FIRe-GAN: a novel deep learning-based infrared-visible fusion method for wildfire imagery

J. F. Ciprián-Sánchez1 • G. Ochoa-Ruiz2 • M. Gonzalez-Mendoza2 • L. Rossi3

Abstract
Wildfire detection is of paramount importance to avoid as much damage as possible to the environment, properties, and lives. In this regard, the fusion of thermal and visible information into a single image can potentially increase the robustness and accuracy of wildfire detection models. In the field of visible-infrared image fusion, there is a growing interest in Deep Learning (DL)-based image fusion techniques due to their reduced complexity; however, the most DL-based image fusion methods have not been evaluated in the domain of fire imagery. Additionally, to the best of our knowledge, no publicly available dataset contains visible-infrared fused fire images. In the present work, we select three state-of-the-art (SOTA) DL-based image fusion techniques and evaluate them for the specific task of fire image fusion, and compare the performance of these methods on selected metrics. Finally, we also present an extension to one of the said methods, that we called FIRe-GAN, that improves the generation of artificial infrared and fused images.

Keywords Image fusion • Fire • Wildfires • Deep learning • Visible • Infrared

1 Introduction

Wildfires can occur naturally or due to human activities and have the potential to get out of control and have a significant impact on the environment, properties, and lives. Recently, there have been several wildfires of significant proportions worldwide, such as the Australian wildfires of 2019 and 2020. CNN reported that the said fires took the lives of at least 28 people [1]. Another more recent example is the ongoing wildfire season in California in the USA. According to the BBC, as of September 17, 2020, 6.7 million acres have burned and more than 30 people have died [2]. Early wildfire detection enabling technologies are thus crucial in order to avoid as much damage as possible and to help firefighters in their endeavors.

Vision-based fire detection techniques can be divided into visible or infrared fire detection systems [3], according to the spectral range of the cameras employed. In operative scenarios, visible image-based systems display significant false alarm rates and missed detections. The latter is due to constraints present in different situations, such as changing environmental conditions and illumination [4]. In contrast, infrared (IR) cameras can perform flame detection in weak or no light conditions.

Furthermore, smoke is transparent in this type of images, as such, it can be advantageous from a practical perspective to employ infrared-based systems to perform flame detection in both daytime and nighttime conditions [3]. There has been work on fire detection on the near-
infrared (NIR) and long-wave infrared (LWIR) infrared bands. However, it is also not trivial to detect fire on infrared images, as they present problems such as thermal reflections and IR blocking [4].

The fusion of visible and infrared images can thus prove beneficial for improving the robustness, accuracy, and reliability of fire detection systems [3]. Although there have been some approaches in this area [5, 6], as well as for fire segmentation [7], to the best of our knowledge, DL-based visible-infrared fusion methods have not been tested for fire image fusion.

There is a growing interest in DL-based image fusion techniques. This is due to their reduced complexity compared to methods on the multi-scale transform and representation learning domains [8]. In order to assess the applicability of some of the most promising DL-based approaches in IR-visible fusion to wildfire image fusion, we chose three SOTA methods and evaluate their performance on fire image fusion. We also implemented extensions on one of them, generating the proposed FIRe-GAN method to improve its applicability to the Corsican Dataset. The selected methods are as follows: the method proposed by Li et al. [9], the work by Ma et al. [10], and the architecture proposed by Zhao et al. [11]. We evaluate and compare these methods on the Corsican Fire Database [12] using some of the most relevant metrics for assessing image quality found in the literature.

These state-of-the-art methods were selected because they present several desirable features. The method by Li et al. [9] uses a pre-trained VGG19 Deep Convolutional Neural Network (DCNN) as a part of its process for the fusion of detail content. Since the authors employ only selected layers of the network, no further training on new datasets (such as ours) is needed.

The method by Ma et al. [10] represents, to the best of our knowledge, the first approach towards image fusion through the use of Generative Adversarial Networks (GANs). This technique has the advantage of being end-to-end trainable, which significantly reduces its implementation and training complexity.

The method by Zhao et al. [11] is a GAN-based approach as well, with the additional feature of being able to generate approximate infrared images from visible ones. It is relevant to note that the type of infrared images (near-infrared (NIR), short-wavelength (SWIR), mid-wavelength (MWIR), or long-wavelength (LWIR)) that the model learns to generate depends on the types of images present in the training set.

Finally, it is relevant to note that many of the existing visible-infrared fusion methods [8–11, 13, 14] output grayscale fused images, which means that the color information of the visible image is lost. In the present paper, we present the FIRe-GAN model, an extended version of the method proposed by Zhao et al. [11], which allows for the processing of higher resolution images and the generation of color fused images as outputs. The latter is relevant due to color being one of the most used features in visible image-based fire detection methods [3].

The main contributions of this article are three-fold:

1. We carry out a thorough analysis and comparison of existing DL-fusion methods for conventional imagery.
2. We provide a quantitative demonstration of the feasibility of applying DL-based fusion methods for infrared imagery from wildfires.
3. We introduce a novel artificial IR and fused image generator that has been tested both in conventional and fire imagery.

We believe that these contributions can potentially boost developments in the wildfire fighting community that makes use of visible and infrared fire imagery to perform tasks such as detection and segmentation of wildfires for monitoring purposes more accurately and reliably. It must be noted that this work is part of a larger endeavour in which the proposed fusion methods plays only a small but vital role. The generation of infrared images, which is a component of this system, is to be used in other related research efforts and could prove to be useful for the research community at large.

The rest of this paper proceeds as follows. In Sect. 2 we present the three evaluated methods, the evaluation metrics, and the datasets employed. Section 3 describes the experiments and shows the evaluation results and a quantitative comparison of the selected methods. In Sect. 4 we provide a discussion on the obtained results. Finally, Sect. 5 concludes the paper and outlines potential future work.

2 Methods and data

2.1 Evaluated methods

2.1.1 Infrared and visible image fusion using a deep learning framework

This method, proposed by Li et al. [9] employs a DL framework for the generation of a single image that contains all the features present in visible and infrared images.

First, the authors decompose the original images into base parts and detail content. Then, Li et al. fuse these base parts through weight-averaging. For the fusion of the detail parts, the authors employ a DL framework in which they first use selected layers of a pre-trained VGG19 model to extract deep features [15]. Then, they use a multi-layer fusion strategy to extract weight maps. Next, the authors
use the deep features and weight maps to reconstruct the fused detail content. Lastly, they construct the final output image by combining the fused detail and base contents. Figure 1 depicts the framework of the method proposed by Li et al. [9] with a sample image pair from the Corsican Fire Database, and Fig. 2 shows the DL framework that Li et al. use for the fusion of detail content.

Finally, it is worth noting that the authors employ the fixed VGG19 network pre-trained in ImageNet, for the extraction of the multi-layer features. We will be referring to this method as the VGG19 method.

2.1.2 FusionGAN: a generative adversarial network for infrared and visible image fusion

This method, proposed by Ma et al. [10] introduces an image fusion method based on a Generative Adversarial Network (GAN). To the best of our knowledge, this work is the first to propose a GAN model for image fusion tasks. The architecture is an end-to-end model that generates fused images automatically from the source images without the need of defining fusion rules.

First, the generator attempts to produce a fused image with thermal information from the IR image and gradients from the visible one. The discriminator, in turn, forces the generated image to contain more details from the visible image. This process enables the model to produce fused images that retain both thermal and textural information. Finally, the authors generalize their proposed model for images with different resolutions, with the final image free of the noise caused by the upsampling of infrared information. Ma et al. named this model the FusionGAN model. Figure 3 shows the structure of the FusionGAN model with samples from the Corsican Fire Database.

2.1.3 Fusion of Unmatched Infrared and Visible Images Based on Generative Adversarial Networks

This method was proposed by Zhao et al. [11]. In this work, the authors propose a network model based on generative
adversarial networks (GANs) to fuse unmatched infrared and visible images. First, the visible image is given as an input to the generator $G_1$ to create a synthetic infrared image. Then, the visible image and the synthetic infrared image are concatenated and fed into generator $G_2$, generating the fused image as the output. The discriminator $D_1$ distinguishes between the source visible image and the generated fused image so that it is closer to the source visible one, containing more textural details. Simultaneously, the discriminator $D_2$ distinguishes the source infrared image, the generated infrared image, and the fused image. Through this updating cycle, the generated infrared image becomes closer to the source infrared image. This process enables the final fused image to contain more thermal information. We will be referring to this model as the UnmatchGAN model. Finally, Fig. 4 shows the structure of the UnmatchGAN model.

2.1.4 Summary

The previously mentioned methods have both advantages and disadvantages. The framework proposed by Li et al. [9] has the advantage of only needing some layers of an already pre-trained VGG19 network to perform feature extraction. As such, there is no need for further training for the particular application of image fusion. However, it is not an end-to-end method, and the required intermediate steps increase its implementation complexity.

The model presented by Ma et al. [10] has the advantage of being an end-to-end model, significantly reducing its implementation complexity. However, this GAN-based method needs to be trained on visible-infrared image pairs, and in consequence, its performance depends on the quality of the training process. It is also relevant to note that using GANs has the additional challenge of training stability [16].

The model proposed by Zhao et al. [11], a GAN-based model too, has the advantage of being an end-to-end procedure; however, the challenge lies on the training stability, as well as the need for a good training dataset. This method has the additional capability of learning to generate approximate infrared images based on source visible ones. Additionally, the fusion process requires perfectly aligned source images [17]. For the particular context of fire images, this could prove a significant advantage for the research community given the burden of obtaining perfectly matched visible-infrared fire images on realistic operative scenarios. Finally, it is also relevant to note that the three methods output grayscale fused images. In the context of fire imagery, the preservation of color could prove beneficial.

2.2 Data

For the present paper, we employ the visible-infrared image pairs of the Corsican Fire Dataset, first presented by Toulouse et al. [12]. This dataset contains 640 pairs of visible and near-infrared (NIR) fire images, alongside their corresponding ground truths for fire region segmentation.

We also employ the publicly available RGB-NIR dataset, developed by Brown et al. [18]. This dataset contains 477 non-fire visible-NIR image pairs. Figure 5 shows sample images from both datasets. Finally, it is relevant to remark that both datasets have NIR infrared images, that is, with a wavelength between 0.75 and 1.4 µm.

2.3 Metrics

The metrics selected for the present paper are the information entropy (EN), the correlation coefficient (CC), the peak signal-to-noise-ratio (PSNR), and the structural similarity index measure (SSIM); these metrics are by far the most common in the image fusion area, more details can be found in [19]. In the following subsections, we succinctly describe these metrics.
2.3.1 Information entropy

EN reflects the average amount of information in an image. It is defined as:

$$EN = \sum_{l=1}^{L-1} p_l \log_2 p_l,$$

where $L$ stands for the gray levels of the image, and $p_l$ represents the proportion of gray-valued pixels $i$ in the total number of pixels. The larger EN is, the more information is in the fused image [11].

2.3.2 Correlation coefficient

The CC measures the degree of linear correlation between the fused image and either the visible or infrared image. It is defined as:

$$CC(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}},$$

where $Cov(X, Y)$ is the covariance between the fused image and the reference images, and $Var(X), Var(Y)$ represent the variance of the two images. The larger the value of CC, the

Fig. 4 Structure of the model proposed by Zhao et al. [11]

Fig. 5 Sample images for the RGB-NIR and Corsican Fire Database datasets
higher the correlation between the fused and the reference images [11].

### 2.3.3 Peak signal-to-noise ratio

The PSNR assumes that the difference between the fused image and the reference image is noise. It is defined as:

\[
PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right),
\]

where \( \text{MAX} \) is the maximum value of the image color, and \( \text{MSE} \) is the mean squared error. An accepted benchmark for this metric is 30 dB; a PSNR value lower than this threshold means that the fused image presents significant deterioration [11].

### 2.3.4 Structural similarity index measure

The SSIM is a method for measuring the similarity between two images [20]. It is based on the degradation of structural information [21] and is defined as follows:

\[
\text{SSIM}(X, Y) = \frac{2u_x u_y + c_1}{u_x^2 + u_y^2 + c_1} \cdot \frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \cdot \frac{\sigma_{xy} + c_3}{
\frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} + c_3},
\]

where \( x \) and \( y \) are the reference and fused images, respectively; \( u_x, u_y, \sigma_x^2, \sigma_y^2, \) and \( \sigma_{xy} \) represent the mean value, variance, and covariance of images \( x \) and \( y \), respectively. Finally, \( c_1, c_2, \) and \( c_3 \) are small numbers that help to avoid a division by zero, and \( \alpha, \beta, \) and \( \gamma \) are used to adjust the proportions [11].

The range of values for SSIM goes from 0 to 1, with 1 being the best possible one.

### 3 Results

We evaluate the methods by Li et al. [9] and Ma et al. [10] through their corresponding open-source implementations. Li et al. and Ma et al. provide implementations of their methods pre-trained on the ImageNet and the TNO Dataset, respectively. In the following subsections, we introduce our proposed \( \text{FIRe-GAN} \) model, which extends on the \( \text{UnmatchGAN} \) one, and present the results of the comparison between the evaluated methods.

#### 3.1 Architectural adjustment for wildfire imagery fusion

Since the method proposed by Zhao et al. [11] has no available open-source implementation, we implemented it from the ground-up, extending it into our proposed \( \text{FIRe-GAN} \) model. We refer to the previous work by Isola et al. [22] for relevant implementation details on \( \text{G1} \) and the work by Ma et al. [10] for \( \text{G2} \) and both discriminators. We also modified the final layer of both generators from 1 to 3 filters; the latter allows our architecture to output 3-channel images. Since fire images are generally of a high size and

![Fig. 6 Implemented architecture for G1 of the method by Zhao et al. [11] with the mentioned considerations](image-url)
resolution, we made use of a U-Net architecture for G1 in contrast to the original method that makes use of a simple encoder-decoder architecture. Additionally, we integrated the Two Time-Scale Update Rule (TTUR) module proposed by Heusel et al. [23] and spectral normalization modules into the discriminators, as per the work by Miyato et al. [16] to increase the training stability.

Figure 6 shows the architecture of G1 with the original encoder-decoder architecture; Fig. 7 presents G1 with the proposed U-Net architecture; Fig. 8 shows the architecture of G2, and Fig. 9 the architecture of the discriminators, as per the said considerations.

To determine if the proposed U-Net architecture on G1 improves the quality of the generated infrared images, we pre-train both the original encoder-decoder and the proposed U-Net architectures of G1 on the RGB-NIR dataset and compare the obtained results for the generated infrared images. We split the RGB-NIR dataset into train and a validation sets. The training set contains 6112 image pairs after performing data augmentation, and the validation set consists of 96 image pairs. We train the model with a batch size of 4, 40 epochs, a learning rate for both generators of 5e-5 and for both discriminators of 1e-4, and spectral normalization on both discriminators. Additionally, the discriminators were updated once every two generator updates.

In Table 1, we present the average results for both architectures in terms of the selected metrics. In this case, the CC, PSNR, and SSIM metrics refer to the comparison of the source and generated IR images, whereas Fig. 10 displays sample images produced by both architectures. We can observe improvements on the EN, CC, PSNR, and SSIM metrics for our proposed U-Net architecture. A visual assessment of the produced images allows us to note an increased amount of detail as well.

Fig. 7 Implemented architecture for G1 with the proposed U-Net architecture

Fig. 8 Implemented architecture for G2 with the mentioned changes
3.2 Method comparison

Due to the improvement displayed by the U-Net architecture for G1 on the proposed FIRe-GAN model, we took this extended model and used it for its comparison with the works by Li et al. [9] and Ma et al. [10]. For consistency and to make the comparison fair with these methods, we pre-trained our proposed FIRe-GAN model with the RGB-NIR dataset. Then, we tested the three models on the Corsican Fire Database. In this way, we were able to assess the generalization capabilities of these models on the new domain of fire imagery.

Figure 11 shows the results of these experiments, while Table 2 summarizes the average results on the first three columns, and Fig. 12 shows sample images produces by the three methods. We can observe that the VGG19 method presents the most balanced inclusion of information and, on average, higher quantitative results on the evaluated metrics. Of the GAN-based models, the FusionGAN heavily favors the inclusion of thermal data, while the FIRe-GAN model shows balanced results regarding the inclusion of source information. However, the metric results are lower on average than those of the VGG19 method.

3.3 Transfer learning

The VGG19 and FusionGAN methods can rely on source infrared images to perform the fusion process. In contrast, the UnmatchGAN model and, by extension, the proposed FIRe-GAN model, must generate approximate infrared ones and fuse them with the source visible ones. Then, it stands to reason that it could be more problematic for this model to generalize to new domains. As such, we perform a transfer learning phase with a split of the fire images of the Corsican Fire Database, and then we evaluate its performance.

We split the dataset into a train and a validation sets. The training set has 8192 images after data augmentation, and the validation set consists of 128 image pairs. We set the training epochs to 3. Due to the strong thermal characteristics of fire images, we add a constant term $c$ that multiplies the element of the loss function of $G2$ that represents the adversarial loss between $G2$ and $D1$, thus prioritizing the inclusion of visible information. The final result is a balanced inclusion of visible and infrared information for the fire fused images. Experimentally, we found the best value of $c$ to be 4.5. All other training parameters are the same as those mentioned in Sect. 3.1, and all other loss functions remain the same as those of the UnmatchGAN model. The modified loss function for $G2$ is as follows:

![Fig. 9 Implemented architecture for both discriminators with the mentioned changes](image)

![Fig. 10 Sample resulting synthetic infrared images from both architectures for G1](image)

| Model  | EN   | CC  | PSNR | SSIM |
|--------|------|-----|------|------|
| Original | 9.9158 | 0.8593 | 17.9341 | 0.5506 |
| U-Net   | 9.9474 | 0.9203 | 19.9473 | 0.7541 |
Results from the fire images on the three evaluated methods

Table 2  Average results for the three evaluated methods on the full Corsican Fire Database as specified on Sect. 3.2 on the first three columns, and for the FIRe-GAN method before and after transfer learning on the validation set of the Corsican Fire Database (Sect. 3.3) on the last two columns

| Metric         | VGG19  | FusionGAN | FIRe-GAN | Before | After |
|----------------|--------|-----------|----------|--------|-------|
| EN             | 6.3415 | 6.0072    | 10       | 10     | 10    |
| CC IR-fused    | 0.7976 | 0.9657    | 0.5650   | 0.5774 | 0.7270|
| CC RGB-fused   | 0.7637 | 0.4104    | 0.6651   | 0.6799 | 0.6499|
| PSNR IR-fused  | 19.2014| 22.5090   | 14.2065  | 14.3196| 16.4212|
| PSNR RGB-fused | 19.2255| 15.3900   | 16.7435  | 16.7687| 15.6484|
| SSIM IR-fused  | 0.8337 | 0.8389    | 0.6129   | 0.6187 | 0.7064|
| SSIM RGB-fused | 0.9007 | 0.7783    | 0.8000   | 0.8020 | 0.7935|
\[
L_{G_2} = \gamma \left[ \frac{1}{N} \sum_{n=1}^{N} (D_1(I^n_F) - c_1)^2 \right] + \frac{1}{N} \sum_{n=1}^{N} (D_2(I^n_F) - c_2)^2 \\
+ \frac{\lambda}{HW} (\|I_F - I_R\|^2_F + \xi \|\nabla I_F - \nabla I_V\|^2_F),
\]

where \( N \) represents the number of fused images, \( D_1(I^n_F) \) the output of \( D_1 \) for the fused image \( I^n_F \), \( D_2(I^n_F) \) the output of \( D_2 \) for the fused image \( I^n_F \), \( c_1 \) and \( c_2 \) the values that \( G_2 \) wants the discriminators \( D_1 \) and \( D_2 \), respectively, to believe for fake data (i.e., soft labels), \( H \) and \( W \) the height and width of the input images, respectively, \( \|I_F - I_R\|^2_F \) the square of the Frobenius norm of the difference between the

Fig. 12 Sample resulting images from the three methods. In column 12a are the RGB images, in 12b the IR images, in 12c the fused images from the VGG19 method, in 12d the fused images from the FusionGAN method, and in 12e the fused images from the FIRe-GAN method.

Fig. 13 Results for the FIRe-GAN model from the fire images of the validation set before and after transfer learning.
fused image $I_F$ and the source infrared image $I_R$, $\nabla I_F - \nabla I_V$ the difference between the gradients of $I_F$ and the source visible image $I_V$, and $\xi$ a positive number that controls the trade-off between the term $\|I_F - I_R\|^2_F$ and the term $\|\nabla I_F - \nabla I_V\|^2_F$.

In Fig. 13, we show the results on the 128 images of the validation set before and after the transfer learning phase. Table 2 shows the condensed average results on the last two columns. After only three training epochs, we can observe a marked improvement in the generated NIR images, as well as more accurate inclusion of thermal information on the fused ones.

3.4 Summary

In Table 2 we summarize the average results both for the evaluation of the three methods on the full Corsican Fire Database, as specified in Sect. 3.2, and of the transfer learning results on the validation set of the Corsican Fire Database as specified on Sect. 3.3.

4 Discussion

Of the three evaluated methods, the VGG19 one by Li et al. [9] displayed the best overall performance for the new domain of fire imagery. This model is also the one that shows a more balanced behavior in terms of the inclusion of both visible and thermal information. This can be attributed to the way the authors leverage the VGG19 DL model. As Li et al. employ only the feature extraction capabilities of a pre-trained model, they do not need to train it for the particular task of image fusion. As the model was pre-trained on ImageNet, it demonstrates significant feature extraction capabilities, which explains the superior performance.

The FusionGAN model proposed by Ma et al. [10] is relevant since it is the first attempt to use GANs for image fusion. The simplicity of an end-to-end model is also desirable. However, when applied to the new domain of fire images, this method tends to incorporate more thermal than visible information. This can be due to the fact that fire images have more well-defined thermal information, whereas non-fire images in the training set do not exhibit that strong of a distinction between thermal and visible images.

Our proposed FIRe-GAN model has the advantages of being able to work with higher resolution images and to output three-channel fused images. This last feature allows it to learn to preserve color. Before performing transfer learning, it shows a balanced approach toward the inclusion of visible and thermal information; however, the overall performance is lower compared to the other two methods. Also, the generated IR images are very close to the source visible images; the model does not compensate for thermal information hidden behind features like smoke. Upon visual inspection, we can observe that the fused images preserve colors similar to the visible ones.

When applying transfer learning to our proposed method on a segment of the Corsican Fire Database, after only three training epochs, the model can produce artificial NIR images that are very close to the original ones, and fused

![Figure 14](image_url)
images containing a balanced combination of thermal and visible information. This speaks well of the specialization capabilities of the model. It is also relevant to note that, even though the fused images preserve color information, this color is no longer the same as the visible images, particularly on the fire regions. Since the color of the source visible and infrared images are significantly different for the case of fire images, it appears to be that the model learns to seek an intermediate color representation between the two.

Finally, it is worth recalling that for the present study, we employ only near-infrared (NIR) images, i.e., IR images with a wavelength between 0.75 and 1.4 μm. In this regard, the models themselves do not distinguish between different types of IR images (e.g., NIR, LWIR, among others). Thus, we would expect them to extend seamlessly to different types of IR images if a suitable dataset is provided, although their performance could change for these different types of IR images.

5 Conclusions and future work

In this paper, we demonstrated the feasibility of DL-based methods for the particular task of fire image fusion. In our experiments, the framework proposed by Li et al. [9] displayed the best overall performance. This method takes advantage of the feature extraction capabilities of DCNNs and traditional image fusion techniques for an effective combination.

The evaluated GAN-based methods show promise due to the simplicity of their implementation and their generalization and specialization capabilities. In particular, our proposed FIRE-GAN model displays a balanced approach toward the inclusion of visible and infrared information, with consistent color preservation. Also, it is relevant to note that there is no much data available by DL standards, of visible-infrared image pairs, especially on the fire domain. Therefore, the generation of more visible-infrared image pairs would improve the performance of the models and reduce the risk of overfitting. Finally, further experimentation is needed to assess the significance of color preservation on the fused images for different fire detection techniques.

Acknowledgments The authors would like to thank the University of Corsica for providing access to the Corsican Fire Database.

Funding Information This research is supported in part by Tecnologico de Monterrey and the Mexican National Council of Science and Technology (CONACYT). This research is part of the project 7817-2019 funded by the Jalisco State Council of Science and Technology (COECYTJAL).

Availability of data and material The image fusion methods by Li et al. [9] and Ma et al. [10] are publicly available as Github repositories in [15, 24], respectively. The RGB-NIR dataset employed for pre-training the method by Zhao et al. [11] was developed by Brown et al. [18] and is publicly available at [25]. The Corsican Fire Database [12] is available upon request to the University of Corsica at [26].

Declaration

Code availability The code generated by the authors implementing FIRE-GAN, an extended version of the image fusion method proposed by Zhao et al. [11], is available as an open-source Github repository at https://github.com/JorgeFCS/Image-fusion-GAN.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Yeung J Australia’s deadly wildfires are showing no signs of stopping, here’s what you need to know. https://edition.cnn.com/2020/01/01/australia/australia-fires-explainer-intl-hnk-scli/index.html. Accessed: 2020-04-10
2. The Visual and Data Journalism Team: California and Oregon 2020 wildfires in maps, graphics and images. https://www.bbc.com/news/world-us-canada-54180049. Accessed: 2020-09-30
3. Yuan C, Zhang Y, Liu Z (2015) A survey on technologies for automatic forest fire monitoring, detection and fighting using uavs and remote sensing techniques. Canadian J Forest Res 45:150312143318009. https://doi.org/10.1139/cjfr-2014-0347
4. Çetin AE, Dimitropoulos K, Gouverneur B, Grammalidis N, Güney O, Habiboglu YH, Töreyin BU, Verstockt S (2013) Video fire detection—review. Digital Signal Process 23(6):1827–1843
5. Arrue BC, Ollero A, Matinez de Dios JR (2000) An intelligent system for false alarm reduction in infrared forest-fire detection. IEEE Intell Syst Their Appl 15(3):64–73. https://doi.org/10.1109/5254.846287
6. Martinez-de Dios J, Arrue B, Ollero A, Merino L, Gómez-Rodríguez F (2008) Computer vision techniques for forest fire perception. Image Vision Comput 26(4):550–562
7. Nemalidinne SM, Gupta D (2018) Nonsubsampled contourlet domain visible and infrared image fusion framework for fire detection using pulse coupled neural network and spatial fuzzy clustering. Fire Saf J 101:84–101
8. Li H, Wu XJ, Kittler J (2020) Mdlatlr: a novel decomposition method for infrared and visible image fusion. IEEE Trans Image Process 29:4733–4746. https://doi.org/10.1109/TIP.2020.307975984
9. Li H, Wu X, Kittler J (2018) Infrared and visible image fusion using a deep learning framework. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp 2705–2710
10. Ma J, Yu W, Li X, Ji C, Jiang J (2019) Fusiongan: a generative adversarial network for infrared and visible image fusion. Information Fusion 48:11–26
11. Zhao Y, Fu G, Wang H, Zhang S (2020) The fusion of unmatched infrared and visible images based on generative adversarial networks. Math Probl Eng 2020:3739040
12. Toulouse T, Rossi L, Campana A, Celik T, Akhlioufi MA (2017) Computer vision for wildfire research: An evolving image database for processing and analysis. Fire Saf J 92:188–194
13. Zhao Z, Xu S, Feng R, Zhang C, Liu J, Zhang J (2020) When image decomposition meets deep learning: A novel infrared and visible image fusion method
14. Ma J, Liang P, Yu W, Chen C, Guo X, Wu J, Jiang J (2020) Infrared and visible image fusion via detail preserving adversarial learning. Information Fusion 54:85–98
15. Infrared and visible image fusion using a deep learning framework. https://github.com/hli1221/imagefusion_deeplearning. Accessed: 2020-08-22
16. Miyato T, Kataoka T, Koyama M, Yoshida Y (2018) Spectral normalization for generative adversarial networks
17. Ma J, Ma Y, Li C (2019) Infrared and visible image fusion methods and applications: a survey. Information Fusion 45:153–178
18. Brown M, Süssstrunk S (2011) Multi-spectral sift for scene category recognition. In: CVPR 2011, pp 177–184
19. Liu Z, Blasch E, Xue Z, Zhao J, Laganiere R, Wu W (2012) Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study. IEEE Trans Pattern Anal Mach Intell 34(1):94–109. https://doi.org/10.1109/TPAMI.2011.109
20. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP The ssim index for image quality assessment. https://www.cns.nyu.edu/~lcv/ssim/. Accessed: 2020-08-26
21. Wang Zhou, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
22. Isola P, Zhu J.Y., Zhou T, Efros A.A. (2017) Image-to-image translation with conditional adversarial networks. CVPR
23. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S (2017) Gans trained by a two time-scale update rule converge to a local nash equilibrium. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS’17, p. 6629-6640. Curran Associates Inc., Red Hook, NY, USA
24. Codes for fusiongan, a gan model for infrared and visible image fusion. https://github.com/jiayi-ma/FusionGAN. Accessed: 2020-09-13
25. Rgb-nir scene dataset. https://ivrlwww.epfl.ch/supplementary_material/cvpr11/index.html. Accessed: 2020-12-17
26. Corsican fire database. http://cfdb.univ-corse.fr/. Accessed: 2020-12-17

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.