Past to Present (P2P): Road Thermal Image Colorization

Yuseong Park
South Dakota State University

Follow this and additional works at: https://openprairie.sdstate.edu/etd

Part of the Computer Engineering Commons, and the Electrical and Computer Engineering Commons

Recommended Citation
https://openprairie.sdstate.edu/etd/4074

This Thesis - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.
PAST TO PRESENT (P2P): ROAD THERMAL IMAGE COLORIZATION

BY

YUSEONG PARK

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Computer Science

South Dakota State University

2020
This thesis is approved as a creditable and independent investigation by a candidate for the master’s degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Sung Shin
Advisor

Siddharth Suryanarayanan
Department Head

Dean, Graduate School
This thesis is dedicated to the development of autonomous vehicles.
ACKNOWLEDGEMENTS

I would like to say thank you to my advisor, Dr. Sung, Shin. He has always provided valuable guidance in the field of computer science. I cannot forget Prof. Sung, Shin’s teachings forever. I would like to also take this time to appreciate the committee members of my thesis who are Dr. Kwanghee, Won and Dr. Wookjae, Heo, and. I greatly appreciate all the graduate faculty representatives and committee members. Dr. Won has given me enthusiasm towards computer vision field research. Dr. Heo made it possible for me to experience various computer field research. Dr. Shin did everything he could as an advisor to aid and assist me.

Finally, I want to appreciate all my family and friends for supporting me.
CONTENTS

ABBREVIATIONS........................................................................................................... vii

LIST OF FIGURES .......................................................................................................... viii

LIST OF TABLES ............................................................................................................. ix

ABSTRACT....................................................................................................................... x

1. INSTRUCTION ............................................................................................................ 1

2. RELATED WORK ......................................................................................................... 4

  2.1. Thermal camera ....................................................................................................... 4

  2.2. Colorization method .............................................................................................. 5

3. BACKGROUND ............................................................................................................ 7

  3.1. GENERATIVE ADVERSARIAL NETWORKS ............................................................ 7

  3.2. pix2pix .................................................................................................................... 8

4. PROPOSED METHOD .................................................................................................... 9

5. EXPERIMENTAL .......................................................................................................... 12

  5.1 EXPERIMENTAL SETUP ........................................................................................ 12

  5.2 EXPERIMENT RESULT .......................................................................................... 15

  5.3 RESULT AND ANALYSIS ....................................................................................... 23

    5.3.1 Quantitative measure ....................................................................................... 23
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAS</td>
<td>Advanced Driver Assistance System</td>
</tr>
<tr>
<td>AV</td>
<td>Autonomous Vehicle</td>
</tr>
<tr>
<td>BW</td>
<td>Black and White</td>
</tr>
<tr>
<td>cGAN</td>
<td>conditional Generative Adversarial Nets</td>
</tr>
<tr>
<td>DCGAN</td>
<td>Deep Convolutional Generative Adversarial Nets</td>
</tr>
<tr>
<td>FPS</td>
<td>Frame Per Second</td>
</tr>
<tr>
<td>GAN</td>
<td>Generative Adversarial Nets</td>
</tr>
<tr>
<td>LWIR</td>
<td>Long-Wave Infrared</td>
</tr>
<tr>
<td>NIR</td>
<td>Near-Infrared</td>
</tr>
<tr>
<td>PSNR</td>
<td>Peak Signal-to-noise ratio</td>
</tr>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>SSIM</td>
<td>Structural Similarity Index Map</td>
</tr>
<tr>
<td>SWIR</td>
<td>Short-Wave Infrared</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. The example of KAIST-MS dataset ................................................. 1

Figure 2. Example of fog tunnel with the for different cameras ..................... 4

Figure 3. Generative Adversarial Networks Overview ................................... 7

Figure 4. P2P (proposed method) Overview .................................................. 9

Figure 5. Overview concatenation ................................................................. 10

Figure 6. Result of Test set sample ............................................................... 10

Figure 7. Example input of P2P (proposed method) ..................................... 11

Figure 8. 60th Input and result of 5 models .................................................. 16

Figure 9. 437th Input and result of 5 models ................................................ 17

Figure 10. 549th Input and result of 5 models ................................................. 18

Figure 11. 293rd Input and result of 5 models ................................................. 19

Figure 12. 765th Input and result of 5 models ............................................... 20

Figure 13. 1969th Input and result of 5 models .............................................. 21

Figure 14. 837th Input and result of 5 models ................................................. 22
LIST OF TABLES

Table 1. Train set and Test set of KAIST-MS dataset........................................ 12

Table 2. List of input................................................................................................ 13

Table 3. List of train set and test set.................................................................... 14

Table 4. SSIM and PSNR score of 5 model’s result............................................. 24

Table 5. p-value of proposed method.................................................................... 25
ABSTRACT

P2P (PAST TO PRESENT): ROAD THERMAL IMAGE COLORIZATION

YUSEONG PARK

2020

Thermal image colorization into realistic RGB image is a challenging task. Thermal cameras are easily to detect objects in particular situation (e.g. darkness and fog) that the human eyes cannot detect. However, it is difficult to interpret the thermal image with human eyes. Enhancing thermal image colorization is an important task to improve these areas. The results of the existing colorization method still have color ambiguities, distortion, and blurriness problems.

This paper focused on thermal image colorization using pix2pix network architecture based on Generative Adversarial Net (GAN). Pix2pix is a model that transforms thermal image into RGB image, but our proposed model used three input types of images which are present as frame thermal image, present frame RGB image, and previous frame RGB image. By extracting the color information (i.e. luminance and chrominance) of the previous frame RGB image, the result obtained a more realistic RGB image.

Experiments use two kinds of evaluation method, which are quantitative measure and qualitative measure. First, quantitative measure is the calculation of specific numerical scores, the method names are PSNR and SSIM. Second, qualitative measure is human subjective evaluation. Evaluation method compared and evaluated pix2pix and our proposed method with the two types of measuring method.
1. INSTRUCTION

Recently, thermal cameras are widely equipped in vehicles as ADAS and AV system, which the cameras work for the driver’s convenience to improve recognizing objects on the road. A thermal camera is a device that captures the heat emitted from an object, and it is used very efficiently to detect objects not only at nighttime but also during the daytime. Especially, it has the advantage of being able to detect objects that are difficult to recognize (e.g. darkness, fog, dust, and sun glare etc.) [14]. The thermal image made of grayscale has a heterogeneity that is difficult to interpret the object on the road by the drivers, so enhancing thermal image colorization into RGB image is very important to improve the viewing performance of the thermal image for the driver [1, 2, 3, 4, 5, 6].

Transforming thermal image into RGB image is color ambiguity problem. In Figure 1, the thermal image is difficult to distinguish color difference, so there is an ambiguous problem, which is the vehicle color (i.e. A vehicle can have many colors). Therefore, it is

Figure 1: The example of KAIST-MS dataset [15]:

a) Thermal image (left) and b) RGB image (right).
important to apply the same color as the color in the RGB image. Recent work has shown impressive colorization results through GAN, CNN, and RL [1, 2, 3, 4, 5, 6, 7, 9, 16]. The GAN-based model methods were adjusted for generator, discriminator network and loss function to solve colorization problem of ambiguity, blurriness, and distortion [1, 2, 3, 7, 9, 16]. In [5, 6], thermal image was colorized through development of CNN. [4] used the method to allocate color to each object by object segmentation through RL. These adjustments have shown impressive results.

In this paper, our proposed method used KAIST-MS dataset [15] which road dataset composed of RGB images and thermal images. Also, the KAIST-MS road dataset is video-based image dataset. Most of recent GAN based methods uses paired image dataset (i.e. present thermal image and present RGB image) [1, 2, 7, 9, 16]. However, our model is the colorization thermal image method using the pix2pix [16] network architecture after classifying present frame thermal image, present frame RGB image, and previous frame RGB image of road dataset. The previous frame RGB image is for extracting and using color information during colorization processing. This method solved not only the color ambiguity problem but also distortion and blurriness problems simultaneously.

These results are evaluated by two types of evaluation methods which are quantitative measure and qualitative measure [17]. Quantitative measure is the calculation of specific numerical scores, the method names are Peak Signal-to-noise ratio (PSNR) and Structural Similarity Index Map (SSIM). PSNR is a method for evaluating the quality of the generated image [17]. SSIM is compared with luminance, contrast, and structure of real and generated image [17, 19]. Qualitative measure is human subjective evaluation. The experiment was evaluated with 5 results and consists of pix2pix and four types of proposed
method. The four results of proposed methods. The four results of the proposed methods are composed of down pixel of the previous frame RGB image input and limitation number of frames that can only load colors to find an environment that can solve colorization problems.
2. RELATED WORK

2.1. Thermal camera

Figure 2. Example of thermal image [14]: a. fog tunnel with the for different cameras, b. sunny day condition (above) and rainy-day condition (below).

As shown in Figure 2. a, the thermal camera is composed of three types which are Near-Infrared (NIR), Short-Wave Infrared (SWIR), and Long-Wave Infrared (LWIR). The three types of thermal images have different wave lengths, and the wavelengths are as follows : the wavelength of NIR is from 0.7 to 1.4 (micron-meter), SWIR is from 1.4 to 3.0 (micron-meter), and LWIR is from 8 to 15 (micron-meter) [14]. As shown in Figure 2. b, thermal camera vulnerability on the road, which shows the heat emitted from the object. The lane is invisible during the rainy-day condition because of the temperature difference between the lane and the road surface [14]. Thermal camera wavelength of KAIST-MS dataset [15] is LWIR which is suitable for use on the road because wavelength camera is very sensitive to the emitted heat of the objects.
2.2. Colorization method

Recently, thermal image colorization is one of the fields that is being vigorously researched in the computer vision field. The colorization method showed a lot of impressive results. These studies have been trying to improve the colorization problem [1,2,3,4,5,6,7,9,16]. The method of In Limmer et al. [4] is based on Markov Decision Processes (MDP) of RL, they colorize thermal image using object recognition. Segmentation of the objects (e.g. plants, buildings, sky, water, roads and others) in the thermal image, colorization is performed by calculating the color for each class object. However, the result is still an unnatural color. Recently, the colorization of the thermal image using CNN have been widely researched [5,7]. However, it is difficult to solve the colorization problem with the CNN model by itself. Unlike CNN based colorization, GAN can solve the colorization problem because the GAN based model effectively minimizes the calculation of loss by indicating the output direction desired by the user [16]. Therefore, most of the method of thermal image colorization is based on GAN model in order to solve colorization problem.

This section will introduce the GAN based model of thermal image colorization [1,2,3,7,9]. In Isola et al. [16], they have already solved colorization problem to some extent by using L1 loss. However, in Berg et al. [1], the Loss function was separated into a luminance and a chrominance loss. In Kuang et al. [2], they employed cGAN. The result of their method shown improved spatial smoothness and more realistic details of colorization by utilizing TV loss and perceptual loss. Unlike Kuang et al. [2] method, Nyberg et al. [3] method performed thermal image colorization using cycleGAN [18] of style transfer base. Suárez et al. [9] method based on DCGAN Radford et al. [13]. They
separated generator into three (i.e. Red, Green, and Blue channel) of RGB channels. They all showed good performance, but their method still has colorization problems.
3. BACKGROUND

3.1. GENERATIVE ADVERSARIAL NETWORKS

Generative adversarial network (GAN) was published by Goodfellow et al. [11].

GAN showed an outstanding performance in the computer vision field. The framework of GAN is composed with discriminator and generator model.

The generative model makes an effort to generate the counterfeits that are indistinguishable from the genuine articles, and the discriminative model judges whether a counterfeit made by a generator is genuine or counterfeit. Both of the models compete to improve their methods until the counterfeits are indistinguishable from the genuine articles. The adversarial loss function of GAN is calculated as follows [11]. The pix2pix [16] to be introduced next is also based on the GAN model.
3.2. pix2pix

Pix2pix (Image-to-image translation with conditional adversarial networks) was published by Isola et al. [16]. Pix2pix showed several images transforming results (e.g. Map photo to aerial photo, black and white photo to color photo, sketch photo to photo, Day photo to night photo, and thermal photo to color photos etc). Pix2pix is a model that based on the Conditional Generative Adversarial network (cGAN) [12] architecture. The cGAN model is also based on GAN. The difference between GAN and cGAN is that conditional information is additionally entered, and the generative model generates data by referring to conditional information. The difference is that the input of the GAN is a noise vector, whereas the input of the pix2pix is paired image. The input of pix2pix uses paired image of Black and White (BW) image and color image. Pix2pix solved the blurry problem by adding L1 loss, and the generative model was shown to have improved performance by using U-net, and discriminative model showed fast computation by analyzing the image in patch units using Markovian discriminator. Despite the results of pix2pix, it still has colorization problems, which are color ambiguity, image distortion, and blurriness problems.
4. PROPOSED METHOD

In this section, proposed method is introduced. The name of proposed method is P2P (past to present), which means using the past RGB image frame to colorize the present RGB images. Proposed method employed pix2pix [16], which method is a comparative model of the P2P. As explained in related work, pix2pix had shown outperformance in many fields of colorization. However, it still had a colorization problem. Most of the thermal image colorization method [1,2,3,7,9,16] were performed with only two inputs which are thermal image and RGB image as ground truth.

In the proposed method, the method employs three types of input which are previous frame RGB image, present thermal image, and present RGB image as ground truth. The difference from the existing methods [1,2,3,7,9,16] is that the previous frame RGB image has additional set as an input of the generative model. The reason of input of previous frame RGB image is to extract color information. The generator input,
present frame thermal image and previous frame RGB image is concatenated as generative model inputs. Both of input image concatenation is shown in Figure 6, the input image size is converted from (256, 256, 3) to (256, 256, 6) [20]. (256, 256, #) means image size, ( #, #, 6) means color information of RGB, and since two images are included, the RGB values of the two images are concatenated, so the RGB value is 6. In fact, the concatenation idea of GAN model is also a technique used to input thermal and RGB image in the discriminator model of pix2pix [16]. However, P2P has two inputs discriminator model as well as the generator model.

![Figure 6. The example of results.](image)
In Figure 7, P2P outperformed in the training sample. In the P2P training sample, a good result was obtained, but output of test set was a problem of coloring the object that was not in the thermal image. Therefore, by reducing the pixel size of the previous frame RGB image from \((256 \times 256)\) pixels to \((32 \times 32)\) pixels, previous frame RGB image minimized the loading of the object structure, and proceeded to bring only the color information. Figure 8 shows the type of P2P input image, the first row is the input image of \((256 \times 256)\) pixels, and the second row is the input of the \((32 \times 32)\) pixels image. The next section will show the experimental methods, results, and the analyzation.

**Figure 7. Example of input P2P (proposed method).**
5. EXPERIMENTAL

5.1 EXPERIMENTAL SETUP

In this section, I will introduce the experiments setup. The experiment uses KAIST-MS dataset [15], which is composed of road thermal image and road RGB image, nighttime and daytime. Table 1 shows the number of trainsets and test-sets of KAIST-MS dataset. The dataset is a frame image format, originally a video format. The video’s FPS is 2. And the image size of dataset is (640 x 512).

<table>
<thead>
<tr>
<th></th>
<th>Test set</th>
<th>Train set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day-time image</td>
<td>33,399</td>
<td>29,179</td>
</tr>
<tr>
<td>Night-time image</td>
<td>16,788</td>
<td>15,962</td>
</tr>
<tr>
<td>RGB image</td>
<td>25,093</td>
<td>22,570</td>
</tr>
<tr>
<td>Thermal image</td>
<td>25,093</td>
<td>22,570</td>
</tr>
</tbody>
</table>

The experiment is divided into five models to train. Each model is shown in Table 2. The experiment of proposed method has a difference between the number of previous frames and the image pixel size. The previous frame RGB images are experimented with 10th and 3rd previous frame RGB images. This is to observe how much the frame is performing. Also, the experiment implemented with (32 x 32) and (256 x 256) pixel size of previous frame RGB image. There was a problem of bringing not only the color
information but also the structure of the object, when the previous frame RGB image size was set to (256 × 256) pixel size in the proposed model. Therefore, to only receive the color information, previous frame RGB image (256 × 256) decreased the pixel size to (32 × 32).

Table 2. List of input.

<table>
<thead>
<tr>
<th>Model</th>
<th>Input</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Generator (train&amp;test)</strong></td>
</tr>
<tr>
<td></td>
<td>• Present frame <strong>thermal</strong> image</td>
</tr>
<tr>
<td></td>
<td>• Present frame <strong>thermal</strong> image and previous frame <strong>RGB</strong> image</td>
</tr>
<tr>
<td>Pix2pix[16]</td>
<td>• Present frame <strong>thermal</strong> image and previous frame <strong>RGB</strong> image</td>
</tr>
<tr>
<td>P2P32_10Frame</td>
<td>• Present frame <strong>thermal</strong> image and previous frame <strong>RGB</strong> image</td>
</tr>
<tr>
<td></td>
<td>• Previous frame 10th <strong>RGB</strong> image in (32 × 32) pixel size of (256 × 256) image size.</td>
</tr>
<tr>
<td>P2P256_10Frame</td>
<td>• Present frame <strong>thermal</strong> image and previous frame <strong>RGB</strong> image</td>
</tr>
<tr>
<td></td>
<td>• Previous frame 10th <strong>RGB</strong> image in (256 × 256) pixel size of (256 × 256) image size.</td>
</tr>
<tr>
<td>P2P32_3Frame</td>
<td>• Present frame <strong>thermal</strong> image and previous frame 3rd <strong>RGB</strong> image in (32 × 32) pixel size of (256 × 256) image size.</td>
</tr>
<tr>
<td>P2P256_3Frame</td>
<td>• Present frame <strong>thermal</strong> image and previous frame 3rd <strong>RGB</strong> image in (256 × 256) pixel size of (256 × 256) image size.</td>
</tr>
</tbody>
</table>
Table 3. List of train set and test set

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>• pix2pix [16]</td>
<td>• Train set&lt;br&gt;RGB image – 1,738 / Thermal image - 1,738&lt;br&gt;• Test set&lt;br&gt;Thermal image – 258&lt;br&gt;• Size: 256 × 256</td>
</tr>
<tr>
<td>• P2P32_10Frame&lt;br&gt;• P2P256_10Frame&lt;br&gt;• P2P32_3Frame&lt;br&gt;• P2P256_3Frame</td>
<td>• Train set&lt;br&gt;RGB image – 1,738 / Thermal image – 1,738&lt;br&gt;• Test set&lt;br&gt;Thermal image – 258 / previous RGB image – 258&lt;br&gt;• Size: 256 × 256</td>
</tr>
</tbody>
</table>

Train and test used the same number of images as Table 3. The train process uses inputs as shown in Table 2, and the test process uses only the generator inputs. The five models set 200 iteration epochs, 1 batch size, 256 by 256 input image size and Adam optimizer which learning rate is 0.0002 and bata_l is 0.5.

All models used keras library, python language, and implemented on the Google Colab pro server. The environment of Google Colab pro server is CPU: Intel (R) Xeon (R) CPU @ 2.30GHz, GPU: GPU : T4&P100 (video memory : 16GB), and Memory : 25GB.
5.2 EXPERIMENT RESULT

In this section, I will show the output result of 5 models. The Figure is composed of input image and output image. Each column of result figure consists input image and output for each model. It is composed as follows.

- First column is present frame thermal image.
- Second column is previous frame RGB image.
- Third column is present frame RGB image.
- Fourth column is output image of each model.

Each row of result figure consists output of five models. It is composed as follows.

- First row is the input and output of pix2pix
- Second row is the input and output of P2P32_10Frame
- Third row is the input and output of P2P256_10Frame
- Fourth row is the input and output of P2P32_3Frame
- Fifth row is the input and output of P2P256_3Frame
<table>
<thead>
<tr>
<th>Present frame</th>
<th>Previous frame</th>
<th>Present frame</th>
<th>Output Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal image</td>
<td>RGB image</td>
<td>RGB image</td>
<td>Image</td>
</tr>
<tr>
<td>pix2pix</td>
<td>P2P32.10Frame</td>
<td>P2P256.10Frame</td>
<td>P2P32.10Frame</td>
</tr>
<tr>
<td></td>
<td>P2P256.3Frame</td>
<td></td>
<td>P2P256.3Frame</td>
</tr>
</tbody>
</table>

Figure 8. 60th Input and result of 5 models.
Figure 9. 437th Input and result of 5 models.
<table>
<thead>
<tr>
<th>Present frame</th>
<th>Previous frame</th>
<th>Present frame</th>
<th>Output Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal image</td>
<td>RGB image</td>
<td>RGB image</td>
<td>Image</td>
</tr>
</tbody>
</table>

**Figure 10.** 549th Input and result of 5 models.
Figure 11. 293rd Input and result of 5 models.
Figure 12. 765th Input and result of 5 models.
Figure 13. 1969\textsuperscript{th} Input and result of 5 models.
Figure 14. 837th Input and result of 5 models.
5.3 RESULT AND ANALYSIS

GAN still has ambiguity about the method of evaluation. However, in recent years, many evaluation methods have been developed to evaluate the result of GAN. In [17], they introduced pros and cons of GAN evaluation method, which is composed of two types of evaluation. These are quantitative and qualitative measure. The quantitative measure is the calculation of specific numerical scores used to summarize the quality of generated images. The qualitative measure is human subjective evaluation or evaluation via comparison.

5.3.1 Quantitative measure

Five models were evaluated with quantitative measure. The representative image comparison evaluation method is introduced in quantitative measure, which are PSNR and SSIM. First, PSNR is a method for evaluating the quality of the generated image. The higher the PSNR, the better quality of the generated image. The evaluation method unit is dB [17]. Second, SSIM is compared with luminance, contrast, and structure of real and generated image. This evaluation method has advantage to catch distorted or blurry part of image. The SSIM score ranges between 0 and 1, which 0 means low similarity and 1 means high similarity [17,19]. Table 4 shows the results of each model, this is the average of the results of 258 test images. The ground truth scores are 1.0 (SSIM) and 361.202 (PSNR) respectively, and the closer the generated image is to the present RGB image, the score will be closer the ground truth score. P2P256_3Frame model shows higher SSIM score and P2P32_3Frame model shows higher PSNR score. Each of the two models showed better scores compared to pix2pix. The P2P32_3Frame’s PSNR scores were higher than
P2P256_3Frame in Fig 8, 10 and 12. However, the P2P256_3Frame seems to show better performance, because the P2P256_3Frame model with a higher SSIM and the P2P32_3Frame model’s PSNR score only showed difference with about 0.45 of PSNR score. Therefore, the P2P256_3Frame (proposed method) model performed better than pix2pix in thermal image colorization.

Table 4. SSIM and PSNR score of 5 model’s result.

<table>
<thead>
<tr>
<th></th>
<th>SSIM</th>
<th>PSNR (unit dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground Truth</td>
<td>1.0</td>
<td>361.202</td>
</tr>
<tr>
<td>Pix2pix</td>
<td>0.425293114</td>
<td>15.78640983</td>
</tr>
<tr>
<td>P2P32_10Frame</td>
<td>0.455655731</td>
<td>16.07653739</td>
</tr>
<tr>
<td>P2P256_10Frame</td>
<td>0.467947046</td>
<td>16.40540534</td>
</tr>
<tr>
<td>P2P32_3Frame</td>
<td>0.596821912</td>
<td>20.68152836</td>
</tr>
<tr>
<td>P2P256_3Frame</td>
<td><strong>0.649609842</strong></td>
<td>20.22756994</td>
</tr>
</tbody>
</table>

Also, paired student’s t-Test of statistical significance was used as a method to test the difference between the data of pix2pix and the results of four proposed methods. This method is not simply the average of SSIM and PSNR scores. The p-value can be judged probably to get a better percentage than pix2pix. P2P32_3Frame and P2P256_3Frame obtained a 0 of p-value, which means that the proposed method has a significant probability distribution than the SSIM and PSNR scores of pix2pix with 100% probability.
Table 5. p-value of proposed method.

<table>
<thead>
<tr>
<th>p-value</th>
<th>SSIM</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2P32_10Frame</td>
<td>8.3e-14</td>
<td>1.51e-3</td>
</tr>
<tr>
<td>P2P256_10Frame</td>
<td>1.91e-23</td>
<td>3.44e-9</td>
</tr>
<tr>
<td>P2P32_3Frame</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P2P256_3Frame</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
5.3.2 Qualitative measure

The qualitative measure is human subjective evaluation or evaluation via comparison [17]. This section will look at the experimental results image and compare the results of the proposed method with pix2pix.

This measure compare the color ambiguity problem. The object in the proposed method output image is properly colorized in the tree color of Fig. 8, 9, 10, 11, 12, the vehicle color of Fig. 8, 10, 13, 14, and the crosswalk on the road of Fig. 12. Proposed method properly colorized object on the road than pix2pix. Better results were obtained in previous 3rd frame RGB image input than previous 10th frame RGB image input. In previous 3rd frame RGB image, the problem of blurriness and distortion problem is much less. Therefore, the closer previous frame RGB image to the present frame RGB image can better colorize the thermal image.
6. CONCLUSION

Many studies of transforming thermal image into RGB image methods have made great efforts to solve the colorization problems. To solve colorization problems, unlike recent research, we used the previous frame RGB image. The RGB image is extracted color information to solve the colorization problems. Our proposed model using two input images and a generator can solve the colorization problem, which could be color ambiguous, blurry, and distort. And then the proposed method showed better performance thermal image colorization result compared to pix2pix.

Nevertheless, as seen in the experiment section, we are not still perfectly imitating with ground truth. We did four types experiments with our proposed method, but there was a problem of extracting the color information as well as the object structure of the previous frame RGB image. And we got better results than target model, but we have not completely solved the blurriness and distortion problems. To improve these parts in the future, we think that it can improve the colorization problem by using RL or object recognition [4] and GAN together.

Finally, road thermal image colorization is a valuable topic which research will be able to improve ADAS and AV system. Also, these studies will provide drivers with a lot of information in difficult situations when the road is hard to be recognized (e.g. darkness, fog, dust, and sun glare etc.).
REFERENCES


