In: 2023 5th International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 20-23 February 2023, Bali, Indonesia
DOI: 10.1109/ICAIIC57133.2023.10067042

Multi-Spectral Fusion using Generative Adversarial Networks for UAV Detection of Wild Fires

1st Tanmay Kacker
Sch. of Aerospace Transport Manufacturing
Cranfield University, Bedford, UK
tanmay.kacker.465@cranfield.ac.uk

2nd Adolfo Perrusquia
Centre for Auto. Cyberphysical Sys.
Cranfield University, Bedford, UK
Adolfo.Perrusquia-Guzman@cranfield.ac.uk

3rd Weisi Guo
Centre for Auto. Cyberphysical Sys.
Cranfield University, Bedford, UK
weisi.guo@cranfield.ac.uk

***Abstract*—Wild fires are now increasingly responsible for immense ecological damage. Unmanned aerials vehicles (UAVs) are being used for monitoring and early-detection of wild fires. Recently, significant research has been conducted for using Deep Learning (DL) vision models for fire and smoke segmentation. Such models predominantly use images from the visible spectrum, which are operationally prone to large false-positive rates and sub-optimal performance across environmental conditions. In comparison, fire detection using infrared (IR) images has shown to be robust to lighting and environmental variations, but long range IR sensors remain expensive. There is an increasing interest in the fusion of visible and IR images since a fused representation would combine the visual as well as thermal information of the image. This yields significant benefits especially towards reducing false positive scenarios and increasing robustness of the model. However, the impact of fusion of the two spectrum on the performance of fire segmentation has not been extensively investigated. In this paper, we assess multiple image fusion techniques and evaluate the performance of a U-Net based segmentation model on each of the three image representations - visible, IR and fused. We also identify subsets of fire classes that are observed to have better results using the fused representation.**

***Index Terms*—fire detection, deep learning, UAV, drone, GAN**

I. Introduction

A wildfire is an unplanned and uncontrolled fire that burns in natural areas such as grasslands and forests. There have been 62 major wild fires on an average every year for the past decade. Wild fires (mainly in forests) pose immense social, economic and ecological challenges to countries around the world. For example, in the state of California, in the USA, had worsening wildfires in the season of 2017 and 2018 that led to a total financial loss of $ 40 billion. The Camp Fire in 2018 led to the death of 88 people, burning of 8.7 million acres of land and destruction of 18,500 buildings and other structures. There is increasing severity of forest fires in the past decade, both in the sheer number of forest fires increased and the affected acres of area burnt.

Combustion when it has the necessary requirements of fuel, oxygen and a source of ignition like heat. The spread of fire is based on a chain reaction when would depend on other environmental factors like geography, wind etc. This is known as the fire triangle. The spread of a fire in a forest is more likely due to the abundant presence of oxygen and combustible materials like tress and grass. The spread also reduces the chances of mitigating a fire and even what strategies to employ. Van Der Werf et al. [20] identified that the rate of spread, the fire intensity can greatly impact forest fires. It also means that a fire should be tackled or even detected before it reaches the fourth stage of crown fires. Such fires can spread extremely fast and cause widespread destruction.

We would like to thank funding from EC H2020 DAWN4IoE - Data Aware Wireless Network for Internet-of-Everything (778305). All data used in this paper is openly available.

A. State of the Art & Gaps

1) Image Segmentation for Early Fire Detection: Fire segmentation as opposed to fire detection is more relevant due to the fact that we can obtain pixel by pixel information. This helps us in predicting various attributes of a fire like spread, position, rate of growth and volume. These statistics can be extremely useful for predicting the growth of fire and how it can be mitigated. Along with fire, smoke detection is also equally important because smoke can be detected from larger distance and can be used when the fire is not even visible. Hence, fire and smoke segmentation provides more granular insights than simple classification where the system will only alert us about the presence of fire but not impart any other characteristics that would be useful in tacking them.

Rule-based approaches use image processing techniques along with pixel values in different colour spaces to establish a rule-set to generate segmentation masks. Chen et al. [5] were the first to propose an early fire detection approach using chromatic analysis to extract fire and smoke pixels from images with the following observations: (1) Fire generally display colours from red to yellow, (2) Red values should be greater than a threshold RT since red is the major component and fire also emits light, and (3) Background illumination could lead to false positive results due to either presenting fire similar aliases or negatively impact the saturation of fire. Qian et al. [16] improved the rule-set by modifying the decision function to use average saturation instead of the difficult to tune saturation threshold. The proposed change removes the sensitivity of the approach to environmental conditions.

Deep-learning approaches use Convolutional neural networks (CNNs) such as ResNet [9] and VGG [18] architectures, with strong power in representation and solving inverse problems [7]. He et al. [9] investigated the effects of the vanishing gradient problem and proposed the usage of residual blocks in CNN. Such as model is the ResNet model which has demonstrated good performance with segmentation tasks and is even robust to perturbations. Other such architectures using residual blocks are the FusionNet [17] and faster region based CNN F-RCNN are also robust and high performance when it comes to segmentation tasks. Such architectures suffer from a large memory footprint, especially due to the extensive use of skip connections which makes it infeasible use such models in a system using UAVs which have practical limitations on available memory. Bulo et al. [2] suggested the usage of multiple layers of batch normalisation with an activation function - In-place activated batch normalisation model (IN-PLACEABN), which attempts reduce the memory consumption of the model by reducing the range of the incoming features. Multiple such architectures have even been combined using multiple streams such as the FRRN [15] model which uses full resolution residual networks by having a ResNet model in one stream and a VGG model. The full resolution or residual stream is used for identifying and respecting the boundaries of the segments while the other stream undergoes pooling and unpooling sequence of operations for feature extraction.

Despite these advances, these approaches [21] have certain limitations: (1) the CNN only performs binary classification. It can only identify if a fire is present in the image or not, (2) the rule-based approach is similar to the one proposed by Celik et al. [3] – it does not contain rules specific for smoke segmentation and can also be enhanced by adding rules for fused images, and (3) the dataset used was small and cropped around the fire pixels. It did not contain significant variations in environment and illumination to be viable for real-time forest fire detection applications.

2) Multi-Spectral Fusion for Robust Detection: On the basis of the spectral range, fire and smoke segmentation systems generally employ images either from the visible or the IR spectrum. The visible images have high spatial resolution and contain textural information. [19] identified that segmentation models are sensitive to textural and colour variations such as different illumination and environmental conditions. Varying amounts and colour of smoke also impact the performance of the model. Hence, the segmentation model using visible images is prone to high false positive rates Since the fire has large difference in temperature when compared to the its surroundings, fire pixels can be easily distinguished. Systems using IR cameras are also not prone to lighting conditions. However, smoke is not discernible in IR images which does improve the performance of fire segmentation but discounts systems requiring smoke segmentation. Reflections can significantly impact the performance of the model. The fusion of the images of the two modalities or spectrum poses as viable option where the fused representation borrows characteristics from both the modalities. Deep learning based image fusion techniques present a far simpler and robust alternative to representation learning. As shown by Yuan et al. [22], the fused representation of the visible and the IR images yield outputs that could potentially retain both the textural and thermal properties of the original modalities.

Techniques using traditional image processing methods (e.g., non-sub-sampled contourlet transform (NSCT) with a pulse-coupled NN (PCNN) [12], [14]) are high performance but are highly complex and sensitive when it comes to variations. DL-based image fusion techniques have garnered notable interest due to their robustness and simplicity compared to the complex techniques such as multi-scale transforms. Li et al. [11] propose a DL-based framework for fusing images from the visible and IR spectrum. Firstly, the images undergo decomposition to be split into base parts and detail content.

The advantage of using Generative Adversarial Network (GANs) for image fusion is that they can be trained end-to-end that can reduce the complexity of implementation and training. It can generate samples automatically from original images without the need to implement fusion rules. Both the models here are simple 5-layered CNNs. Zhao et al. [23] also uses GAN to fuse images from the visible and IR spectrum across wavelengths. Similar to method suggested by Ma et al. [13], the multi-layered GAN is an end-to-end method. The authors report the best structured similarity (SSIM) score of 0.83 amongst other metrics on the TNO data set and around 0.94 the VEDAI dataset. The method requires the incoming visible and IR images to be aligned. The method also has the benefit of generating IR images which can be further used. GANs have a demonstrated that they have do not show training stability which will have to be dealt with.

B. Innovation

Image fusion has not been used for fire segmentation. Indeed it is much more challenging as fire is a very dynamic and not well defined imagery process, unlike prior work on infrastructure damage for example [8], [10]. It has not been observed or established whether the performance of the segmentation model would be enhanced by using fused fire images.

Here, we propose to solve these challenges by the following innovations:

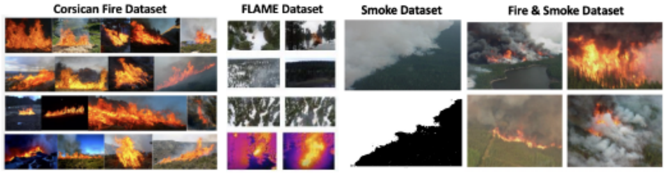
- Use GANs to improve multi-spectral representation of fires. The models will have to be modified to give multi-channel output and training stability will have to be improved.
- Fuse GAN generated IR with true visual data to produce more reliable fire early detection
- Compare our approach against the three techniques developed by Li et al. [11], Ma et al. [13] and [23].

We believe the GAN approach here are end to end trainable and the last method has the added benefit of generating IR images which can be extremely useful for a UAV.

II. METHODOLOGY

A. Data Sets & Pre-Processing

There has been substantial research conducted for fire and smoke detection which is evident from the amount of data that is available. However, most of these datasets focus on classification rather than segmentation. Our task was to



Name	Type	Classes	Tags	Size
Corsican Fire Dataset	RGB Images, Infrared	Fire, Smoke	Segmented	600
Fire and Smoke 2 Dataset	RGB Images	Fire, Smoke, Control	Annotated	3831
MIVIA Fire Dataset	RGB Videos	Fire, Control		31
FireNet	RGB Images, RGB Videos	Fire, Control		46
Center for Wildfire Research	RGB Images	Smoke	Segmented	49
Smoke dataset	RGB Images, RGB Videos	Smoke	Segmented	5000
FLAME Dataset	RGB Videos, Infrared	Fire, Control	Segmented	2000
AI for Mankind	RGB Images	Smoke		150

Fig. 1. Datasets of Wild Fires.

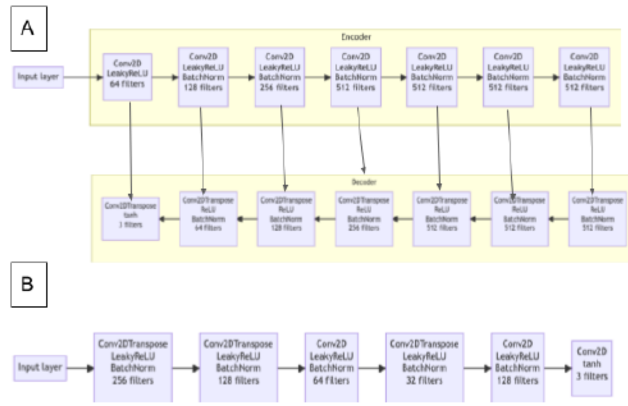


Fig. 2. Architecture of (a) GAN 1, and (b) GAN 2.

identify relevant datasets for fire and smoke segmentation. This still yielded a large list of datasets. Next filtering was done on the basis of the classes present. Most of the datasets focus on the fire classification or segmentation, while smoke is generally not identified. We are also looking for images taken from a high camera angle and set in a forest setting. Images from urban settings tend to overfit the model to background details. The images need to have significant variations across illumination, background, fire spread etc. Figure 1 lists the most viable datasets that we found. The next few sections list the chosen datasets along with their information.

The Corsican dataset is a standard dataset when it comes fire and smoke detection and is maintained by the Corsican University. It contains 600 images of fire and smoke across different background settings. It also contains images from the IR spectrum and hence is extremely useful for testing our hypothesis. It also contains segmentation ground truths which can be used for semantic segmentation. The only issue of the dataset (see Figure 1) is that it does not contain images taken from a high angle rather most of the images come from focused shots of the fire.

The FLAME dataset comprises of images shot from a UAV surveying fire around the North Arizona region in the USA. It contains around 40000 frames which are each labelled as fire or not. It also contains around 2000 frames specifically for the task of segmentation. There exists segmentation masks or ground truths for each of these 2000 frames. They also contain images taken from IR spectrum. However, the lack of large fire bulbs could hamper the performance of the model and hence the images are cropped to increase the fire pixel ratio to around 15%. The analysis of luminosity also reveals that the pixels are bright enough while lesser than that of the Corsican dataset. This dataset matches our requirements however, the dataset does not significant variations in frames and is taken during the winter and hence does not represent other geographies.

The smoke dataset is maintained by Chen et al. [4] and was used for their work on smoke detection using DeepLab architecture. The dataset contains around 5000 images specifically targeting smoke and also has the segmentation masks. For further analysis as shown in Figure 1, we conduct the luminosity comparison of the dataset which reveals to be more balanced than the others. This can be attributed to the quality of smoke since it is not as bright as fire. It can be seen there a quite a few images in the lower ranges of luminosity.

The Fire and Smoke dataset is crowdfunded dataset with images from varying background settings. It contains around 5000 images with fire and smoke. It contains the annotations and ground truths required for segmentation. We chose a non-urban subset of the dataset which comprises of around 1500 images which are in forest or rural setting. An example of chosen images can be seen in the Figure 1.

B. System Overview

1) Segmentation: Akhloufi et al. [1] proposed the modification seen in 3.6 for the U-Net architecture. It is known as the DeepFire model and it outputs a binary segmentation mask detecting dire pixels. The images are down-sampled through four pooling layers with a kernel of size 1. The activation function used to generate the mask is the sigmoid function. The model has extremely low number of trainable parameters i.e. 2 million and thus has faster training and inference times. Since our problem definition is based on UAVs, using such a lightweight model reduces the load on the system resources. The authors achieved the best F1 score of 0.99 on the Corsican dataset.

2) IR Generation GAN: We are inspired by the implementation by Ma et al. [13], the generator G1 is DCNN is an encoder-decoder network. Our implementation of G1 uses the U-Net architecture. The final tanh activated layer?s dimensionality of output space was modified to 3 to allow the generation of 3-channel output images. The same changes were made to the other generators. Similar to Ciprian-Sanchez et al. [6], the input for each layer was modified to use spectral normalisation to improve training stability. As seen in Figure 2, the model G1 has:

- 7 convolutional layers for each of the encoder-decoder blocks followed by a single transpose convolutional activated by the tanh function with 3 filters and a 4x4 kernel and step size of 2.

- Fuse GAN generated IR with true visual data to produce more reliable fire early detection
- Compare our approach against the three techniques developed by Li et al. [11], Ma et al. [13] and [23].

and G2 has:

- An input layer of shape 384 x 512 x 6.
- 5 convolutional layers down-sampling the input to 3 channels. The first 2 layers use a 5 x 5 kernel, the two next use a 3 x 3 while the final layer has a 1 x 1 kernel.
- All layers are spectral normalised, are batch normalised apart from the output layer, and apart from the last layer use the LeakyReLU activation function. The last layer uses tanh.

Both the discriminators D1 and D2 use the same network structure. The U-Net model was pre-trained on the ImageNet and RGB-NIR datasets. The model structure is as follows:

- An input layer of shape 384 x 512 x 6.
- A final dense layer which output the binary classification label
- For middle convolutional layers with the LeakyReLU activation function and a filter of size 3 with step size 2.
- All layers are batch normalised apart from the first layer.

C. Training and Cross-Validation

The models are trained on the datasets containing multi-spectral images specifically, the Corsican, FLAME, and FIRE-2 datasets. The entire dataset comprises of 2032 pairs of images for each of the visible and IR spectrum with images of varying levels of complexity. The images were zoomed, flipped along with other translations. The training along with other experiments were conducted on Google Colab. The CPU of the instance was an Intel Xeon CPU with 2.20GHz and the GPU was Tesla GPU P100 which has 16 memory and 3584 cores. The training parameters were subjected to hyper-parameter optimisation using cross-validation and the various choices are listed in Table 5.1. The batch size was set to 32 and Adam was chosen as the optimiser. The learning rate was set to 10?4 and for every two times a discriminator was trained, the generator was trained once i.e the discriminator training step. The adjustment factors ? and ? are set to 120 and 10 respectively. The discriminators converged in 24 training epochs and hence the generators were finalised in 12 epochs. The average training and validation losses were 0.12 and 0.18.

III. Results

A. Multi-Spectral Fusion

As mentioned, the fusion of images was tested on the Corsican, Flame and Fire datasets. The IR and visible images generated the fused images which can be seen in the Figure 3 and 4. It is noticed that the VGG method led to the most balanced images where both of thermal and textural details are included. This is also confirmed via the quantitative

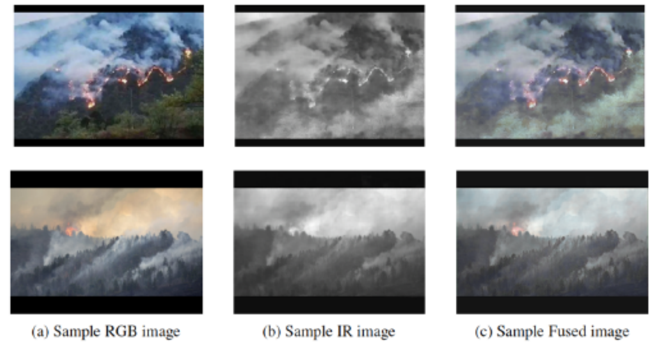


Fig. 3. Results from GAN enabled Multi-Spectral Fusion using the Fire dataset.

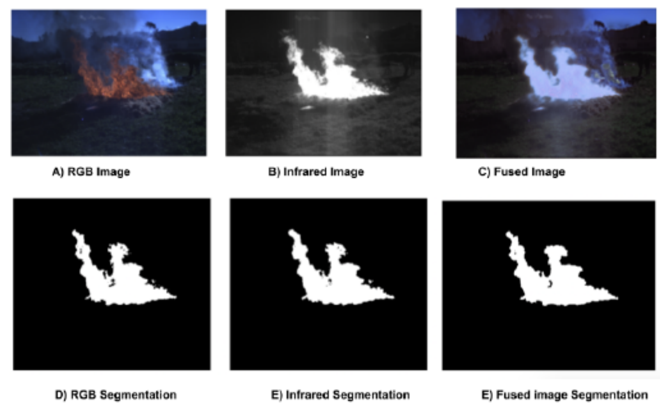


Fig. 4. Comparison between GAN enabled Multi-Spectral Fusion Methods.

results. Our solution, which is an adapted version of the GAN by [23] adds multi-channel output along with spectral normalisation. The GAN by Ma et al. [13] has much more thermal information than the textural details. Zhao et al. [23] even led to a balanced inclusion of thermal information then the textural details but lesser than the VGG method. This can be attributed to the fact that fire images generally have much more discernible thermal information than textural details. Images that do not contain fire do not possess such discernible thermal data. It can also be seen that the images using our method are high resolution and multi-channel rather than the single channel output of the other methods. It is also observed that the fused images are capable of retaining proper colour information. The model has been able to specialise due to the balance between thermal and textural information. An interesting thing to note the coloured fused images are not the same as the images from the visible spectrum rather they have an intermediate representation of the colour information which indicates a balance between the two sources.

The model was evaluated across the datasets on 5 metrics that are EN, SD, CC, PSNR and SSIM. The computed values of the metrics. It can be seen that across the various methods, our proposed method has the best average values for most of the metrics. We now look at and analyse these metrics

Dataset	F1	Accuracy	Precision	Recall	FN rate	FP rate
Corsican	0.75	0.79	0.70	0.92	0.05	0.05
FLAME	0.24	0.68	0.14	0.79	0.02	0.35
FIRE	0.68	0.79	0.52	0.94	0.04	0.26
Total	0.55	0.75	0.44	0.94	0.03	0.32

Fig. 5. Results from using the combined approaches.

across the methods. The multi-layered GAN has the largest information entropy (EN) of 9.8, that means the methods transferred the most information from the original images to the generated image. The VGG method outperforms our approach when it comes to CC where VGG gets a score 0.72 while our methods gets a score of 0.62. Here, the single layered GAN performs the worst. It should be noted that this is comparing fused to the original RGB image and not to the IR. When comparing IR images, the GAN by Ma et al. [13] performs the best actually.

It can be seen that all methods perform similarly on the SSIM metric however the VGG method does perform the best here with a score of 0.9. It means when it comes to visible images which means that our method generated images which were similar to the original visible images. Our method however shows lower PSNR values when compared to other methods and considering the IR image indicating deterioration. However, it performs similarly when considering fusion of images with an RGB image.

Looking at the results, it can be concluded that the proposed methods scores the best on nearly all the metrics considered. The VGG method tends to show more balance but our proposed solution has the largest information and structural similarity. Our method gives the best entropy meaning a lot more information is encompasses in the image compared to other methods. Also, our method has the advantage of being multi-channel and end to end trainable. Our method however scores less on metrics comparing the fusion with respect to original IR images. This is expected since we adjusted the loss function with a parameter than motivates more visible information to be included.

B. Segmentation and Detection

1) Rule-based: For generating the segmented masks we compared the following methods: (1) the method proposed by Qian et al. [16] which is an adapted version of the method by Chen et al. [5], and (2) the YbCbCr method proposed by Celik et al. [3]. It can be seen in the Figure 5 that the algorithm yields the lowest false-negative rate across all the methods and datasets. It is interesting to see that it is independent of the dataset used and hence suggesting variations in image complexity do not impact it. However, the results for other metrics degraded since the two methods now combine to identify a pixel. The accuracy has decreased to 0.75. The false positive rate has also increased indicating the two methods combine can accurately tell if a pixel is not fire but will also give results where the pixel was wrongly segmented as a fire pixel. These cases will be handled by our CNN layer.

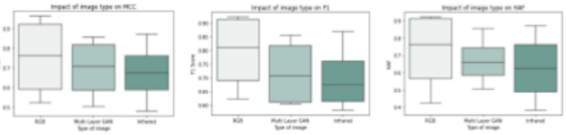
The method for generating masks for smoke pixels was adapted from Celik et al. [3]. The smoke threshold was set to 25 for the datasets. The results for the segmenting the smoke pixels can be seen in Table 6.4. It can be seen that the performance of the algorithm is much worse for segmenting smoke. The false positive and false negative rates are extremely high while the accuracy is extremely low. Rule-based approaches are good for pre-selection of fire but not smoke since smoke is extremely dynamic and has varying shapes and colours.

2) Deep Learning: Figure 6 shows the segmentation of a few sample images from our dataset. It can be seen that the fire pixels have been accurately segmented. The model has been able to pick up disconnectedness in the fire which can be seen in the respective masks. There are some areas where isolated pixels appear which can be removed by adding a smoothening process after the masks are generated. The interesting thing to note is that high saturation areas that were responsible for false positives in the rule-based approach do not get segmented as fire pixels here. This is why adding a CNN to the process leads to robust performance.

One of the hypothesis to be tested is whether the pre-selection of pixels improves the overall segmentation performance and specifically the false positive rate of the system. To test this, we train two models where the first one is trained without any prior segmentation while the other one uses the rule-based approach. Both of these models are then evaluated on their segmentation metrics such as F1, accuracy, precision, recall, loss, FP and FN and MCC. This experiment is restricted to the Corsican dataset due other datasets adding redundant information. As it can be seen, the F1 score of the model without any pre-selection is slightly higher than the model with pre-selection. The F1 score of the model without pre-selection is 0.85 while the other model has an F1 score of 0.82. However, the accuracy of the model is slightly lower than compared to the model with pre-selection. The more relevant portion of this analysis is the comparison of the false positives and negatives. As it can be seen, the model without pre-selection yields much higher false positives i.e., 0.023 compared to 0.0018. The model also shows slightly lesser false negative i.e. 0.12 compared to 0.18. Hence, pre-selecting pixels using a rule-based approach leads to reducing of the false-positive rate of the model. The false-negative rate can be further reduced by improving the rule-based approach itself.

IV. Conclusions & Future Work

In this paper, our innovation was to use image fusion combining state-of-the-art approaches [11], [13], [23] to improve wildfire detection. Our model was adapted to contain spectral normalisation along, multi-channel output along with modifications in the network structure. The training and evaluation of the models led to the analysis that the VGG method has the greatest balance when it comes to fusion but our model outperformed the others on the total amount of



Metric	Value	Loss	Image Type
MCC	0.92	Focal-Tversky	Visible
F1	0.93	Dice	Visible
HAF	0.90	Dice	Visible

Fig. 6. Results from deep learning segmentation with GAN-enabled fusion: (a) demonstration, (b) statistical testing.

information that was transferred and the similarity coefficient (SSIM). The fused images were added to the dataset and then their segmentation performance was evaluated.The quantitative results from testing performance of the segmentation model on visible and fused images suggest that the image categories do not impact segmentation performance. However, we were able to identify cases such as images which contains reflections from fire, high saturated backgrounds or similar colours tend to benefit from image fusion and it can be seen that F1 score for such cases is higher when using the fused images for segmentation. Hence, image fusion does not have global impact on segmentation of fire pixels but certain scenarios which include reflections and similarly coloured pixels benefit from fusing images from the visible and IR spectrum.

Our future work will focus on integrating capabilities for humanitarian science, such as with our prior work on detecting natural disaster damage on infrastructures [8], [10].

References

[1] Akhloufi, M. A., Tokime, R. B., and Elassady, H. (2018). Wildland fires detection and segmentation using deep learning. In Pattern recognition and tracking xxix, volume 10649, pages 86-97. SPIE.

[2] Bulo, S. R., Porzi, L., and Kontschieder, P. (2018). In-place activated batch norm for memory-optimized training of DNNs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[3] Celik, T., Demirel, H., and Ozkaramanli, H. (2006). Automatic fire detection in video sequences. IEEE European Signal Processing Conf.

[4] Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv:1412.7062.

[5] Chen, T.-H.,Wu, P.-H., and Chiou, Y.-C. (2004). An early fire-detection method based on image processing. IEEE International Conference on Image Processing, pages 1707-1710.

[6] Ciprian-Sanchez, J. F., Ochoa-Ruiz, G., Gonzalez-Mendoza, M., and Rossi, L. (2021). Fire-gan: A novel deep learning-based infrared-visible fusion method for wildfire imagery. Neural Computing and Applications.

[7] Bai, Y., Chen, W., Chen, J., Guo, W. (2020). Deep learning methods for solving linear inverse problems: Research directions and paradigms. Signal Processing, 177.

[8] Pamuncak, A., Guo, W., Soliman Khaled A., Laory, I., (2019). Deep learning for bridge load capacity estimation in post-disaster and -conflict zones. Royal Society Open Science.

[9] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770-778.

[10] Lu, L., Guo, W. (2021). Automatic Quantification of Settlement Damage using Deep Learning of Satellite Images. IEEE International Smart Cities Conference (ISC2).

[11] Li, H., Wu, X.-J., and Kittler, J. (2018). Infrared and visible image fusion using a deep learning framework. In 2018 24th international conference on pattern recognition (ICPR), pages 2705-2710. IEEE.

[12] Luo, X., Zhang, Z., Zhang, B., and Wu, X. (2016). Image fusion with contextual statistical similarity and non-subsampled shearlet transform. IEEE sensors journal, 17(6):1760-1771.

[13] Ma, J., Yu, W., Liang, P., Li, C., and Jiang, J. (2019). Fusiongan: A generative adversarial network for infrared and visible image fusion. Information fusion, 48:11-26.

[14] Nemalidinne, S. M. and Gupta, D. (2018). Non-subsampled contourlet domain visible and infrared image fusion framework for fire detection using pulse coupled neural network and spatial fuzzy clustering. Fire Safety Journal, 101:84-101.

[15] Pohlen, T., Hermans, A., Mathias, M., and Leibe, B. (2017). Full-resolution residual networks for semantic segmentation in street scenes. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4151-4160.

[16] Qian, Z., Xiao-jun, L., and Lei, H. (2020). Video image fire recognition based on color space and moving object detection. In 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE), pages 367?371. IEEE.

[17] Quan, T. M., Hildebrand, D. G. C., and Jeong, W.-K. (2021). Fusionnet: A deep fully residual convolutional neural network for image segmentation in connectomics. Frontiers in Computer Science, page 34.

[18] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large scale image recognition. arXiv preprint arXiv:1409.1556.

[19] Toulouse, T., Rossi, L., Akhloufi, M., Celik, T., and Maldague, X. (2015). Benchmarking of wildland fire colour segmentation algorithms. IET Image Processing, 9(12):1064-1072.

[20] Van DerWerf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., Collatz, G. J., et al. (2017). Global fire emissions estimates during 1997-2016. Earth System Science Data, 9(2):697-720.

[21] Wang, Y., Dang, L., and Ren, J. (2019). Forest fire image recognition based on convolutional neural network. Journal of Algorithms & Computational Technology, 13:1748302619887689.

[22] Yuan, C., Zhang, Y., and Liu, Z. (2015). A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. Canadian journal of forest research, 45(7):783?792.

[23] Zhao, Y., Fu, G., Wang, H., and Zhang, S. (2020). The fusion of unmatched infrared and visible images based on generative adversarial networks. Mathematical Problems in Engineering, 2020.

2023-03-23

Multi-spectral fusion using generative adversarial networks for UAV detection of wild fires

Kacker, Tanmay

IEEE