iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.66096

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

SAR Image Colorization with a Deep Learning Model for All-Inclusion Understanding (H)

M. Jaya Lakshmi[1], C. Ahalya[2], B. Shubha Deepika[3], M. Manisha[4], J. Pallavi[5]
[1, 2, 3, 4, 5]*Electronic And Communication Engineering, Ravindra College of Engineering for Women*

Abstract: ***The procedure for enhancing halftone photographs with color to create eye-catching pigmented pictures is known as image colorization. This involves the application of colors in line with the intended purpose of the image. However, traditional methods of colorization available today often include the need for end user's intervention in the form providing color points and doodle drawings, as well as use of color images for reference in which case colors are transferred to the target image or using multiple color images to predict what the colored result would look like; however, the majority of colorizes developed tend to produce colorized images that are not realistic in most instances due to unskilled users who attempt to apply the devices, wrong color transfer or limited color image library. In this paper, to overcome these weaknesses, a new method that consists of two modules is proposed to colorize certain regions of an image by fusing semantic segmentation and seamless region filling. In the first module, input image foreground and background regions masks and categories are extracted each extracted region's reference image is taken from a per-sorted color image database using a Mask R-CNN model. Using colorization approaches based on the VGG and U-Net models, respectively, the second module entails coloring the image's backdrop and other areas. Afterward, the final, fully colored image is created by combining the photographs using the poission editing process. The tests show that our method enables not only the proper reference image selection based on the semantic information of various image components, but also incorporation of the colored results to create a believable colored image. Proposing to use different CNN-based models in a combination where each is used according to its own strength, In contrast to other methods now in use, our scheme gives a deeper creative visual impression while avoiding the disadvantage of single-step techniques' failure frequency.***

Keywords: ***Mask R-CNN, Input Photograph, Image Sections, Coloured Photographs, Conceptual Data, Visuals, Black and White Images, Semantic Pix elation, Reference Picture, Base Image, Images Stockpile, Outlying Areas, Image Library, Segregation Methods, Doodle, Conventional Neural Network Based Designs, U-Net Architecture, One-shot Approach, Strategy Present in This Manuscript, Zonal Regions, Principal Zones, Large Image, Color space, Competitive Half toning Algorithm, Zonal Context, Focus Area, Artistic Incorporation, Partially Automatic Process, Cutting Out Section of the Picture, Instance Level Segmentation.***

I. INTRODUCTION

Image Colorization can be referred to as an artistic style in burning the colors to the images such as images in color, bones, black-and-white photos, ink paintings, and so on in order to create the enhanced version of the image which is fulfilling and vibrant in a room [1], [2]. Colorization was first brought into idea by Wilson Markle in the year of 1970 where he outlined the process of introducing color to televisions and films that are in black and white [3]. Presently there is an upsurge in the demand for colorization in the fields of graphic design, animation, film making, art font rendering, and many other creative endeavors which have not been enjoyed in the past. In basic computer vision, coloring pictures is recognized as an ill-posed problem. This is because most black regions in the input image can be visualized in many possible color variants according to one's imagination and culture [1]. Therefore, a colorization solution cannot be correct on its own. The initial stages of colorization are very painstaking as every work is done manually redesigned. This is an elaborate, costly and lengthy procedure. Hence many techniques of computer-aided design have come up to ease the cognoactivity stages of the coloring processes among them techniques that rely on user cues, transfers, auto-prediction, and hybrid approaches [2]. These currently available practices yield excellent performance characteristics; however, they have some shortcomings. The user-hints based technique is the most troublesome as it requires a great deal of labour-intensive artistic scribbles; the transfer-based method needs very vibrant and appealing pictures for the process to work, but such pictures are not easy to come by; the auto-prediction method systems learn from a very comprehensive database of images and will colour images, but that will only be images within a training database. It seems that instead of making a better colorization algorithm, hybrid methods which combine several single-step techniques decided to obtain better results. However, this is not satisfactory either since combining different single methods without any consideration of how they relate to the main regions in an image gives sub optimal results for many cases due to poor aesthetics. Hence, there is a need for a better hybrid scheme which will counter the inefficiencies of eastern and western modes of practice.

©IJRASET: All Rights are Reserved

Incorporating in addition to picture semantic segmentation and fusion techniques, we present a hybrid method for regional coloring of images that enables users to apply different colors to different areas of the target image, potentially producing a more diversified artistic color impact.

The region segmentation and colorization modules make up the majority of the suggested system. First, a well-trained CNN network, Mask R-CNN, is used in the segmentation module to extract the areas and background masks and categories from the input gray picture. Next, region reference images are either manually selected or randomly picked out from a sorted images database in context of the category of every region. Second, a U-Net model is used for the automatic colorization of defined background regions, and a VGG model is employed to fill each distinct target region with color using appropriate images. Ultimately, Poisson blending is used to smoothly integrate the colored backdrop and areas to create an appealing and satisfying colorful image.

Our approach can give different target areas, their unique natural semantic colour, with careful selection of the comparable reference photos from the classified database that is now accessible, semantic colour using just one reference image. Aiming at the background colorization and image region colorization respectively, how two different CNN based colorization techniques are fused is presented in a way that single step colorization method is avoided which is highly prone to failure. The experiments' findings show that the suggested approach outperformed other approaches in terms of visual appeal. The remainder of the document is organized as follows. A survey of the literature on image colorization techniques is given in Section 2. In Section 3, an advanced hybrid approach for colorizing regions of an image is detailed. In Section 4, performance assessment of the suggested approach is completed. Finally, in Section 5, a summary is provided along with the outlining of the future works.

II. ASSOCIATED RESEARCH

Classic reproduction with colors is good for drawing pictures, retouching photographs, or even doing films but is to a large extent, also very boring, engrossing and costly [4]. Hence, Numerous Computing approaches have come up for rapid and efficient colorization which can be subdivided further along different parameters into various types for easier understanding and application.

Depending on how much the user is involved, colorization techniques are divided into three categories: automatic, manually operated, and user-guided [4]. The automatic methods use large-scale image dataset trained CNN networks to colorize black and white images. The semi automatic techniques involve mainly copying hues from a few pictures to the black and white picture. The user-guided ones allow placing of colors in certain areas of image based on the directions of the users.

Depending on the source of color information, one could distinguish between hybrid pigmentation, auto-prediction, transfer-based, and user-hint-based methods. The company mentioned that these will be discussed below.

A. Techniques Based on User Suggestions

The work of Levin et al. in 2004 [5] was the first to apply this approach that mainly centers on the diffusion of user input made in the forms of colored points and strokes [1]. Such methods are based on the common assumption that the neighboring Normalized cuts are an effective way to fix pixels with similar brightness because they will probably have similar hues [6]. For further enhancement of the quality of the pictures, the colorized images can be additionally and interactively improved with the users' extra strokes [7]. Huang et al. describe a method of dynamic edge extraction to speed up optimization and bleeding for coloring [8]. Yatziv and Sapiro explain how to use extended zones defined by other colors to "blend" colors over a picture [9]. Luan et al. address this problem by taking into consideration both nearby pixels of the same color as well as faraway pixels with the same texture in order to improve the quality of coloring when few strokes are given by users [10]. Xu et al. apply a method where each pixel is assigned a color according to the color distribution of other pixels which is weighted by the distribution of strokes [11]. The above helpers are very interactive but users themselves have to specify where and how colors and strokes are applied and intense practice is still needed to get a decent colorized image due to long processing time [4],[12].

B. Approaches Based on Transfer Learning

Transfer-based methods fully highlight the picture using the colors of a chosen reference image or pictures that are comparable to the black and white image in order to reduce the amount of work required by the user [12]. Local descriptors help to derive the correspondence comparing reference and input pictures naturally, or it can be done manually with some human intervention. The first approach proposed by Welsh et al. transfers color information by utilizing worldwide pigment data. In many cases, this approach is unsatisfactory because it ignores the spatial placement of pixels [1], [13]. To enhance the color transfer techniques, various techniques for correspondence are proposed which are encompassing the pixel connection approaches, super-pixel level connection, and segmentation area level correspondence [1], [14]– [17].

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Chia and others use the Internet to easily locate an appropriate reference image for colorization. Given a user's segmentation of main objects within a scene, and appealing semantic text labels, numerous pictures can be collected from image banks and appropriate ones for colorization of a supplied gray-scale image, those that are useful for colorization of the pictures by applying certain reference images, can be chosen [16]. Morimoto et al. introduced an automatic coloring technique based on web-image information, which takes numerous images from the web and performs intelligent imaging and coloring of the grayscale images [17]. Two more recent works use characteristics from a VGG network that has already been trained while performing colorizing images [18], [19].

He et al. trains a CNN's avatar to perform direct image translation Gray-Semantic input to a Colorized output image by referring to better suited Recoloring image. An image retrieval algorithm is described which helps in recommending the references without much manual effort, by utilizing the semantic and luminance information [1]. Luan et al. create a VGG-based scheme for photography style transfer that enables the color changes of the reference picture to the target image while maintaining the pixel-by-pixel information of the reference and target pictures. This however prevents the distortion and provides clear vivid images under multiple photogenic conditions which is otherwise very challenging to achieve [2]. Nevertheless, the methodologies which rely on transferring the art do require all the selected reference images to also contain similar using the supplied monochromatic pictures for scene terminology, which is still time-consuming and a hindrance to the users [12], [22]. In reality, it is very difficult to find a reference image that approximates the original gray-scale image, resulting in unsatisfactory colorized images.

C. Approaches of Self Prediction

The auto-prediction approaches focus solely on the notion of learning to color from large samples of colored images. The majority of colorization systems are developed on the basis of enormous databases in a completely unsupervised fashion [3]. Deshpande et al. formulate the process examine colorization as a linear system and discover its parameters [23]. Cheng et al. take the approach of joining many built-in image features which are the input of a three layer's dense brain network. Another approach was taken by in order to simulate per-pixel color histograms, Larsson et al. developed a deep network that can serve as an auxiliary entity in the network dedicated to color image creation or can be used as a middle step before generating an image. Other techniques for end-to-end planning Construct several CNN-based models with matching loss functions in order to gather information and forecast color outcomes. Iizuka et al. display a sophisticated structure with a fusion layer that allows for the integration of both local and global picture characteristics, and it is trained end-to-end using L2 loss to yield visually pleasing results [12]. Zhang et al. mention a solution to the multi-modal colorization issue by training a feed-forward convolutional neural network with the pertinent absence function (classification loss) [25]. Isola et al. explain a scenario in which an L1+GAN loss is used to train a U-Net based GAN network for image-to-image translation. The network's output may include image colorization, picture construction from edge maps, and image generation from intermediate label maps [26]. Whereas many of the auto prediction approaches simply output a single sensible image to an input given to it since, again, colorization is simply an open-ended task of generating more than one image given some input [1]. While they are aided by training images in the database as and when required for the maximum possible dispensation of their space, the dangled effects cannot create any other color to luminance ratio variation that is not covered in the database.

D. Hybrid Approaches

Hybrid approaches integrate some or fully the above-mentioned techniques in order to attain satisfying colorization effects [1]. The hybrid frameworks put out by Zhang et al. [27] and Sangkloy et al. [28] inherit management from the user hints-based method and resilience from the auto-prediction technique, accordingly. From this, it is considerate to include in the solutions portion of the paper. Zhang et al suggested a colorization framework based on U-Net architecture, which takes user-supplied color points and a source picture in monochromatic, which helps in reducing users' involvement as well as enhancing the efficiency of pigmentation [27]. The deep adversarial architecture for image synthesis is provided by Sangkloy et al. which is capable of producing objects like a car or a face in the region bounded by the sketch and color patterns [28]. In his research, Cheng et al. apply a post processing step with a joint bilateral filtering technique. He further proposes an image clustering method that is adaptive to the image and exploits global picture data that is utilized to determine a three-layer CNN pigmentation model's weights [24]. Many test results show that this technique achieves better quickness and quality, but the color looks dimmer with some dullness. Although these hybrid approaches which take the benefits of different one-shot techniques tend to enhance the overall outlook of the colored images, they still possess some limitations of the one-step approaches. In particular, they colorize the image without regard to the regional semantics, which is also very far from ideal performance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Analysis and Problems

Having outlined the strengths and limitations of the existing systems, it is clear that the single-step colorization techniques have severe efficiency limitations and the hybrid systems in existence have their own problems too. In just over the last several years, notable advancements have been made in design and picture analysis graphical techniques as well as in other aspects such as audio input systems, whereby speech can be recognized and reproduced vividly as imagined.

Nevertheless, the current methods of colorization have been developed by incorporating various models based on CNN because of its ease of application of deep learning. However, because of the variations the purpose and effectiveness of various CNN models, the network designs, and the objective loss function used vary as well. In other words, CNN algorithms cannot be said to be general-purpose algorithms since it is impossible for one CNN model to satisfy all the colorization task requirements.

Hence it is imperative to appreciate the merits and demerits of various colorization methods so as to rationally integrate and improve upon a range of current CNN-based designs and to create an automated hybrid pigmentation solution of a region in an image which is not subject to the shortcomings of existing techniques, and is able to source reference images better suited for the principal portions of the input black and white image, thereby enhancing the natural richness and variety of the resultant colorization.

III. PROPOSED SCHEME

For the region pigmentation of auto/semi-auto images, an enhanced hybrid system comprising segmentation and colorization modules is suggested. As seen in Figure 1, the scheme is a combination of regions segmentation and colorization-based methods. The colorization module works with a gray input image and the emphasis on the pertinent elements of the image is achieved through the image is masked in the Mask R-CNN method, and corresponding categories are provided. In accordance with the disbursed area of the area, the appropriate reference picture is selected manually or at random from a library of colored images that have already been identified. The coloring generator uses a VGG-based transfer system using the reference pictures chosen in the first module to color the main parts of the pics, while the U-Net-based auto-prediction model colorizes the image backdrop naturally.

Lastly, the picture fusion method optimally fuses the colored background and regions to produce a coherent and visually pleasing colored image. The specification of the scheme is given in the sections that follow.

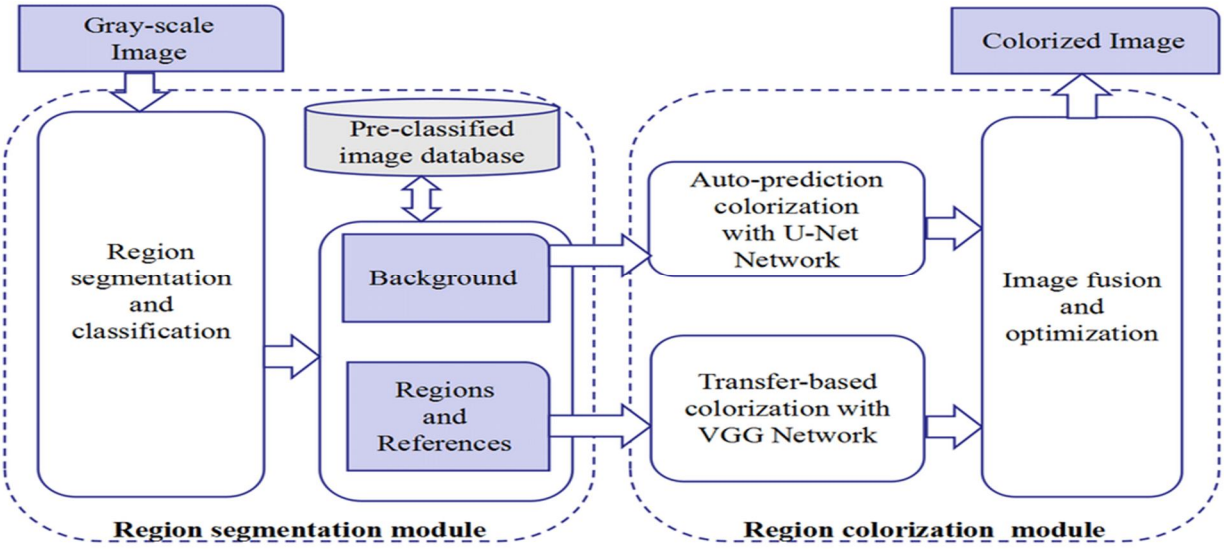


FIGURE 1.A hybrid image colorization system designed for regional colorization of an image.

A. Module of Region Segmentation

The concept behind object recognition involves localizing and categorizing the target objects as seen in the image through the principles of pattern detection and processing of pictures. In other words, Semantic division is an advanced application than object recognition where the goal is the labeling and separation of all detected pixels that belong to the same category of objects in an image. This constitutes a fundamental technology for content-concentric multimedia applications. In other words, the term instance segmentation demands more detailed segmentation of identical objects in comparison to the partitioning of semantics. In this study, we concentrate on dividing instances for a specific purpose, and it is to demarcate the distinct sub-regions in an image along with their respective classes.

1) Segmentation Classification de Regions

The structure designed for instance segmentation is called Mask R-CNN and it was proposed by He et al in the year 2016 [29]. It is a neural network-built end to end and can perform several tasks by incorporating other categories that include, between others, target grouping, recognizing objects, posture recognition, and semantic segmentation. This architecture is also incredibly flexible, efficient, and extendable.

Figure 2 displays that Mask R-CNN contains three principal components [30] and an additional spam filtering component which we incorporated in order to restrict masks of the primary regions. The fewest feature extraction is a foundational network called Res Net-FPN, which combines the feature pyramid network FPN with the multi layer residual neural network Res Net. In other words, the feature maps are the input photos after the backbone network are used in this component. The second component is called the region-proposal network (RPN), a thin neural network, examines the features of the photo's feature maps looking for the target area proposals with a sliding window technique. The last part is made up of a mask sector, the boundaries of the regression, and a ROI classifier. The regions of interest (ROIs) are classified into one of the predefined classes (like human, vehicle, chair, etc.) and also into the surrounding class using the ROI classifier that the region proposal network (RPN) has produced. To accurately modify the size and location of the border box enclosing the area of interest, the border box must be filled in. A neural network is used in the mask branch. Using the area regarded as positive the ROI classifier can be used as input to construct a mask of each item.

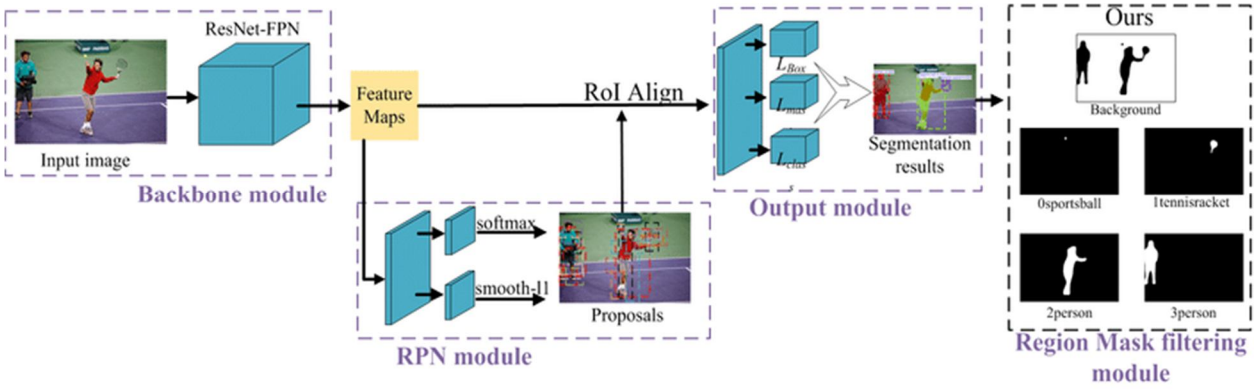


Figure 2. The updated MASK R-CNN method framework.

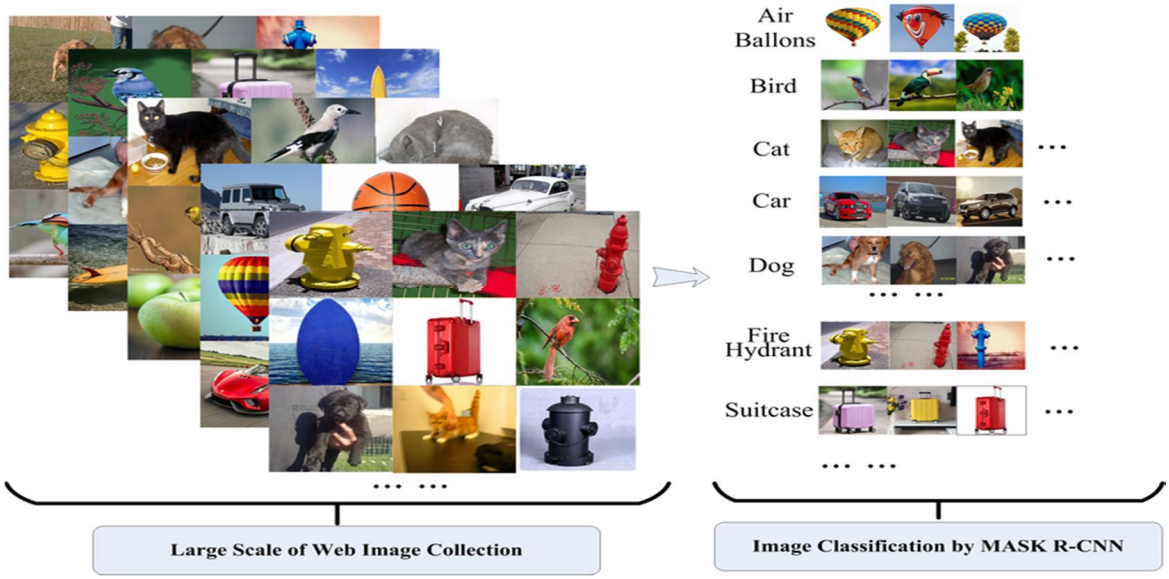


DIAGRAM 3. A database of visuals with a predilection to a particular class.

Multiple regional targets may be present in a picture; only the backdrop, the mask, and the primary target category are necessary. In order to filter the masks of the specific key regions in accordance with the appropriate rules, a portion of the region mask filtering part is created.

The anticipated results are given below.

- Covers of each region of the significant instance.
- N+1 groups, where every instance region has its corresponding class denoted within the range of 0 to N, whereas N+1 is the background class.

The filter rules proposed can be summarized in this order.
A) The masks are arranged according to target location and type.
B) A threshold limit for the screening of the mask image is formulated. Masks corresponding to regions whose areas are greater than the threshold may be output, otherwise, it belongs to the backdrop.

2) Establishing a Collection of Pre-classified Pics

Building the database of pre-classified pics facilitates the selection of a reference image for each area in the database. This database includes images from the web on a large scale collected using crawling software. These images need to be classified using the MASK R-CNN technique and later put in this database.
Based on the classification of geographical area, an either manually or at random, the relevant reference the picture is selected from the relevant annotated database. This helps to ensure that each geographical area has its own reference image color wise and also the images used are color toned with what is expected in reality.

B. Mechanism for Colorization by Region

Two distinct methods are used in this section to color the categorized foreground and the primary areas of a given grayscale image. These methods are shown below.

1) U-Net-Based Auto-Prediction System

Here, the U-Net structure for splitting images, which Zhang et al. [27] implemented, is modified for the task of colorizing the given input background as it has shown to be very effective in a number of conditional generation jobs that provide access to valuable information through skip connections at higher layers while still retaining lower level details.
As depicted in Fig. 4 shows the U-Net network's construction, which consists of 10 multi-layer blocks designated conv1 through conv10.The properties of the tensors in the inversion 1–4 portions are spatially lowered by a factor of 2 while at the same time the feature dimension is increased by a factor of 2; each block consists of two to three pairings of conv-relu. While the feature dimensions are coarsened, the pattern spatial clarity is restored in the conv7 through conv10 blocks. In turn, in paragraphs conv5 and conv6, the spatial resolution does not decrease, instead, diluted convolutions serve to substitute for the decrease in quality. The purpose of the bidirectional bypass links is to make the network's ability to retain spatial features and provide easy reach for the essential lower-level details housed in the subsequent sections [27].

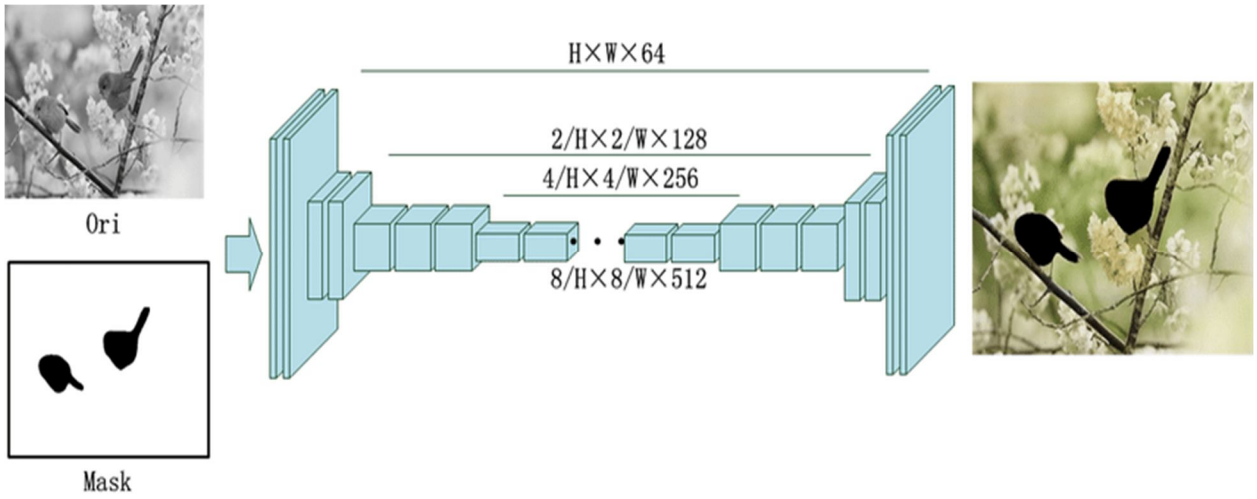


FIGURE 4. An illustration of the architecture used in auto background color prediction using a U-Net.

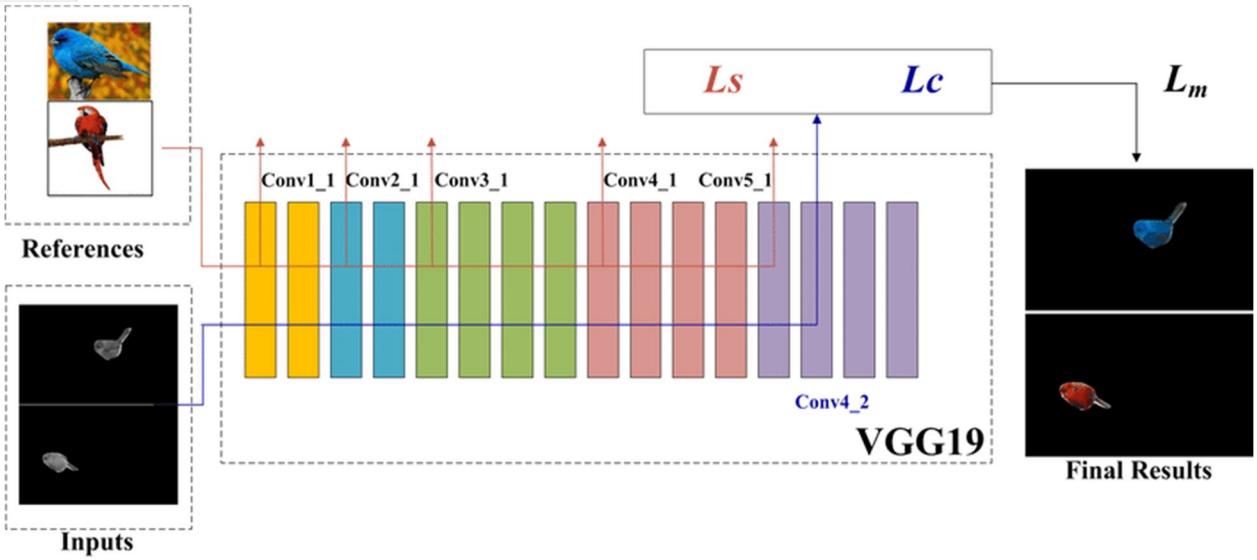


FIGURE 5. The procedure of colorization through transfer technique.

The main purpose of training the U-Net network is to determine the pixel-by-pixel correlation between a colored image represented by θ and a picture in grayscale and perform optimization on the objective function (1) so as to achieve an optimal value of the element θ*. D denotes a group of a set of black and white images, coloring requests from users, and the corresponding-colored images. X is the initial black and white visual with an extra user tensor input U; M is the mask for the foreground; Y is the photo with an underlying-colored Y; The loss that L has described is the difference between the actual-colored image as provided by the network and the desired image.

$$\theta^* = arg\ min\ E_{X,U,Y \to D}[L(F(X,M,U,\theta),Y)] \qquad (1)$$

θ

2) Transfer Learning Based Model using VGG

To add shade to some areas of the image while maintaining its textured qualities, the color transfer technique that was developed by Luan et al. [2] is utilized. This technique is able to incorporate the shade contents of the source image into the relevant subject picture through the use of the Matting Laplacian image to minimize the deformation of the surfaces in the content graphics during style transfer.

The above procedure is used in the YIQ color environment, in which chrominance element is signified by I , Q standards and its brightness factor is denoted by Y. Initially regions of the picture and their respective mask images are fed to the learned network VGG_{19} and L_{total} which is the final value of loss that has to be minimized by the network is calculated on the equation (2), the network is then optimized. Eventually in this model a transferred color image is achieved by this process in the course of user defined number of iterations until a transferring threshold is reached.

In the equation below, L_c refers to the content loss obtained from conv4 2 layer; L_s is the style loss obtained from layers conv1 1, conv2 1, conv3 1, conv4 1, and conv5 1 of VGG_{19}; L_m is the additional photorealism regularization loss integrated within the model, It seeks to restrict the textural changes of the content images; additionally, α, β, and γ are user-defined graded factors that can be adjusted to fine-tune a final shipment effects.

$$L_{total} = \alpha L_s + \beta L_c + \gamma L_m \qquad (2)$$

3) Amalgamation Of Sections of Image

The section of image colorization and optimizing incorporates the colored background and every area to form the last-colored image without any distortions. However, such an approach where all the colored pictures are simply combined would end up looking artificial as the boundary edges of all these regions would be distinct and thus, a jarring contrast would exist between the colored regions and the surrounding background.

Poisson editing [31] is a fusion optimization approach that successfully places the target images into the backdrop image smoothly. Here, a user-defined visual set of boundary conditions is used to solve a Poisson equation. Additional features like texture smoothing, editing local lighting changes, local color changes, and so forth can also be obtained with this technique. In order to achieve a final colorized image with smooth edge blending and well-combined regional colors, we use Poisson editing to deal with the background and the colorized sections. The following are the algorithm's main steps.

a) Find the gradient fields for the N target region pictures and the background image, accordingly;

b) Using their category information, apply the image dilation technique to the N regions masks. Each mask's intricate operations are carried out as follows: Every pixel in the mask image is scanned using a structural element of size, which is a 3x3 matrix, and all of its element values are set to 1 by default. The computer then performs a "and" operation between the structure element itself and the mask image beneath it; if all are 0, the mask image pixel is 0; if not, it is 1. Consequently, a single circular band enlarges the mask image.

c) The N dilated masks are used to merge the pictures of the target region with the gradient field of the source background image.

d) The gradient field of an image that requires reconstruction is obtained by calculating the gradient value for each pixel point. The partial derivative of the gradient is then computed to determine divergence.

e) With b representing the divergence calculated in the third step above, A the coefficient matrix, and X solved to obtain the R, G, and B values of each pixel of the fused colorized image as x, this Poisson equation can be expressed as follows.

$$Ax = b \quad (3)$$

IV. ASSESSMENT

A. *Data Collections and Environment*

This work employs the MS-COCO [32] dataset for training the instance segmentation models and the Image Net [33] dataset for training the image colorization model. On account of limited real-world images available for testing various colorization algorithms' outcomes, the source provided input grayscale images for the colorization model's training have, in fact, been obtained from their paired color images, which are the anticipated results and ground truth respectively. The experiments are based on a workstation running CentOS $_{7.0}$ and equipped with Open CV and Tensor flow software with deep learning capabilities and an Nvidia P_{40} GPU with dual Xeon $E_{5\text{-}6280}$ processor hardware.

B. *Assessment of Gray Level Image Segmentation Performance of Mask R-CNN*

Two distinct models, the color-based model and the gray-based model, have been trained on a color image dataset and a gray image dataset, respectively, to evaluate the effect of the Mask R-CNN model's segmentation on grayscale pictures. As seen in Figure 6, their ability to segment a color image with contents similar to the gray cutout is evaluated.

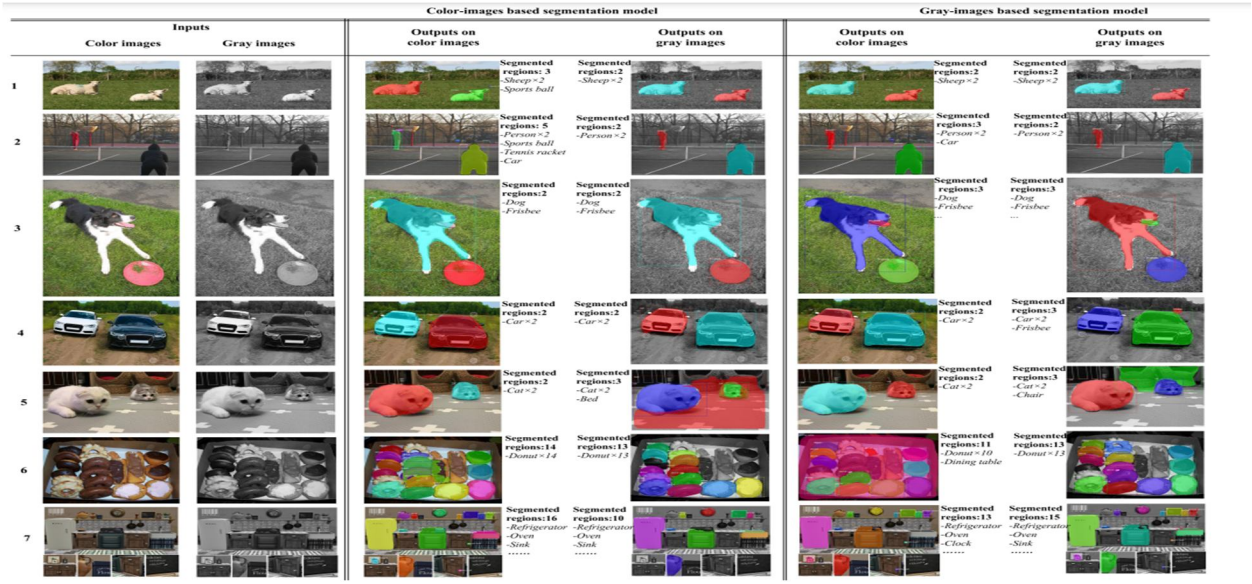


Figure 6: Mask R-CNN's segmentation and masking effects on color and grayscale pictures.

The study's findings demonstrate that both algorithms can correctly pinpoint the main target regions of the provided photos. A greater portion of the color image than the gray image can be obscured by the color model, particularly small areas and targets that are "hidden" in cluttered backgrounds. The gray model is better at clearly segmenting gray images because it learns to extract more pertinent information from them. The gray model was used for this study because of this.

C. Evaluation Of the Image Impact of The Current Proposal

This subsection evaluates the effectiveness of the schemes that we have proposed. There are four distinct schemes based on the processing of the input and output in our article. The scheme 1 operates on one reference image only and critically without optimizing the edges; the scheme 2 operates also on one reference image but optimizing the edges; the scheme 3 works on more than one reference images and does not optimize the edges; the scheme 4 works on more than one reference images but optimizes the edges. Figure 7 illustrates the filtering principle where the primary region masks produced by Mask R-CNN are picked for additional colorization. Scheme 1 and scheme 2 use a reference image and apply the same coloring process on all target areas; whereas the scheme 3 and scheme 4 are capable of identifying selective reference images from the idea bank built on each region's semantics, and colorizing the respective area of the gray image with the appropriate reference. And, the resulting images are more colorful and more varieties of subjects are present.

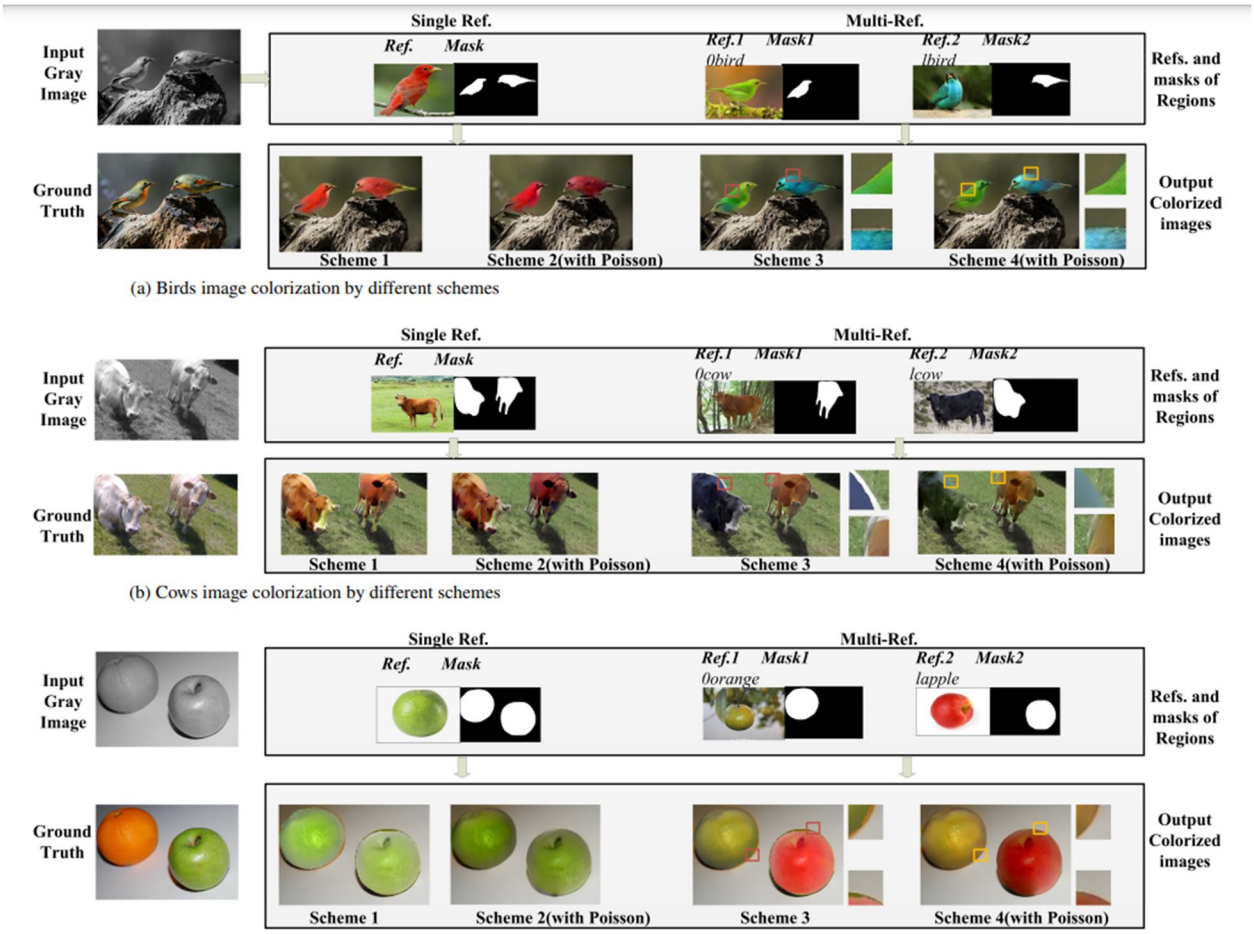


FIGURE 7. Condensed visual impact assessment of the proposed approach under varying conditions.

The regions' boundaries are not clearly defined, creating a sharp edge gap and a difficult transition between the regions, even though the third scheme is better than the first and has the most appealing visual appeal because of the rich colors used. The Poisson editing technique is used optimally in Scheme 4 in a similar way to Scheme 2, not only on the edges of the corresponding regions but also on their brightness and color. This ensures that all of the colored regions have a smooth transition with the colored background, making the final image more aesthetically pleasing with richer colors. For this reason, scheme 4 is chosen.in order to add color to this work.

D. Evaluation of Performance Against Current Systems

This section assesses the suggested scheme's performance in comparison to other schemes. It is important to take note of Figure 8, which demonstrates that while auto-prediction techniques can accurately forecast the general color of a given grayscale image, in certain cases, the colors are limited by the training datasets. Specifically, the aesthetic appeal lacks vibrancy and color.



FIGURE 8. Comparison of performance of Various Existing Approaches (Rows 1 and 10 represent input and ground truth, respectively; the fourth row's reference image is utilized by the approaches presented in paper [1] and paper [2]. Row 8 is extra reference images incorporated in our method on top of the previous one to provide colorizing the different parts).

The transfer-based approach can be used to add color to black and white pictures. Utilizing a single reference image, but it neglects the inherent semantics of the regions within the image. Consequently, the newly introduced colors in the primary areas of the image exhibit the same shades, and some of them even spill over the edges of the area. Drawing on the strengths of both the auto-prediction approach and transfer-based techniques, the design provides for the automatic or manual retrieval of relevant reference pictures from a big database of images for every place based on the specific semantics of various geographical areas. In order to achieve wider and more complicated backgrounds, auto-prediction method helps to achieve color prediction. Results achieved after coloring with our method are stable in quality, contain a variety of colors, and combine different picture regions naturally.

One of the artistic creation processes, by adding color to a gray picture colorization is another process which does not have a unique ground truth. This is because the process depends on the imagination of the person doing the work. Different people are likely to have varying opinions if not completely opposing opinions on similar colorization results. How then should the visual outcomes of such a algorithm for coloring be understood remains a significant question. Generally speaking, qualitative assessment (QA) is predominantly relied on in assessing the coloration effects. The evaluation in QA is done in the form of questionnaire which warrants participants to assess the outputs of various algorithms which is observed by the participants. But the evaluation results could differ depending on the status of the participants (age, occupation, etc). There is a methodical inconsistency inherent in the QA technique however this technique can at the very least give some insight of the audience's liking towards color blending.

This work employs an Internet-based questionnaire survey of users from various demographics. A total of thirty people participated in our survey, most of who consisted of university attendees and educators. The questionnaire's contents were to arrange the impact of coloring of all the input gray scale images shown in Fig. 8. Ultimately, the qualitative mean rank of the performance of different methods is arrived at by ordering the varying colorization effects of a single image. It is noted in Table 1 that the suggested scheme provides a better rating in terms of human visual perception than any other method discussed in literature.

TABLE 1: The Mean Rank of Reported Effects

Method	Classification categories	Average rankings
[7]	Auto-prediction	4.28
[12]		3.57
[2]	Transfer-based	5.28
[1]		2.26
[27]	Hybrid	4
Our	**Hybrid**	**1.29**

When applied to color prediction, our system produces rich hues, steady quality, and organic fusing of image regions in visual effects.

TABLE 2 Existing Method Comparisons

Scheme	Classification	Network architecture	Datasets	Interactivity	Large datasets	Region coloring	Multi refs.
[7]	Auto-prediction	- VGG+hypercolumn descriptor	ImageNet; ctest10k	x	✓	x	x
[12]		- Low-Level Features network; - Global Features Network; - Mid-Level Features Network; - Colorization Network	Places	x	✓	x	x
[2]	Transfer-based	- VGG+Photorealism regularization	-	✓	x	x	x
[1]		- Similarity sub-net(VGG); - Colorization sub-net(U-Net);	ImageNet	✓	x	x	x
[27]	Hybrid	- Main colorization network layers (U-Net); - Local Hints Network; - Global Hints Network	ImageNet	✓	✓	✓	x
Our		- Mask R-CNN; Poisson editing; - Net for Background; - VGG for regions	CoCo ImageNet;	✓	✓	✓	✓

Table 2 provides a detailed comparison of various existing techniques. The majority of these are developed using the standard Image Net images, and employing sophisticated deep learning techniques like VGG for image feature extraction. Auto-predication methods are functional without the intervention of any user in its operations, and also have to undergo a learning process of color mapping from grayscale pictures to colored ones; thus, they necessitate large datasets. The techniques presented in [1], [2], [7] various key areas of a picture with multiple reference photos. Further necessitate the use of an extensive collection of images since their approach involves the picking of appropriate reference images to be used for the input images. In particular, our methodology can produce animated images as these can be obtained by collage of key areas of a picture with multiple reference photos.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. CONCLUSION

The challenge of coloring images is to appropriately infuse lifeless shades of greyscale/monochrome images, photos, videos, and films, artistic work becomes richer in content and more engaging. Typically, this problem is not a well-posed one as it consists of painting colors in an RGB projected image, wherein the whole image is described by only luminance values, and often requires a lot of manual correction to achieve realistic quality standards.

There have been developed computer-assisted technologies for alleviating the colorization task as well as speeding it up. However, they usually come with certain limitations. These approaches usually entail color scribbling over the target image, defining a set of target reference images or having a rich image database. Colorized images may also seem unnatural and unrealistic due to poorly developed colorization skills of the user, poor color application, or small image databases. All existing hybrid approaches which integrate different single step techniques work on the premise that images do not have any regional semantics, which is an inherent limitation since images are often colorized according to regional details. There is need to come up with an enhanced hybrid approach for region-based image colorization.

In this paper, a hybrid technique for image region colorization is proposed with a view to surpassing the above limitations. This scheme allows users to add color to several target areas of the greyscale picture with many reference pictures based on details about the regions category while eliminating the shortcomings of one-stepped methods and hence, the arrangement of picture hues can be executed logically. The suggested plan has two basic components: modules for colorization and segmentation. In the initial module, the back ground and target areas of the given monochrome picture are segmented utilizing the R-CNN model for Mask.In line with the classification of primary areas, suitable pictures pertaining to every area are chosen either manually or at random from a pre-categorized images database. The second component of the system manually cuts out sections of the picture that do not include the target object trained using a U-Net image segmentation model, and for each target region, traditional colorization is performed with a deep VGG net. In the end, colorful backgrounds and regions are blended together and refined utilizing Poisson editing techniques to create an appealing image.Through the extensive utilization of the segmentation and categorization information obtained from the segmentation based on Mask R-CNN methodology, our approach not only establishes a database that is organized according to the pre-classification of images but also selects the relevant reference image from the database for every region based on the semantic category and target areas so that all the regions of the image have a relevant reference image assigned to them rather than being in a situation whereby only one relevant image is to be used for coloring all regions in the given grayscale image.The authors offer an interesting solution for colorization of old films which simply integrates two different methods: both the auto-prediction approach and the transfer-based strategy.In this case, both methods are depending on various CNN architectures and have distinct purposes. The outline of this method enables fill-in color in the background and the corresponding regions with respect to the subject or image, effectively overcoming the drawbacks of mono step approaches. In addition, the presented approach preserves high manageability as well as a good level of stability. Finally, we apply the algorithm of seamless fusion to smooth out the joints of every rendered color region in an image. Hence, the produced colored picture has a wider range of possible visual effects than the one presented in the generation.

In this paper we nonetheless highlight the limitations of our scheme. First of all, we need to look for a better reference selection algorithm to search for most appropriate reference images for each region. In addition, there are still certain limitations with the existing transfer-based approaches e.g. texture blur, color distribution not matching with reference image which require further enhancement. The suggested system cannot perform real-time operations, due to high processing times associated with R-CNN and VGG masks. Third, we may expand our scheme to colorize different regions of an image using multiple reference images, which implies looking into the automatic referencing and ordering of the images, and dealing with the fusion and transfer of image colors as well. Finally, approaches to integrating the different methods used still need to discussed later on.

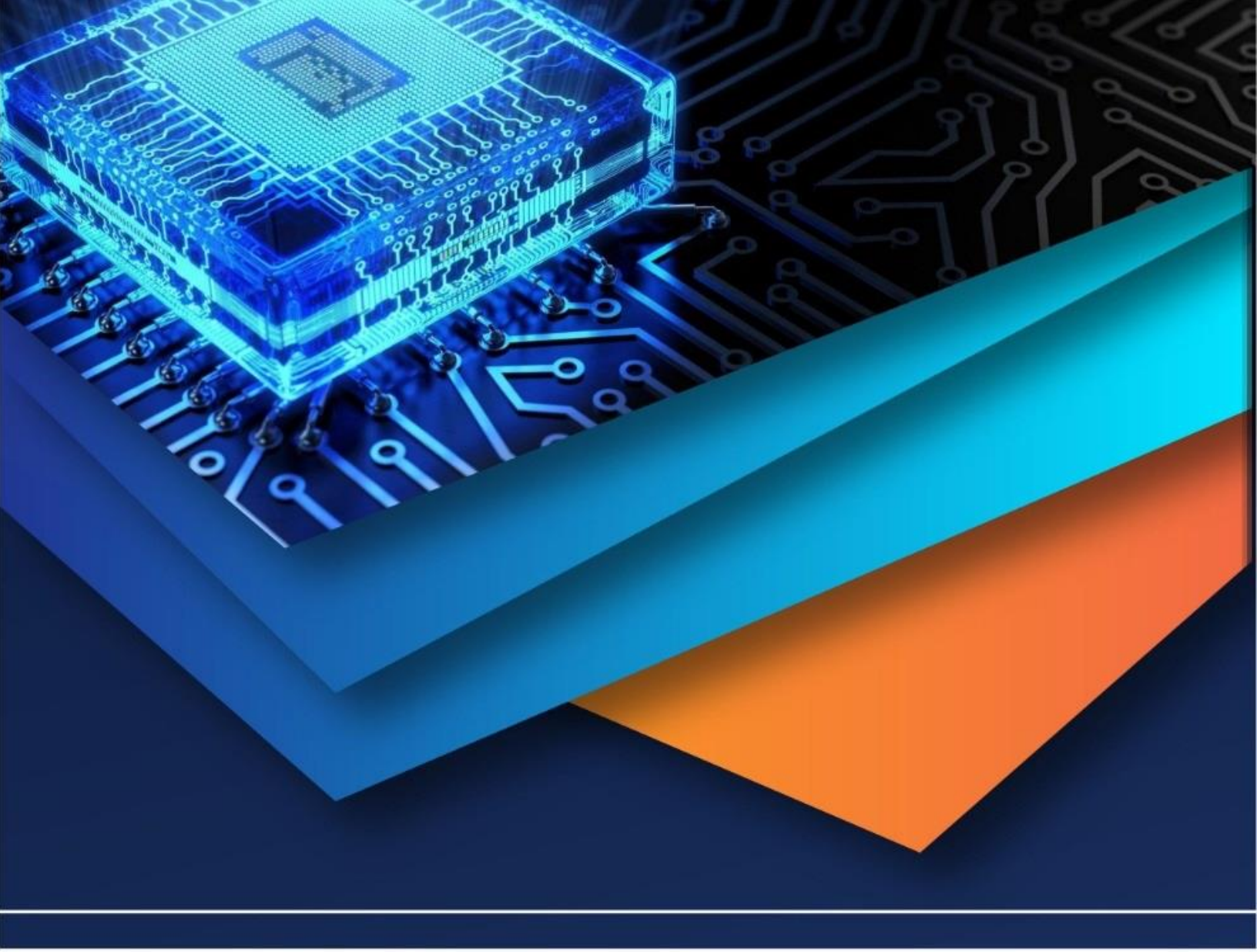
REFERENCES

[1] He, Mingming, Dongdong Chen, Jing Liao, Pedro V. Sander, and Lu Yuan. "Deep exemplar-based colorization." ACM Transactions on Graphics (TOG) 37, no. 4 (2018): 1-16.
[2] Luan, Fujun, Sylvain Paris, Eli Shechtman, and Kavita Bala. "Deep photo style transfer." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4990-4998. 2017.
[3] Markle, Wilson. "The development and application of colorization." SMPTE journal 93, no. 7 (1984): 632-635.
[4] Li, Simeng, Qiaofeng Liu, and Haoyu Yuan. "Overview of scribbled-based colorization." Art and Design Review 6, no. 04 (2018): 169.
[5] Levin, Anat, Dani Lischinski, and Yair Weiss. "Colorization using optimization." In ACM SIGGRAPH 2004 Papers, pp. 689-694. 2004.
[6] Shi, Jianbo, and Jitendra Malik. "Normalized cuts and image segmentation." IEEE Transactions on pattern analysis and machine intelligence 22, no. 8 (2000): 888-905.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[7] Larsson, Gustav, Michael Maire, and Gregory Shakhnarovich. "Learning representations for automatic colorization." In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14, pp. 577-593. Springer International Publishing, 2016.
[8] Qu, Yingge, Tien-Tsin Wong, and Pheng-Ann Heng. "Manga colorization." ACM Transactions on Graphics (ToG) 25, no. 3 (2006): 1214-1220.
[9] Yatziv, Liron, and Guillermo Sapiro. "Fast image and video colorization using chrominance blending." IEEE transactions on image processing 15, no. 5 (2006): 1120-1129.
[10] Luan, Qing, Fang Wen, Daniel Cohen-Or, Lin Liang, Ying-Qing Xu, and Heng-Yeung Shum. "Natural image colorization." In Proceedings of the 18th Euro-graphics conference on Rendering Techniques, pp. 309-320. 2007.
[11] Xu, Li, Qiong Yan, and Jiaya Jia. "A sparse control model for image and video editing." ACM Transactions on Graphics (TOG) 32, no. 6 (2013): 1-10.
[12] Iizuka, Satoshi, 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 Transactions on Graphics (ToG) 35, no. 4 (2016): 1-11.
[13] Welsh, Tomihisa, Michael Ashikhmin, and Klaus Mueller. "Transferring color to greyscale images." In Proceedings of the 29th annual conference on Computer graphics and interactive techniques, pp. 277-280. 2002.
[14] Ironi, Revital, Daniel Cohen-Or, and Dani Lischinski. "Colorization by Example." Rendering techniques 29 (2005): 201-210.
[15] Charpiat, Guillaume, Matthias Hofmann, and Bernhard Schölkopf. "Automatic image colorization via multimodal predictions." In Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III 10, pp. 126-139. Springer Berlin Heidelberg, 2008.
[16] Gupta, Raj Kumar, Alex Yong-Sang Chia, Deepu Rajan, Ee Sin Ng, and Huang Zhiyong. "Image colorization using similar images." In Proceedings of the 20th ACM international conference on Multimedia, pp. 369-378. 2012.
[17] Bugeau, Aurelia, Vinh-Thong Ta, and Nicolas Papadakis. "Variational exemplar-based image colorization." IEEE Transactions on Image Processing 23, no. 1 (2013): 298-307.
[18] Chia, Alex Yong-Sang, Shaojie Zhuo, Raj Kumar Gupta, Yu-Wing Tai, Siu-Yeung Cho, Ping Tan, and Stephen Lin. "Semantic colorization with internet images." ACM Transactions on Graphics (ToG) 30, no. 6 (2011): 1-8.
[19] Morimoto, Yuji, Yuichi Taguchi, and Takeshi Naemura. "Automatic colorization of grayscale images using multiple images on the web." In SIGGRAPH 2009: Talks, pp. 1-1. 2009.
[20] Liao, Jing, Yuan Yao, Lu Yuan, Gang Hua, and Sing Bing Kang. "Visual attribute transfer through deep image analogy." arXiv preprint arXiv:1705.01088 (2017).
[21] He, Mingming, Jing Liao, Dongdong Chen, Lu Yuan, and Pedro V. Sander. "Progressive color transfer with dense semantic correspondences." ACM Transactions on Graphics (TOG) 38, no. 2 (2019): 1-18.
[22] Wang, Chung-Ming, and Yao-Hsien Huang. "A novel automatic color transfer algorithm between images." Journal of the Chinese Institute of Engineers 29, no. 6 (2006): 1051-1060.
[23] Deshpande, Aditya, Jason Rock, and David Forsyth. "Learning large-scale automatic image colorization." In Proceedings of the IEEE international conference on computer vision, pp. 567-575. 2015.
[24] Cheng, Zezhou, Qingxiong Yang, and Bin Sheng. "Deep colorization." In Proceedings of the IEEE international conference on computer vision, pp. 415-423. 2015.
[25] Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14, pp. 649-666. Springer International Publishing, 2016.
[26] Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125-1134. 2017.
[27] Zhang, Richard, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S. Lin, Tianhe Yu, and Alexei A. Efros. "Real-time user-guided image colorization with learned deep priors." arXiv preprint arXiv:1705.02999 (2017).
[28] Sangkloy, Patsorn, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. "Scribbler: Controlling deep image synthesis with sketch and color." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5400-5409. 2017.
[29] Danielczuk, Michael, Matthew Matl, Saurabh Gupta, Andrew Li, Andrew Lee, Jeffrey Mahler, and Ken Goldberg. "Segmenting unknown 3d objects from real depth images using mask r-cnn trained on synthetic point clouds." arXiv preprint arXiv:1809.05825 16 (2018).
[30] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks." IEEE transactions on pattern analysis and machine intelligence 39, no. 6 (2016): 1137-1149.
[31] Gang net, Michel, and Andrew Blake. "Poisson image editing." In Acm Siggraph, pp. 313-318. 2003.
[32] Lin, T sung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, and C. Lawrence Zit nick. "Microsoft coco: Common objects in context." In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, pp. 740-755. Springer International Publishing, 2014.
[33] Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al. "Image net large scale visual recognition challenge." International journal of computer vision 115 (2015): 211-252.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)