Support Vector Machine (SVM) for Colorization the Grayscale Image

Shymaa Akram Alrubaie a, Israa Mohammed Hassoonb

a Middle Technical University, Continuous Education Center, Baghdad-Iraq
b Al-Mustansiriya University, College of Science, Mathematics Science, Baghdad-Iraq

ARTICLE INFO

Article history:
Received 05 June 2020
Received in revised form 15 July 2020
Accepted 20 July 2020

Keywords:
Support Vector Machine (SVM)
Discrete Wavelet Transform (DWT)
Haar-DWT
Red Green Blue color system (RGB)

ABSTRACT

Recently, there have been several automatic approaches to color grayscale images, which depend on the internal features of the grayscale images. There are several scales which are considered as a prominent key to extract the corresponding chromatic value of the gray level. In this aspect, colorizing methods that rely on automatic algorithms are still under investigation, especially after the development of neural networks used for recognizing the features of images. This paper develops a new model to obtain a color image from an original grayscale image through the use of the Support Vector Machine to recognize the features of grayscale images which are extracted from two stages: the first stage is Haar Discrete Wavelets Transform used to configure the vector that combines six of Statistical Measurements: (Mean, Variance, Skewness, Kurtosis, Energy and Standard Deviation) extracts from the grayscale image in the second stage. After the Support Vector Machine recognition has been done, the colorization process uses the result of Support Vector Machine to recover the color to grayscale images by using YCbCr color system then it converts the color to default color system (RGB) to be more clear. The proposed model will be able to move away from relying on the user to identify the source image which matches in color levels and it exceeds the network determinants of image types with similar color levels. In addition, Support Vector Machine is considered to be more reliable than neural networks in classification algorithms. The model performance is evaluated by using the Root Mean Squared Error (RMSE) in measuring the success of the assumed modal of matching the coloring (resulting) images and the original color images. So, a reality-related result has been obtained at a good rate for all the tested images. This model has proved to be successful in the process of recognizing the chromatic values of grayscale images then retrieving it. It takes less time complexity in trained data, and it isn’t complex in working.

1. Introduction

The colorization process can be done automatically by using a computer to recover the color of black-and-white and grayscale images or films. There is a wide use of this technique in image processing and it can be used scientifically to recover the color of black-and-white films or to restore old color films. Currently, most of the images that need to be colored are grayscale images and this makes them a reliable source of interest for researchers in this field. There are multi-processes for colorizing images...
and this depends on different methods and each process requires time and technical accuracy to restore the colors suitable for grayscale images Casaca et al. [1], Colorizing the grayscale image add three-dimensional (RGB) pixel values to the image that has only one dimension (luminance or intensity), while there are multi hue or saturation for the same luminance value that determines the color. For this reason, the challenges of colorizing grayscale images have appeared while no inherently "correct" solution to take into account. Usually, human vision plays an important role in the colorization method and the process of the colorization will manually depend on human vision Kumar et al. [2]. In some domains, colorizing grayscale image could be automatic, the automatic colorization is very beneficial as it decreases the time and increases accuracy. This is because it depends on information that is collected from an image. Several techniques have been used for colorization, some of the methods used reference images that have similar features to the target grayscale image to transfer its colors Futschik et al. [3].

Occasionally, there are completely automated solutions for colorizing images, one of them is performed by artificial neural networks. Many Research papers used multi kinds of Artificial Neural Networks ANN, these techniques make the automatic colorization easier and shorter processing time, and it can be applied on a larger scale An et al. [4]. With the development of ANN, the process has to gain more attention as a performed issue for perfect visual realization of artificial intelligence. It is represented as an object recognition process. The coloring process passes through several stages to reach true color image Futschik et al. [3]. In this paper, we propose an automatic image colorization algorithm using the Support Vector Machine to recognize the features of grayscale images, SVM used to recognize the features of grayscale images, its work depends on the principle of Structural Risk Minimization from statistical learning theory, SVM represents a learning algorithm and is widely used in the classification of the objects in images Awad et al. [5]. This paper is organized as follows: section 2 displays Related Work, section 3 presents Methodology, section 4 illustrates Experimental Results, while Performance measurement in section 5, while Discussions & Comparisons explained in section 6, conclusions are presented in section 7, Acknowledgments shown in section 8, Lastly references appear.

2. Related Work

Many studies have attempted to treat the colorization process for the grayscale images to get the most appropriate results. Kumar et al. [2], in this texture features, have dealt with information between the images, five statistical measurements were used as texture features that were extracted from the correlation matrix of the images which are: energy, entropy, homogeneity, contrast, and correlation. Zhang et al. [6] depended on an easy premise so that in test time, a feed-forward CNN was trained on over a million color images. The algorithm was evaluated by utilizing a "colorization Turing test,". The participants were asked to choose between the output color image and the true color image. This algorithm was successful by a higher rate in deceit 32% humans of the trials. A novel approach to colorize grayscale images that allows user intervention presented by Wallace Casaca et al. [1], this methodology makes use of multidimensional projection to map high-dimensional data to a visual space. A user can manipulate projected data in the visual space to improve further clusters and thus the colorization result, the easy use of interactive tools while still being flexible enough to enable local color modification. The approach was proved to be effective through experiment, set of examples, and comparisons. Matias Richart et al. [7] depended on training a simple classifier using back propagation that training over a set of colors and corresponding grayscale images. Each pixel depends on the gray level of the pixels surrounding it to predict the color of it. Self-Organizing Maps used to reduce the colors so the predictor domain will be kept small, this reduction process of chroma values with enough variation can generate good results for all colors in the training set. A convolutional-neural-network CNN system is built by Jeff Hwang et al. [8] to colorize the black and white image without direct human assistance. The classification-generates model to colorize images that are more aesthetically pleasing than those created by the baseline regression-based model. Thus, this shows the efficiency of the method and its future development for better work.

Domonkos Varga, et al. [9] presents an automatic approach that can produce reality color images from an input grayscale image. The method employs deep learning techniques depend on VGG-16 and a two-stage of Convolutional Neural Network (CNN) that predicts the channels of color. In this method, there was no human intervention in the resulting data, rather it was a fully automatic method, the results measured by found that Quaternion Structural Similarity (QSSIM) that gives a good evaluation for the quality of colorization. Jiancheng An et al. [4] developed the work of Domonkos [9] by using the loss of cross entropy that is given a good a plausible colorization. Therefore, the method enhanced the colorization process. Joharia, et al [10] in this work, states that the colorization process consists of two stages; the first stage depends on an auto-encoder colorization network to determine the similarity of the images and produce an initial colored image. The second stage consists of several specialist networks, whose tasks are improving the quality of the initial colored image and producing real-looking, plausible images. The method is compared with several previous methods in order to demonstrate the capability of the proposed method.
3. Methodology

The proposed model is based on SVM neural network to recognize the colors of the grayscale images, this model works by input the extracted features from the grayscale images as inputs. SVM is training by multi color images, these images are converted to grayscale then the features were extraction from them which are used to train SVM. The trained SVM will be used to recognize the features of grayscale images and classify it to any of the trainer images in a similar way, then the process of recovering color will be done to get the appropriate colors of it depending on SVM recognition. The success of SVM depends on the type of entering features, these features are extracted by the use of dual stages which they are Haar-DWT and six of statistical measurements. By these two stages, one vector obtained to be the input of training SVM, these features are proved to be appropriate in SVM usage, the model is illustrated as following block diagram in Fig. 1:

![Block Diagram of the Proposed Model](image)

**3.1. Extract Features**

The important characteristics (features) will be extracted from the grayscale image. So, to reduce the non-useful information from the image, the features are extracted by two stages:

**3.1.1. Haar-DWT**

The first stage of feature extraction is Haar-DWT. It is used to find good features, Haar-DWT represents the simplest kind of DWT in addition to its speed and easy processing [11] [12]. The Haar-DWT, like other wavelet transforms, consists of a discrete signal into two sub-signals. Four sub-images (LL, LH, HL, and HH) decomposed the plain image from the two-dimensional wavelet transform. These sub-images are: the low-frequency signal and high frequency signal Ahadpour et al. [13]. Well, the low-frequency signal is the most important part, since the high frequency part of an “additive” impact, for more details about Haar-DWT see [11] [12]. Haar-DWT is decomposed to four levels, in each level the low-frequency signal(LL) is decomposed to be the input to the next level of the Haar-DWT, so the LL image can be decomposed to four new sub-images which make a tree of sub-images, then the four decomposition of LL sub-image can be converted to be vector of this stage. As seen in Fig. 1

**3.1.2. Statistical Measurements**

The second stage is used to increase the strength of the resulting vector of Haar-DWT. In this stage, six of Statistical Measurements were used, which are: Mean, Variance, Skewness, Kurtosis, Energy and Standard Deviation Sachdeva et al. [14]. These statistical measurements are considered used more to extract special features from the images, so the statistical measurements are extracted directly from the image, more details about these statistical measurements found in [14] [15]. The extracted statistical measurements are added to the Haar-DWT vector to compose the final vector, this vector acts as an input to the next step. See Fig. 1

**3.2. Support Vector Machine**

Support vector machine (SVM) represents a supervised learning method which can be able to classify objects by employing the machine learning theory to increase prediction accuracy with avoiding over-load of data Sakhthivel et al. [16], SVM depends on the concepts of Structural Risk Minimization from statistical learning theory; it utilizes a learning algorithm in feature space to train the features with optimization theory that performs a bias learning theory. The objective of the SVM is to find the highest minimum distance between any two points of data quickly Sakhthivel et al. [16]. SVM locates the closest data vectors called support vectors (SV) to the decision boundary in the training set, so it depends on these vectors to classify a given new test vector Xia et al. [17]. SVM can be handled two problems of input space: linear and nonlinear Awad et al. [5], the default formulation of SVM is used to classifier linear problem while the nonlinear problem of input space will be separable, therefore the kernel function can be used to transform the data to a higher-dimensional space, which referred to as “kernel space”. In the kernel space, the classification task will be done while in the input space in a high-order separating hyperplane will be solved. Awad et al. [5].in nonlinear SVM, the kernel K is represented as a hyperplane in a feature space induced by its “the kernel is defined as a dot product in that space”). Therefore, the hypothesis space is defined as “a set of “hyperplanes” in the feature space induced by K”, thus SVM depends on two important things: the hypothesis spaces and the kernel function. There are many kernel function which can be used with SVM, for example: Linear kernel, Polynomial function, Gaussian radial basis function (RBF), and Laplacian radial basis function. The Kernel selection that is used in SVM is based on the data specifics Awad et al. [5]. In this proposed work, nonlinear SVM will be used to classify the input data features with “Gaussian Radial Basis Function kernel RBF”. SVM with RBF-Kernel is considered to be the best solution for the classification
The trained SVM function to identify the support vectors $s_i$ which has been used as an optimization method, according to equation (1) [19]:

$$c = \sum_i a_i k(s_i, x) + b \quad (1)$$

Where $i$ is the class number, $k$ is a kernel function, $a_i$ is a weight, and $b$ represents bias that are used to classify vectors $x$. If $c \geq 0$, $x$ is classified as a member of the first group, otherwise it is classified as a member of the second group.

"Gaussian Radial Basis Function kernel RBF" is the kernel function of trained SVM that uses to map the training data into kernel space the value of it depends on a scaling factor ($\sigma$), "Gaussian Radial Basis Function kernel RBF" represented as the equation (2) [11]

$$K(x, u) = \exp \left[ \frac{-||x-u||^2}{\sigma^2} \right] \quad (2)$$

Where $x$ is the test vector, $u$ is the train vector, and $\sigma$ is the represented as a scaling factor (standard deviation).

### 3.3. Color Recovering

The process of converting the color from the source (color image) to a target (grayscale image) needs matching luminance and texture features between the images. In this process, two stages are followed: first, YCbCr color system will be used to recover colors. YCbCr represents the decorrelated color space that is used for colorization process. YCbCr color space also consists of three channels Y, Cb, and Cr. So Y Channel: it is an achromatic luminance channel, Cb represents the chromatic channel of difference between blue component and a reference value, finally Cr represents the chromatic channel of difference between red components and a reference value Kumar et al. [2] while a grayscale image represents one channel (luminance channels) which it is matched with channels of colors images. After the target image matches a source image, color information is transferred between the two images of pixels while retaining the value of the original lighting. For more details about YCbCr color system, see Kumar et al. [2]. In the second stage, the YCbCr color image is converted from the YCbCr color system to the default color system (RGB) which is a compound of three chromatic channels (Red, Green, Blue), in this method, the color image will be more clear and the colors are considered nearest to the reality.

### 4. Experimental Results

In this paper, SVM training with 20 images, these images are selected to be different in size, type format (BMP, and jpg), and resolution. The diversity in selecting images has proved that SVM is more reliable and more recognized to test different images. Training images have been selected on five color spectra: 5 images with chromatic spectrum for orange gradients, 3 images with chromatic spectrum for blue gradient, 3 images with chromatic spectrum of red and syphilis, 4 with chromatic spectrum for green gradients, and 5 images with chromatic spectrum overlapping blue and green. These images are subject to the unification of each size in order to obtain a vector of equal length for all images when extracted the vector using the first stage (Haar-DWT) in extract features process. Thus, the resizing of an image is to be about 160*250 then the image is converted to be the grayscale image, this stage produces a vector with 160 long, addition to this vector the vector obtained from the second stage of the extract features process to obtain a vector at 166 long for each image. SVM used Gaussian Radial Basis "RBF" as a Function kernel with a 1 scaling factor ($\sigma$) for training with these vectors to classify the test images which are matching color images that training with SVM. The training images showed in Fig. 2.

Then SVM was tested with 40 images that are different from the training images. The strength of the SVM was detected in recognizing the color spectra of images according to the characteristics (features) that were prepared for them. The tested image represents a grayscale image, then the features which were extracted from it would be the inputs to SVM to recognize the nearest matching image. The colors were taken from the matching image and turned into the new (tested) image by the use of YCYR color system method then they are converted to RGB color system. Some of the results of the tested images showed in Fig. 3.
Figure 2- Original color images used for training SVM (20 images)
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Figure 3- (a) Original Color Images, (b) Grayscale Images, (c) RGB Color Images, (d) The Matching Training Image
5. Performance Measurement

In this paper, the Root Mean Squared Error (RMSE) is used to measure the success of the assumed method of matching the coloring (resulting) images and the original color images, this scale is used to measure the quality of images and it is very commonly used and makes an excellent general purpose error metric for numerical predictions Rajasree et al. [20]. It measured RMSE between the original color image and the result color image. It represents the percentage of error measurement of the matching between color images and original images. The formula of RMSE is in (3) Futschik et al. [3]

$$RMSE(p_{a\beta}, \tilde{p}_{a\beta}) = \frac{1}{1+\epsilon} \sum_{i=1}^{M} \sum_{j=1}^{N} \left\|\frac{(p_{i,j} - \tilde{p}_{i,j})^2}{\epsilon+y}\right\|^2$$  

(3)

Where \(M\) is the number of images and \(N\) is the number of pixels in an image. The error \(\epsilon\) is calculated in de-correlated 2-channel \(a\beta\) color space, as in equation (4) [3]

$$a\beta = \frac{B - \frac{1}{L} (R + G)}{L + \epsilon}$$  

(4)

Where \(L = (R+G+B)/3\), \(\epsilon = 0.0001\) and R,G,B channels are normalized \(\epsilon [0; 1]\).

The result of the evaluation and the efficiency of the proposed model for the colorization images see in (Table 1) that appears RMSE of the proposed model.

Table 1- The Results of RMSE of the Proposed Method

<table>
<thead>
<tr>
<th>No.</th>
<th>RMSE</th>
<th>No.</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55.80571</td>
<td>21</td>
<td>57.37862</td>
</tr>
<tr>
<td>2</td>
<td>103.7273</td>
<td>22</td>
<td>68.17108</td>
</tr>
<tr>
<td>3</td>
<td>61.61202</td>
<td>23</td>
<td>41.25144</td>
</tr>
<tr>
<td>4</td>
<td>42.57201</td>
<td>24</td>
<td>88.7847</td>
</tr>
<tr>
<td>5</td>
<td>44.79534</td>
<td>25</td>
<td>74.59644</td>
</tr>
<tr>
<td>6</td>
<td>81.40136</td>
<td>26</td>
<td>65.05274</td>
</tr>
<tr>
<td>7</td>
<td>68.11881</td>
<td>27</td>
<td>97.97648</td>
</tr>
<tr>
<td>8</td>
<td>85.64872</td>
<td>28</td>
<td>91.2866</td>
</tr>
<tr>
<td>9</td>
<td>117.5906</td>
<td>29</td>
<td>72.82589</td>
</tr>
<tr>
<td>10</td>
<td>129.8653</td>
<td>30</td>
<td>77.4869</td>
</tr>
<tr>
<td>11</td>
<td>72.79234</td>
<td>31</td>
<td>37.13677</td>
</tr>
<tr>
<td>12</td>
<td>167.9442</td>
<td>32</td>
<td>54.89998</td>
</tr>
<tr>
<td>13</td>
<td>50.31454</td>
<td>33</td>
<td>63.6623</td>
</tr>
<tr>
<td>14</td>
<td>23.27662</td>
<td>34</td>
<td>77.80699</td>
</tr>
<tr>
<td>15</td>
<td>30.54028</td>
<td>35</td>
<td>105.5327</td>
</tr>
<tr>
<td>16</td>
<td>97.76456</td>
<td>36</td>
<td>109.4407</td>
</tr>
<tr>
<td>17</td>
<td>84.39768</td>
<td>37</td>
<td>71.15946</td>
</tr>
<tr>
<td>18</td>
<td>91.47904</td>
<td>38</td>
<td>70.44696</td>
</tr>
<tr>
<td>19</td>
<td>84.6228</td>
<td>39</td>
<td>43.30157</td>
</tr>
<tr>
<td>20</td>
<td>102.5789</td>
<td>40</td>
<td>49.30605</td>
</tr>
</tbody>
</table>

6. Discussions & Comparisons

The colorization process depends on two parameters to obtain the final evaluation: Are the results good or not? First parameter can be deduced from the results of RMSE of the proposed model is presented in (Table 1), the analysis of the one that is presented in Chart in Fig 4. It can be noted from these results that most of the tested images in the SVM did not exceed the error rate of 100 although the SVM trained on multi spectra images. Also, from these results showed that only 7 images exceeded the 100 error rate. Second parameter can be deduced if we take a visual scale of view of the image note that the results color images from the proposed model are acceptable by viewer and not far from the original color of images. By relying on these two parameters for evaluation, it becomes clear for us that the model works well and has good results. So, a reality-related result were obtained at a good rate for all the tested images. In addition, the method deviated of manual comparison to an automatic method and a void the troubles of Artificial Neural Network ANN.

By making a comparison with previous states of art colorization models: the models of: Zhang et al. [6], Jeff Hwang et al. [8], and Domonkos Varga, et al. [9] in all these models Convolutional Neural Networks (CNN) would be used, CNN which considered one of Artificial Neural Network ANN. While Joharia, et al [10] used the ANN by two stages in their work, and this increases the time complexity and made the performance more expensive. ANN suffers from many problems and is regarded as a less performance technique compared with SVM Nitze et al. [18], SVM distinguished by multi options that are not found in ANN [5] [18]. SVM avoids the double curse of dimensionality and overfitting, therefore it can exceed a problem of classical multifocal minima. Overfitting in machine learning model occurs when all training data is aimed to be a zero, that is more likely to occur with machine learning approaches by minimizing the structural risk rather than the empirical risk, as in the case of ANN “the training metrics depend on variants of the sum of squares error” [5]. The classification accuracy is obtained by the SVM classification algorithm is Higher than ANN classification algorithms Kumar et al. [21]. The complexity of SVM is automatically determined by selecting the number of supported vectors, while in ANN the complexity is determined by limiting the feature set, and this made control model complexity of SVM less than ANN Awad et al. [5]. SVM represents the best solution for the different classification, SVM has a higher classification performance, faster calculation time, and more robustness compared to ANN Nitze et al. [18].

![Figure 4- RMSE of the proposed method](image-url)
7. Conclusion

This paper developed an automatic technique to colorize grayscale images by using SVM as a recognition technique for the most similar color images to convert its colors to target grayscale images. The success of SVM Recognition is based on the kind of extracted features which are regarded as inputs to SVM. Features were extracted by two stages: the first stage is HAAR-DWT that product vector was combined with the vector product from the second stage which is; six of statistical measurements: Mean, Variance, Skewness, Kurtosis, Energy, and Standard Deviation. The final vector of features extraction process is represented as an input to SVM in training and testing phases. SVM trained on multi images with various color spectra to be able to recognize various colors images. The performance of this model was evaluated by using Root Mean Square Error (RMSE) measure, the obtained results of this model would be good depending on the values of RMSE and the visual scale of view of the image by the viewers. SVM trained with ‘Gaussian Radial Basis RBF’ as Function kernel and this made the classification more efficient in comparison with ANN, SVM distinguished by it is avoid the troubles of ANN, Also. It has a higher performance and less training time in classification algorithms.

Acknowledgments

We would like to thank the universities that facilitated the task of researchers contributing to the research and affiliated to it, which are: Middle Technical University (www.mtu.edu.iq), Al-Mustansiriyyah University (www uomustansiriyyah.edu.iq), for their support to the present study.

REFERENCES