arrays efficiently. These networks operate by processing data through several layers, each designed for specific operations. These layers typically include convolutional layers, rectified linear unit (ReLU) layers, pooling layers, and fully connected layers..

4 Implementation

Brief Approach

Our study aims to transform grayscale images, represented by a single channel of image data, into standard RGB images consisting of three channels of data.

Toaccomplish this, we employ the CIE Lab color space for both input and output images due to its effective separation of lightness (intensity) and color components.

Our methodology centres around training Convolutional Neural Network (CNN) models to convert grayscale images into colorized Lab space representations, which are subsequently converted to RGB images. The primary goal of the models is to learn a mapping function that predicts the associated color channels (a, b channels) based on the input grayscale image. We utilize the Euclidean loss (L2 loss) function, commonly known as mean squared error (MSE) loss, to quantify the disparity between predicted and ground truth Lab images.

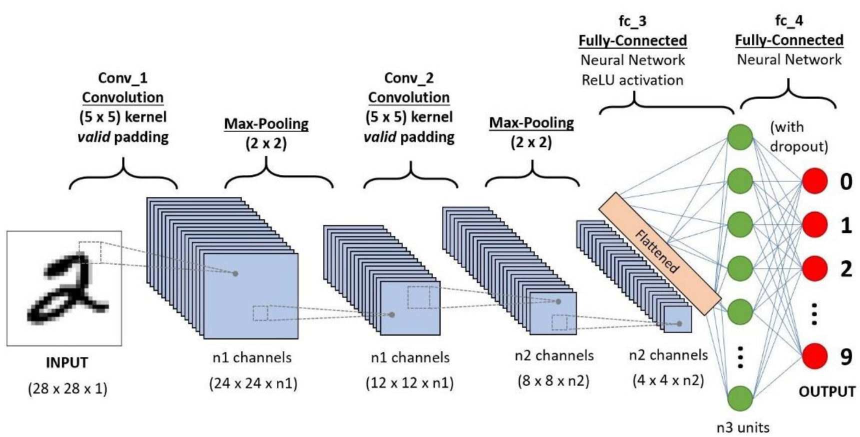
Incorporating Generative Adversarial Networks (GANs), our approach features a generator network tasked with mapping grayscale images to the a and b channels, mirroring the functionality of the CNN model. Concurrently, a discriminator network is employed to differentiate between generated images and real images, facilitating the training of the generator to generate authentic colorized images.

Layer	Filters	Activation	Regularizer
Conv2D	64	ReLU	L2
Conv2D	64	ReLU	L2
Batch Normalization	None	None	None
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Batch Normalization	None	None	None
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Batch Normalization	None	None	None
Conv2D x 3	512	ReLU	L2
Conv2D x 3	512	ReLU	L2
Conv2D x 3	512	ReLU	L2
Batch Normalization x 3	None	None	None
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Batch Normalization	None	None	None
Up Sampling 2D	None	None	None
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Batch Normalization	None	None	None
Conv2D	313	Softmax	None

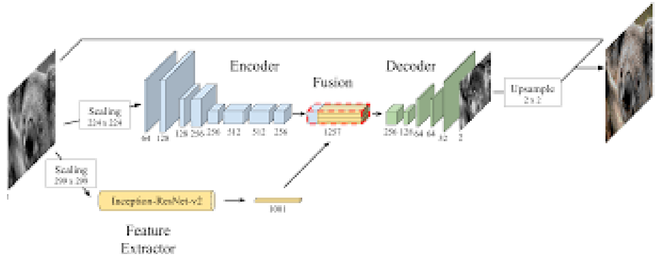
Architecture

Baseline CNN Model:

Architecture
Baseline CNN Model:
The custom CNN architecture involves an Encoder that processes the L channel of the Lab color space, extracting features and downsampling to reduce computation. The Decoder then upsamples and applies convolutional layers to output the a and b channels. We use the tanh activation function to map output values between -1 and 1.



Inception-ResNetv2 based CNN:
This model utilizes a pre-trained Inception-ResNetv2 model as a feature extractor. The Encoder part is similar to the Baseline CNN, while the fusion layer combines features from the encoder and the pre-trained model. The Decoder then generates the final a and b channels.



Pix2pix GAN:
The Pix2pix GAN employs a modified U-net generator with skip connections, utilizing pre-trained models like MobilenetV2 and Densenet121 in the encoder part for feature extraction. The discriminator network distinguishes between real and generated images. The loss function includes modified binary cross-entropy loss and L1 loss, ensuring the generator produces images similar to real targets.

Dataset & Preprocessing
The dataset comprises nature and landscape images, filtered to remove unwanted images. After preprocessing, the dataset consists of RGB images resized to 256x256 and normalized to Lab color space. The L channel is the input, whereas the a and b channels represent the target output.

Training Details
We trained all models using TensorFlow on 90% of the dataset, with various optimizers and learning rates tailored to each model. The CNN models were trained for 150 epochs, while the Pix2pix models were trained for 100 epochs. Batch sizes and optimization parameters were adjusted accordingly.

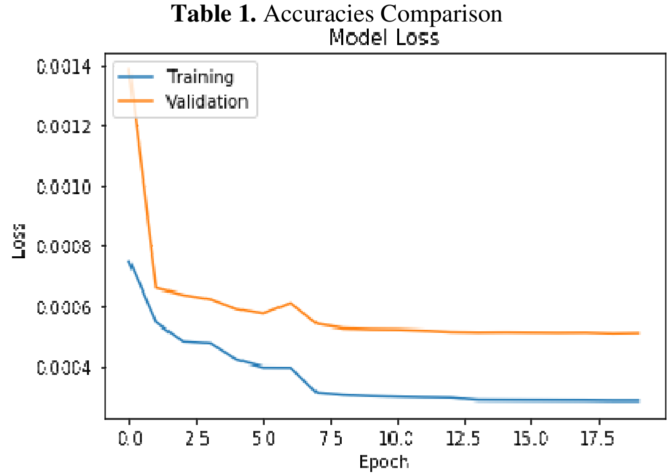
5 Results

The model described in section 4 was trained on a dataset consisting of 8000 randomly selected images from the ImageNet dataset. This dataset was split into 60% for training and 40% for testing purposes. To evaluate the model's performance, we utilized categorical cross-entropy as the loss function, optimized using the StochastiGradient Descent (SGD) optimizer.

To assess the final colorization results, we employed metrics such as Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR). These metrics provide insights into the accuracy and quality of the colorized images generated by the model.

Figure 3 illustrates the model's loss across 18 epochs during training and validation, offering a visual representation of the model's learning process and convergence over time.

Table 1. Accuracies Comparison



6 Conclusion& Future Work

In this study, we explored the realm of automatic image colorization using deep learning methodologies. Our findings underscore the significance of this approach in saving time and reducing human effort, particularly in scenarios where user-based colorization proves to be laborious, especially with complex scenes and numerous objects.

Through the implementation of four distinct deep learning models, including CNN and GAN-based architectures, we gained valuable insights. Notably, our simpler CNN model showcased impressive performance, surpassing larger and more complex models like Inception-resnetv2, especially when dealing with high-level image components such as skies, mountains, and forests. However, the limitations of our dataset's simplicity were apparent in the underperformance of sophisticated models like Inception-resnetv2, despite their potential demonstrated in prior studies.

On the other hand, our Pix2Pix GAN models exhibited superior colorization results, delivering vibrant and nearly artifact-free outputs. Leveraging pre-trained models such as MobileNetv2 and Densenet121 enhanced our ability to create high-performance models within resource constraints. Among these, the Pix2Pix model with MobileNetv2 emerged as a highly effective and compact solution, suitable for deployment even on mobile devices.

While our models excel in certain scenarios, their generalization power is limited due to the dataset's restrictions. To enhance performance and address challenges in colorizing small objects and fine details, future endeavors should focus on expanding the dataset's diversity and complexity. This expansion would enable our models to learn a broader spectrum of color schemes, thereby enhancing their generalization capabilities. Additionally, exploring more sophisticated models could mitigate existing limitations and further refine the colorization process.

References

1. Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. "Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification". ACM Transaction on Graphics (Proc. of SIGGRAPH), 35(4):110, 2016.
2. R. Zhang, P. Isola, and A. A. Efros, “Colorful image colorization,” in ECCV, 2016.
3. Baldassarre, F., Morín, D. G., & Rodés-Guirao, L. (2017). Deep Koalarization: Image Colorization using CNNs and Inception-ResNet-v2. CoRR, abs/1712.03400. Retrieved from http://arxiv.org/abs/1712.03400.
4. Jeff Hwang and You Zhou, “Image Colorization with Deep Convolutional Neural Networks”, Stanford University.
5. David Futschik, “Colorization of black-and-white images using deep neural networks”, January 2018.
6. Alex Avery and Dhruv Amin, “Image Colorization”.CS230 – Winter 2018.
7. G. Larsson, M. Maire, and G. Shakhnarovich, “Learning representations for automatic colorization", CoRR, vol. abs/1603.06668, 2016.
8. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in Proceedings of the IEEE Conference on computer vision and pattern recognition, 2017, pp. 1125–1134.
9. K. Nazeri, E. Ng, and M. Ebrahimi, “Image colorization using generative adversarial networks,” in International Conference on Articulated Motion and Deformable Objects. Springer, 2018, pp. 85–94.
10. Levin, A., D. Lischinski, and Y. Weiss. Colorization using optimization. in ACM transactions on graphics (tog). 2004. ACM.

