International Journal of Early Childhood Special Education (INT-JECSE)
DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

A MACHINE LEARNING APPROACH TO ANISOTROPIC FILTERING FOR INFRARED AND VISIBLE SENSOR IMAGES

[1]BABU GUNDLAPALLY, [2]BOBBALA SREEDEVI,[3]ANIL KUMAR CHIDRA,[4]BATTAGANI MEGHANA,[5]BITLA PRAMUKHA
[123]Associate Professor,[45]Students
Department of ECE
Vaagdevi College of Engineering, Warangal, Telangana

ABSTRACT
The combination of visible and infrared imaging is a significant and common issue. In order to merge the characteristics seen in visible and infrared pictures into a single image, many fusion techniques have recently been developed. These cutting-edge techniques are extensively employed in several applications, such as target detection, picture classification, and image pre-processing. Finding important elements in the original photos and combining them to create the fused image is the main challenge in image fusion. Discrete wavelet transform (DWT), contourlet transform, shift-invariant shearlet transform, quaternion wavelet transform, and other traditional signal processing techniques have been used for decades in the image fusion area to extract picture characteristics. for the purpose of fusing visible and infrared images. However, the fused picture could contain artefacts from various techniques. Optimization-based fusion techniques are suggested as a solution to these issues. To get the best answer (fused picture), these approaches require several iterations. Due to the numerous rounds, these optimisation techniques could oversmooth the fused image. Furthermore, edge-preserving picture fusion techniques are also gaining popularity. In order to identify the borders of visible and infrared sensor images, this work used anisotropic filtering. For the objective of fusion, this approach uses edge-preserving smoothing filtering/process. Additionally, machine learning has emerged as a very active research technique with applications in several image processing domains. In order to create a single image that incorporates all of the information from both visible and infrared images, we suggest an efficient image fusion technique based on an unsupervised machine learning framework based on principal component analysis (PCA). Anisotropic filtering is used to first break down the source pictures into their fundamental components and detail content. We employ PCA to extract multi-layer features for the detail content. We create the final fused detail content using the weighted-average technique and these attributes.

I. INTRODUCTION

1.1 OVERVIEW:

A charge coupled device (CCD) may be used to collect visual information [1] in a scene as a picture. The visible (VI) light wavelengths that a CCD sensor can detect fall between 400 and 700 nm. However, a CCD or VI picture by itself is unable to offer all the scene data in the majority of computer vision applications. Multiple sensors of various modalities should be used to collect complementary pictures of the same scene in order to extract additional details. By taking pictures at wavelengths other than the VI band of the electromagnetic spectrum, this can be accomplished. Infrared (IR) radiation is often emitted by things that are warmer than zero Kelvin. These items produce energy, which infrared sensors can detect and show to users as pictures.Nevertheless, these pictures by themselves are insufficient to accurately describe the scenario in question. Therefore, to properly understand a situation, important information from both visual and infrared photos must be integrated into a single image. Fusion is the technique or procedure of blending different pictures into one image to improve the scene's perception [2]. Digital photography [3], medical imaging [4], remote sensing, the military, the detection of concealed weapons, night vision, target recognition and tracking, and other fields are the primary uses of image fusion. Single sensor image fusion (SSIF) and multisensor image fusion (MSIF) are the two main forms of fusion. In SSIF, several photos of the same scene are captured using a single sensor.

On the other hand, MSIF uses multimodal sensors to record supplementary data. While the remaining applications, including medical imaging, remote sensing, military, concealed weapon detection, target identification, and tracking, are under MSIF, digital photography is an example of SSIF. In this work,

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

we concentrate on MSIF specifically for military, night vision, and concealed weapon detection applications. Generally speaking, VI and IR sensors are employed in each of these applications to gather complementary data about the intended scenario. VI sensors are employed in the military to gather visual data, like combat field conditions. However, depending on temperature changes, infrared sensors are utilised to gather information about soldiers, vehicles, animals, and camouflage. Millimetre wave (MMW) sensors are utilised to detect hidden or concealed weapons, whereas VI sensors are used to gather visual information.

• In this study, a novel NSCT-based image fusion framework for multimodal medical pictures is proposed.

• For the combination of low and high-frequency coefficients, two distinct fusion rules are suggested.

• The phase congruency-based model is utilised to fuse the low-frequency coefficients. Phase congruency's primary advantage is that it chooses and blends the low frequency coefficients' contrast- and brightness-invariant representation.

• In contrast, high-frequency coefficients are combined using a novel definition of directive contrast in the NSCT domain. The most noticeable texture and edge information is chosen from high-frequency coefficients using directional contrast, and then incorporated in the fused ones.

• By adding a visual constant to the SML-based definition of directive contrast, which offers a more comprehensive depiction of the contrast, the concept of directive contrast is strengthened. Additionally, the suggested approach is expanded for multispectral fusion in colour space, which basically corrects undesired cross-channel artefacts in the IHS colour space and generates the highest quality output with natural spectral characteristics and enhanced colour information.

1.2 Potential Applications of Image Fusion

Intelligent robots

Require motion control, based on feedback from the environment from visual, tactile, force/torque, and other types of sensors

Stereo camera fusion

Intelligent viewing control

Automatic target recognition and tracking

Medical image

Fusing X-ray computed tomography (CT) and magnetic resonance (MR) images

Computer assisted surgery

Spatial registration of 3-D surface

Manufacturing

Electronic circuit and component inspection

Product surface measurement and inspection

non-destructive material inspection

Manufacture process monitoring

Complex machine/device diagnostics

Intelligent robots on assembly lines

Military and law enforcement

Detection, tracking, identification of ocean (air,ground)target/event

Concealed weapon detection

Battle-field monitoring

Night pilot guidance

Remote sensing

Using various parts of the electro-magnetic spectrum

Sensors: from black-and-white aerial photography to multi-spectral active microwave space-borne imaging radar

Fusion techniques are classified into photographic method and numerical method.

Image as Matrices:

The preceding discussion leads to the following representation for a digitized image function:

f (0,0)   f(0,1)   ………..   f(0,N-1)
f(1,0)   f(1,1)   …………   f(1,N-1)
f(xylem)=   .   .   .

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

. . .
f(M-1,0) f(M-1,1) ………… f(M-1,N-1)

The right side of this equation is a digital image by definition. Each element of this array is called an image element, picture element, pixel or pel. The terms image and pixel are used throughout the rest of our discussions to denote a digital image and its elements.

A digital image can be represented naturally as a MATLAB matrix:

f(1,1) f(1,2) ……. f(1,N)
f(2,1) f(2,2) …….. f(2,N) . . .
f = f(M,1) f(M,2) …….f(M,N)

Where f(1,1) = f(0,0) (note the use of a monoscope font to denote MATLAB quantities). Clearly the two representations are identical, except for the shift in origin. The notation f(p ,q) denotes the element located in row p and the column q. For example f(6,2) is the element in the sixth row and second column of the matrix f. Typically we use the letters M and N respectively to denote the number of rows and columns in a matrix. A 1xN matrix is called a row vector whereas an Mx1 matrix is called a column vector. A 1x1 matrix is a scalar. Matrices in MATLAB are stored in variables with names such as A, a, RGB, real array and so on. Variables must begin with a letter and contain only letters, numerals and underscores. As noted in the previous paragraph, all MATLAB quantities are written using mono-scope characters. We use conventional Roman, italic notation such as f(x ,y), for mathematical expressions

Reading Images:

Images are read into the MATLAB environment using function imread whose syntax is imread('filename')

Format name	Description	recognized extension
TIFF	Tagged Image File Format	.tif, .ti
JPEG	Joint Photograph Experts Group	.jpg, .jpeg
GIF	Graphics Interchange Format	.gif
BMP	Windows Bitmap	.bmp
PNG	Portable Network Graphics	.png
XWD	X Window Dump	.xwd

Here filename is a spring containing the complete of the image file(including any applicable extension).For example the command line >> f = imread ('8. jpg'); reads the JPEG (above table) image chestxray into image array f. Note the use of single quotes (') to delimit the string filename. The semicolon at the end of a command line is used by MATLAB for suppressing output. If a semicolon is not included.MATLAB displays the results of the operation(s) specified in that line. The prompt symbol(>>) designates the beginning of a command line, as it appears in the MATLAB command window. When as in the preceding command line no path is included in filename, imread reads the file from the current directory and if that fails it tries to find the file in the MATLAB search path. The simplest way to read an image from a specified directory is to include a full or relative path to that directory in filename.

II. LITERATURE REVIEW

In [5] authors presented a convolution-guided transformer framework for infrared and visible image fusion (CGTF), which aims to combine the local features of convolutional network and the long-range dependence features of transformer to produce satisfactory fused image. In CGTF, the local features are calculated by convolution feature extraction module (CFEM), and then, the local features are used to guide the transformer feature extraction module (TFEM) to capture the long-range dependences of the image, which can overcome not only the lack of long-range dependences that exist in convolutional fusion methods but also the deficiency of local feature that exists in transformer models.

In [6] authors presented a new technique is suggested for the fusion of infrared (IR) and visible (VIS) images. The primary problems for image fusion at feature levels are that artefacts and noise are introduced in the fused picture. The weight map generated by the modified naked mole-rat algorithm (mNMRA) is used to retain important information without using artefacts in a final fused image. The proposed FNMRA fusion method is based on a feature-level fusion after the refinement of weight maps, utilising the WLS approach.

In [7] authors presented in order to extract more detailed information through multi-channel inputs. In contrast to conventional unsupervised fusion network, the proposed network contains three channels for extracting infrared features, visible features and common features of infrared and visible images, respectively.

In [8] authors presented an infrared and visible image fusion method based on relative total variation decomposition, which can maintain the contrast information and texture information of source images simultaneously. Firstly, the source images are decomposed into structural layers and texture layers according to the relative total variation.

In [9] authors presented a novel fusion model using a triple-discriminator generative adversarial network, which can achieve the balance. The difference image obtained by image subtraction can highlight the difference information, extract image details, and obtain the target outlines in some scenes. Therefore, besides the visible discriminator and infrared discriminator, a new difference image discriminator is added to retain the difference between infrared and visible images, thereby improving the contrast of infrared targets and keeping the texture details in visible images.

In [10] authors presented a progressive image fusion network based on illumination-aware, termed as PIAFusion, which adaptively maintains the intensity distribution of salient targets and preserves texture information in the background. Specifically, we design an illumination-aware sub-network to estimate the illumination distribution and calculate the illumination probability.

In [11] authors presented ATRE technique is used to efficiently extract the bright object regions from the IR image and retain much of the visual background regions from the VI. An adaptive parameter is introduced for accurate segmentation. A region mapping process is followed to get the fused image.

In [12] authors presented a deep-learning-based infrared-visible images fusion method based on encoder-decoder architecture. The image fusion task is reformulated as a problem of maintaining the structure and intensity ratio of the infrared-visible image.

In [13] authors presented method synthesizes visible and near-infrared images using contourlet transform, principal component analysis, and iCAM06, while the blending method uses color information in a visible image and detailed information in an infrared image.

In [14] authors presented to handle this problem. However, for multi-modality image fusion, using the same network cannot extract effective feature maps from source images that are obtained by different image sensors. In TPFusion, we can avoid this issue. At first, we extract the textural information of the source images. Then two densely connected networks are trained to fuse textural information and source image, respectively. By this way, we can preserve more textural details in the fused image.

In [15] authors presented a novel infrared and low-light visible image fusion method from the perspective of low-light visible image enhancement, weak feature extraction strategy and detail preserved fusion rules. By combining both local and global contrast enhancements, an adaptive light adjustment algorithm is proposed to improve the brightness and texture details of low-light visible images. In addition, we design a hybrid multiscale decomposition model based on guided filters (GFs) and side window guided filters (SWGFs) to decompose the source images into the base layer, large-scale detail layers and small-scale detail layers.

In [16] authors presented a heterogeneous knowledge distillation network (HKDnet) with multilayer attention embedding to jointly implement the fusion and super-resolution of infrared and visible images. Precisely, the proposed method consists of a high-resolution image fusion network (teacher network) and a low-resolution image fusion and super-resolution network (student network). The teacher network mainly fuses the high-resolution input images and guides the student network to obtain the ability of joint implementation of fusion and super-resolution.

In [17] authors presented a novel Image Fusion Transformer (IFT) where we develop a transformer-based multi-scale fusion strategy that attends to both local and long-range information (or global context). The proposed method follows a two-stage training approach. In the first stage, we train an auto-encoder to extract deep features at multiple scales. In the second stage, multi-scale features are fused using a Spatio-Transformer (ST) fusion strategy.

In [18] authors presented a unified gradient and intensity discriminator generative adversarial network for various image fusion tasks, including infrared and visible image fusion, medical image fusion, multi-focus image fusion, and multi-exposure image fusion. On the one hand, we unify all fusion

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

tasks into discriminating a fused image's gradient and intensity distributions based on a generative adversarial network.

In [19] authors presented in recent years have problems such as loss of detailed information and low contrast. In this paper, we propose a novel infrared and visible image fusion method based on visibility enhancement and hybrid multiscale decomposition.

In [20] authors presented a novel infrared and visible image fusion method, i.e., the Y-shape dynamic Trans- former (YDTR). Specifically, a dynamic Transformer module (DTRM) is designed to acquire not only the local features but also the significant context information. Furthermore, the proposed network is devised in a Y-shape to comprehensively maintain the thermal radiation information from the infrared image and scene details from the visible image. Considering the specific information provided by the source images, we design a loss function that consists of two terms to improve fusion quality: a structural similarity (SSIM) term and a spatial frequency (SF) term.

III. EXISTING METHOD

WAVELET TRANSFORM

A signal analysis method similar to image pyramids is the discrete wavelet transform. The main difference is that while image pyramids lead to an over complete set of transform coefficients, the wavelet transform results in a nonredundant image representation. The discrete 2-dim wavelet transform is computed by the recursive application of lowpass and high pass filters in each direction of the input image (i.e. rows and columns) followed by sub sampling. Details on this scheme can be found in the reference section. One major drawback of the wavelet transform when applied to image fusion is itswell known shift dependency, i.e. a simple shift of the input signal may lead to complete different transform coefficients. This results in inconsistent fused images when invoked in image sequence fusion. To overcome the shift dependency of the wavelet fusion scheme, the input images must be decomposed into a shift invariant representation. There are several ways to achieve this: The straightforward way is to compute the wavelet transform for all possible circular shifts of the input signal. In this case, not all shifts are necessary and it is possible to develop an efficient computation scheme for the resulting wavelet representation. Another simple approach is to drop the subsampling in the decomposition process and instead modify the filters at each decomposition level, resulting in a highly redundant signal representation.

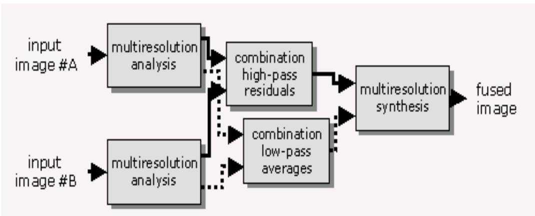

Fig 1: Block diagram of basic image fusion process.

Continuous Wavelet Transform (CWT)

A continuous wavelet transform (CWT) is used to divide a continuous-time function into wavelets. Unlike Fourier transform, the continuous wavelet transform possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. In mathematics, the continuous wavelet transform of a continuous, square-integrable function $x(t)$ at a scale $a > 0$ and translational value $b \in \mathbb{R}$ is expressed by the following integral

$$X_w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt$$

Where $\psi(t)$ is a continuous function in both the time domain and the frequency domain called the mother wavelet and $*$ represents operation of complex conjugate. The main purpose of the mother wavelet is to provide a source function to generate the daughter wavelets which are simply the translated and scaled versions of the mother wavelet. To recover the original signal $x(t)$, inverse continuous wavelet transform can be exploited.

$$x(t) = \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2} X_w(a,b) \frac{1}{\sqrt{|(a)|}} \tilde{\psi}\left(\frac{t-b}{a}\right) db\, da$$

$\tilde{\psi}(t)$ is the dual function of $\psi(t)$. And the dual function should satisfy

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

$$\int_0^{\infty}\int_{-\infty}^{\infty}\frac{1}{|a^3|}\psi\left(\frac{t_1-b}{a}\right)\tilde{\psi}\left(\frac{t-b}{a}\right)\,db\,da=\delta(t-t_1)$$

Sometimes, $\tilde{\psi}(t)=C_{\psi}^{-1}\psi(t)$, where

$$C_{\psi}=\frac{1}{2}\int_{-\infty}^{+\infty}\frac{\left|\hat{\psi}(\zeta)\right|^2}{|\zeta|}d\zeta$$

Is called the admissibility constant and $\hat{\psi}$ is the Fourier transform of ψ. For a successful inverse transform, the admissibility constant has to satisfy the admissibility condition:

$$0<C_{\psi}<+\infty.$$

It is possible to show that the admissibility condition implies that $\hat{\psi}(0)=0$, so that a wavelet must integrate to zero.

METHOD USED

Describes the brief explanation of our proposed fusion frame work. Fused output image is obtained by implementation of NLAF process to obtain the approximate and detail layers with PCA fusion rule. Proposed NLAF-PCA fusion methodology shown in figure.

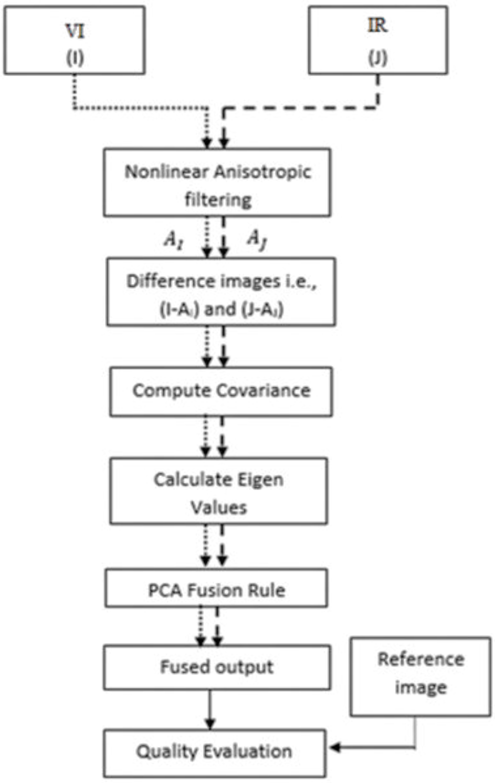

Fig.2: Proposed NLAF-PCA fusion process flow

IV. RESULTS & DISCUSSIONS

All the experiments have been done in MATLAB 2016b version under the high-speed CPU conditions for faster running time. Aim of any fusion algorithm is to integrate required information from both source images in the output image. Fused image cannot be judged exclusively by seeing the output image or by measuring fusion metrics. It should be judged qualitatively using visual display and quantitatively using fusion metrics.

(INPUT A) (INPUT B) (FUSED IMAGE C)

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

Figure 3: Fusion performance on Dataset-1

(INPUT A) (INPUT B) (FUSED IMAGE C)

Figure 4: Fusion performance on Dataset-2

The above fig i.e input-A is an Infrared image and the image taken as input-B is a visible image. With the help of Non-linear anisotropic filtering framework and Principal component analysis both the inputs will be combined to get a fused output.

NLAF has utilized to extract the approximate and detail layers from the IR and VI source images.

In both the inputs the picture's edges will be not clear but when they are fused together the output obtained will be the clear form of the image.

Using Principal component analysis the dimensionality of the images will be reduced and, increasing the interpretability at the same time minimizing the information loss.

The fusing process will be done as follows

Select and read VI and IR source images from the MATLAB current directory.

Convert the source images into gray scale in case of RGB image.

Apply NLAF process to obtain approximate layers of VI and IR images.

Subtract the source images from the obtained approximate layers to get the detailed layers of VI and IR images

Compute the covariance of detailed layers obtained from the above step.

Calculate the eigen vectors for the output of the above step.

Now apply PCA fusion rule to obtain final fused output of VI and IR images.

The PCA method generates a new set of variables, called principal components. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other, so there is no redundant information. The principal components form an orthogonal basis for the space of the data

Table 1: quantitative analysis of fusion methods for dataset

Methodology	PSNR (in dB)	RMSE	CC	SSIM	Entropy
SWT [24]	68.95	0.0909	0.933	0.988	0.9684
DWT [25]	68.98	0.0906	0.934	0.988	0.9683
Proposed method	74.18	0.049	0.973	0.999	5.16

Quantitative analysis with IQA sown in table 1 for the test results presented in figure 6.1, which gives the analysis of dataset . Table 6.1 consists of various fusion metric parameters such as PSNR, RMSE, CC, SSIM and entropy. The best values are highlighted in bold letters. Our proposed method obtained far better values over all the existing fusion methods discussed in the literature. We also tested the qualitative analysis of dataset 2 with the similar fusion metric parameters considered for dataset 1.

V. CONCLUSION

Using the NLAF-PCA methodology, a novel texture-preserving fusion method is suggested for IR and VI pictures. The IR and VI source pictures' approximation and detail layers have been extracted using NLAF. The principle components were then calculated using the PCA technique. Lastly, fusion is utilised to create a combined picture while maintaining texture. A number of medical image fusion techniques that have been published in the literature are used to evaluate the performance of the suggested NLAF-PCA fusion procedure. The suggested NLAF-PCA fusion technique outperformed

DOI:10.48047/intjecse/v13i2.21282 ISSN: 1308-5581 Vol 13, Issue 02 2021

the traditional medical fusion algorithms, according to a comparative analysis based on picture quality parameters.

REFERENCES

[1] *Le, Zhuliang, et al. "UIFGAN: An unsupervised continual-learning generative adversarial network for unified image fusion." Information Fusion 88 (2022): 305-318.*

[2] *Lin, Yingcheng, and Dingxin Cao. "Adaptive infrared and visible image fusion method by using rolling guidance filter and saliency detection." Optik 262 (2022): 169218.*

[3] *Lu, Ruitao, et al. "A novel infrared and visible image fusion method based on multi-level saliency integration." The Visual Computer (2022): 1-15.*

[4] *Wang, Di, et al. "Unsupervised misaligned infrared and visible image fusion via cross-modality image generation and registration." arXiv preprint arXiv:2205.11876 (2022).*

[5] *Li, Jing, et al. "Cgtf: Convolution-guided transformer for infrared and visible image fusion." IEEE Transactions on Instrumentation and Measurement 71 (2022): 1-14.*

[6] *Singh, Simrandeep, Nitin Mittal, and Harbinder Singh. "A feature level image fusion for IR and visible image using mNMRA based segmentation." Neural Computing and Applications 34.10 (2022): 8137-8154.*

[7] *Wang, Hongmei, et al. "Infrared and visible image fusion based on multi-channel convolutional neural network." IET Image Processing 16.6 (2022): 1575-1584.*

[8] *Chen, Jun, Xuejiao Li, and Kangle Wu. "Infrared and visible image fusion based on relative total variation decomposition." Infrared Physics & Technology 123 (2022): 104112.*

[9] *Song, Anyang, et al. "Triple-discriminator generative adversarial network for infrared and visible image fusion." Neurocomputing 483 (2022): 183-194.*

[10] *Tang, Linfeng, et al. "PIAFusion: A progressive infrared and visible image fusion network based on illumination aware." Information Fusion 83 (2022): 79-92.*

[11] *Meher, Bikash, et al. "Visible and infrared image fusion using an efficient adaptive transition region extraction technique." Engineering Science and Technology, an International Journal 29 (2022): 101037.*

[12] *Wang, Bowen, et al. "Multimodal super-resolution reconstruction of infrared and visible images via deep learning." Optics and Lasers in Engineering 156 (2022): 107078.*

[13] *Son, Dong-Min, Hyuk-Ju Kwon, and Sung-Hak Lee. "Visible and Near Infrared Image Fusion Using Base Tone Compression and Detail Transform Fusion." Chemosensors 10.4 (2022): 124.*

[14] *Yang, Zhiguang, and Shan Zeng. "TPFusion: Texture preserving fusion of infrared and visible images via dense networks." Entropy 24.2 (2022): 294.*

[15] *Zou, Dengpeng, and Bin Yang. "Infrared and low-light visible image fusion based on hybrid multiscale decomposition and adaptive light adjustment." Optics and Lasers in Engineering 160 (2023): 107268.*

[16] *Xiao, Wanxin, et al. "Heterogeneous knowledge distillation for simultaneous infrared-visible image fusion and super-resolution." IEEE Transactions on Instrumentation and Measurement 71 (2022): 1-15.*

[17] *Vs, Vibashan, et al. "Image fusion transformer." 2022 IEEE International Conference on Image Processing (ICIP). IEEE, 2022.*

[18] *Zhou, Huabing, et al. "Unified gradient-and intensity-discriminator generative adversarial network for image fusion." Information Fusion 88 (2022): 184-201.*

[19] *Luo, Yueying, et al. "Infrared and visible image fusion based on visibility enhancement and hybrid multiscale decomposition." Optik 258 (2022): 168914.*

[20] *Tang, Wei, Fazhi He, and Yu Liu. "Ydtr: infrared and visible image fusion via y-shape dynamic transformer." IEEE Transactions on Multimedia (2022).*