***Article Received:** 25 July 2024 **Revised:** 12 August 2024 **Accepted:** 30 September 2024*

Optimizing Visual Content Representation Through Semantic Sparse Recoding

Dr. Harsh Lohiya [1]

Research Guide Department of Computer Science & Engineering , Sri Satya Sai University of Technology and Medical Sciences, Bhopal, M.P, India,

Surender Reddy S[2]

Research Scholar, Department of Computer Science & Engineering, Sri Satya Sai University of Technology and Medical Sciences, Bhopal, M.P, India,

ABSTRACT

This study introduces a novel methodology for optimizing visual content representation through Semantic Sparse Recoding (SSR). By leveraging advanced sparse representation techniques and integrating a Global Dictionary Learning approach, the proposed system addresses limitations in conventional image fusion and content retrieval methods. The SSR framework improves the ability to preserve structural details and semantic features, particularly for multi-modal image datasets. Experimental results demonstrate the system's superior performance in terms of edge preservation, visual fidelity, and computational efficiency compared to state-of-the-art techniques. Applications span various domains including medical imaging, surveillance, and multimedia content management.

Keywords: Semantic Sparse Recoding, Visual Content Optimization, Image Fusion, Sparse Representation, Global Dictionary Learning

INTRODUCTION

Humans nowadays have the ability to take and share images using a variety of digital devices, such as smartphones, tablets, and computers, as well as the internet, using a variety of applications, such as image-sharing websites like flickr.com or Facebook. These days, there are a lot of digital images and images being shared on the internet, and many of them are directly related to people, such as facial images, which are intimately linked to the social engagement of human beings in cyberspace. The term cyberspace refers to the Internet environment.

IMAGE FUSION

Images are the only way to describe objects. Some camera images are clear, while others are blurry. Bai et al. (2015) attribute this to the camera's narrow DOF. The narrow depth of field (DOF) prevents appropriate focus on all visual elements. Images with all the relevant features were captured using image fusion technology. This method combines many photos into one (Zhang et al. 2013). Remote sensing, surveillance, medical imaging, warfare monitoring, and hidden weapon identification can benefit from picture fusion.

Visible cameras in surveillance applications can provide exact facts about their surroundings, but if a target is indistinguishable, they won't be disclosed (Aslantas et al. 2011). Although thermal cameras do not collect daytime data, they can reveal hidden targets. Combining thermal and visible pictures with background information to identify hidden targets is crucial for autonomous tracking, face recognition, target detection, action recognition, and object identification.

Due to fluctuating radiation power, each medical technique can only display limited organ features. CT is great for imaging bones, but not soft tissue. However, magnetic resonance imaging (MRI) can distinguish normal and diseased soft tissues better but lacks boundary information. PET and SPECT show blood flow but not organ localization. Doctors desire a single image with information from many imaging modalities to better diagnosis. Remote sensing uses multi-spectral (MS) and panchromatic (PAN) pictures to achieve high spectral and spatial resolution.

Several image fusion algorithms have been published. This category includes spatial and transform algorithms (Goshtasby & Nikolov 2007). Spatial domain fusion (SDF) approaches integrate pixels from many input images to form

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 July 2024 **Revised:** 12 August 2024 **Accepted:** 30 September 2024*

the synthesised image. Incorrect pixel or region selection may cause blocking artifacts and a lack of contrast and sharpness in the fused image. Transform Domain Fusion (TDF) approaches were developed to mitigate these effects. These transforms include the laplacian pyramid (LP), DWT, DTCWT, DCH, DCH DWT, and discrete wavelet transform. NSCT and DCHWT are digital contourlet functions, according to Li et al. (2011). The best transform base relies on the image context and use. In addition to transform choice, fusion rule subband coefficients have negatively affected fusion performance.

Sparse representation (SR) theory has improved picture fusion by overcoming the drawbacks of TDF approaches (Chen et al. 2013). Sparse-based fusion is complex since it requires establishing an over-complete dictionary, accurately determining sparse coefficient activity, formulating the optimum fusion rule, and rebuilding the merged image.

Image retrieval from Local Databases

Using the local database, a user can get a clear image of what is on their hard drive. No automatic features are provided for any institutions; we simply allow the user to see the image. They allowed the user to have complete control over their hard drive and displayed all of the files they had been working on. Organizing these images with a manual organiser gives the customer the most flexibility. There are currently only two primary techniques for manually arranging images: using tags or a directory. However, each treatment has its own unique procedures, goals, and consequences.

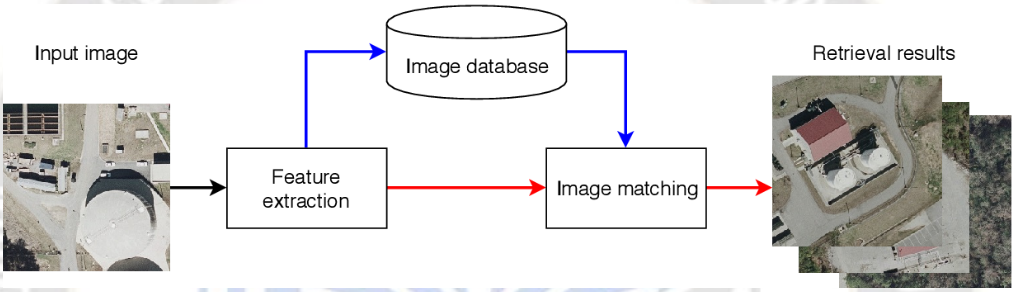


Figure 1: Image Retrieval from the Local Database

Cyberspace Databases:

The term cyberspace refers to the digital world created by a global network of computers for the purpose of facilitating online communication. It is also known as a global network of computers that uses the TCP/IP protocol to exchange data and communicate with other computers around the world. It is possible to interact with a wide range of people in cyberspace, which is an interactive and virtual world. In cyberspace, users can share information, swap ideas, communicate with one another, play online games, and engage in various platforms for discussion, as well as for business objectives. In a book titled Neuromancer by William Gibson in 1984, the phrase was used for the first time. Gibson later called this evocative and essentially pointless in his criticism of it. In spite of this, the word cyberspace is still extensively used to refer to any facility or feature that is directly connected to the Internet.

Many IT experts and specialists, such as Chip Morningstar and F. Randall Farmer, believe that rather than its technological design and execution, cyberspace has caught the popularity of society as a medium for contact. Web images are images that may be accessed by anybody with an internet connection. Non-homogeneous, semi-structured, and large in volume, these images are typically stored in arrays on storage. Images from cyberspace can be accessed via a wide range of web servers. These include pixabay.com, shutterstock.com, everypixal.com, bing.com, fliker.com, and so on. Microsoft owns and operates the bing.com web search engine. The service roots can be traced back to Microsoft prior search engines. Many services are provided by Bing, such as searching for products for videos and images as well as maps of the Internet. There are currently 40 languages available for this service, which was launched on June 3rd, 2009. ASP.NET was used to build the search engine.

Understanding the content is a major issue in computer vision, especially when it comes to image retrieval based on content. A black and white pixel may be recognised by a computer, but it does not perceive white and black pixels as being used for the same purpose. As a result of this, the so-called semantic gap has emerged. As a result, we need to utilise machine or statistical learning to teach the computer to recognise this pattern. The pixel-to-semantic scale Digital images can be decoded using machine learning algorithms that can identify low-level properties such as colour patterns, texture, and shape. However, it was clueless when it came to an image more complex aspects. The semantic gap is referred to as such.

In machine learning, this is a critical flaw. An abstraction of low-level pixel data is required by the machine. It is a tricky issue. When the semantic chasm is closed, we may be able to build machines with actual intelligence. Content-Based Image Retrieval systems have yet to meet the expectations of their customers because of the low-level image properties, usage, and the semantic gap problem.

LITERATURE REVIEW

To solve these important issues, Yang et al. (2022) presented a novel fine-grained visual categorization (FGVC) approach that takes into account both small variations between classes and large changes within them. Their approach enhances spatial feature associations with the help of an attention-based Symmetrized Local Feature Extraction Module (SLFEM) and a Local Feature Extraction Anchor Generator (LFEAG) that mimics the geometric forms of irregular local features. Their methodology outperforms conventional anchor approaches in capturing discriminative characteristics, as demonstrated by its performance on benchmark datasets. In their 2023 study, Preethi and Krupa categorized recent developments in lip reading biometrics, audiovisual speech recognition, silent speech recognition, voice from lips, and lip HCI, and they assessed lower-level facial gesture analysis methods for lip-reading applications. In doing so, they uncovered gaps in current methods and suggested future directions for creating more resilient real-world systems through the use of computer vision, machine learning, and deep learning to problems in a variety of domains.

Ngo et al. (2024) used a mapping review to integrate memory development research findings across frameworks. Over five decades, they examined experimental designs to find methodological convergences and topics for additional study. Pattern separation helped explain memory development cumulatively.

Zhang et al. (2022) introduced the single-image dehazing MFFE network. Their method improves texture clarity and detail recovery using super-resolution algorithms. Their network compensates for contextual information loss with a multiscale feature fusion module and attention mechanisms, exceeding state-of-the-art haze removal and image quality restoration methods.

Autonomous agricultural navigation algorithms based on vision were studied by Zhang et al. (2024). They looked at centerline identification, crop picture capture, and canopy feature extraction. In order to enhance autonomous localization and navigation, they recommended studying vision-based algorithms after discovering problems with accuracy and robustness.

MATERIALS AND METHOD

In the GDFF technique, a pre-classification procedure is applied to the training signals collected from external data sets. From the input multi-focus image pairs, the training samples for this suggested system are generated. Hence, compared to the conventional sparse dictionaries, performance is enhanced due to the improved representation of input image data. In addition, a global dictionary is obtained from prior sparse representations and the sparse representation is executed in the dominant gradient direction. Therefore, representation coefficients let us capture the input data's intrinsic structure more efficiently.

Global Dictionary Learning using Classification (GDLC):

One of the most important challenges in sparse representation modelling is constructing a representation that is both meaningful and stable in light of the input data. A sparse dictionary can help you accomplish this. It is only possible for a dictionary to be more flexible than the structure of the input image if the dictionary is constructed.

The analysis of the properties of training data is extensive. Using gradient operators that can take advantage of this structural knowledge is one example. The gradient information of each patch in the training data set is used to evaluate the focus features of each patch. Figure 1, depicts the GDLC approach complete dictionary learning process.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 July 2024 **Revised:** 12 August 2024 **Accepted:** 30 September 2024*

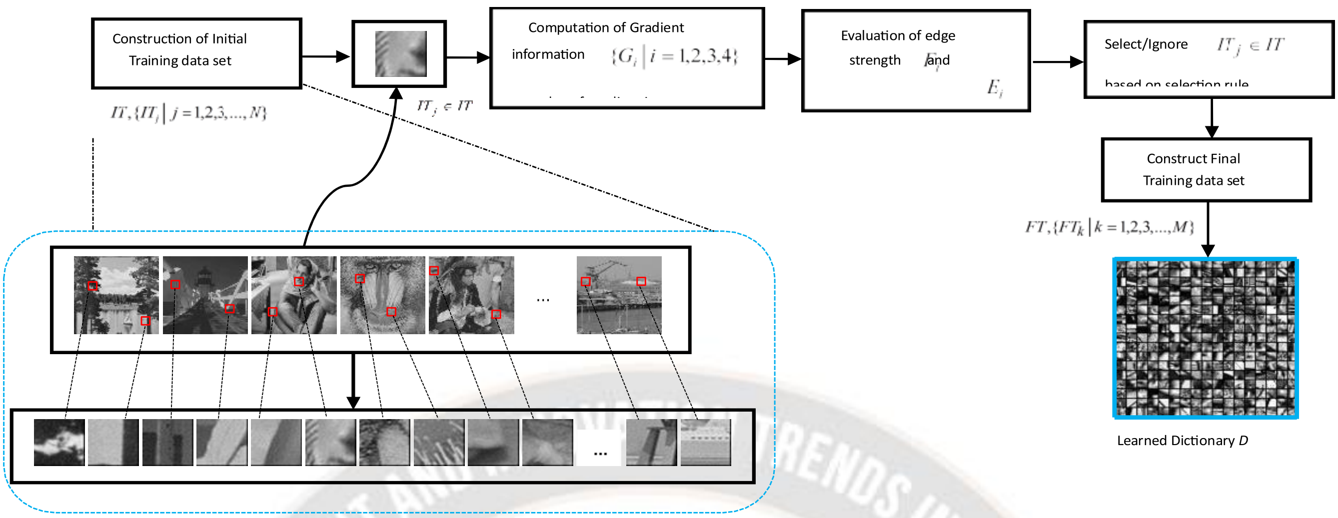


Figure 2: GDLC Framework

Creating a high-quality image with less ambiguity is the main goal of image fusion, which is widely recognized. A number of criteria, including clarity, information density, contrast sensitivity, blocking artifacts, and so on, determine what is considered optimal quality. To achieve the best possible result, this study uses sharpness and information content as its main emphasis features while selecting patches. The sharpness of patch *ITj* □ *IT* , is measured by evaluating the edge strength, preserved along the four gradient directions {*Gi* (*x*, *y*) *i* □ 1, 2,3, 4}. So, the edge strength {*Fi i* □ 1, 2,3, 4} of patch *ITj* □ *IT* , in each of the gradient direction is given by,

$$F1 \square \sqrt{\frac{1}{m} \sum^{m} \sum^{n} [G_1(x, y)]^2}$$

$$F2 \square \sqrt{\frac{1}{mn} \sum_{x=2}^{m} \sum_{y=1}^{n} [G_2(x, y)]^2}$$

$$F3 \square \sqrt{\frac{1}{\sqrt{2}} \cdot \frac{1}{mn} \sum_{x=2}^{m} \sum_{y=2}^{n} [G_3(x, y)]^2}$$

F_4=

$$4 = \sqrt{\frac{1}{\sqrt{2}} \cdot \frac{1}{m} \sum_{x=1}^{m} \sum_{y=2}^{n} [G_4(x, y)]}$$

A picture's structural content can be described by its gradient distribution along different directions. The histogram of oriented gradients (HOG) is commonly employed to make use of this data. Information entropy, denoted as {Eii = 1, 2, 3, 4} (Kvalseth 1987), is utilized as a focal point to assess the data content of a patch, where ITj ∊ IT along the gradient directions. The integer 255Gi The computation of the gradient information with a high Ei value is so i < arg max{Eii < 1, 2, 3, 4}.

Assume the final training data set of the dictionary learning process is denoted as *FT* ,{*FTk k* □ 1, 2,3,..., *M*}.- The patch's ability to retain edge details and information content in its dominant direction is determined using a selection rule, taking into account the factors mentioned above. According to the guideline, when learning a dictionary, use the patch with the most detailed and visually informative gradient direction. This procedure is iterated upon until every training data set has been classified. Before learning starts, we remove the average values of each patch from the final training data set to make sure that only the patches' edge structures are incorporated. A global dictionary, D, is created using the K-SVD technique. A popular and well-known method for learning

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 July 2024 **Revised:** 12 August 2024 **Accepted:** 30 September 2024*

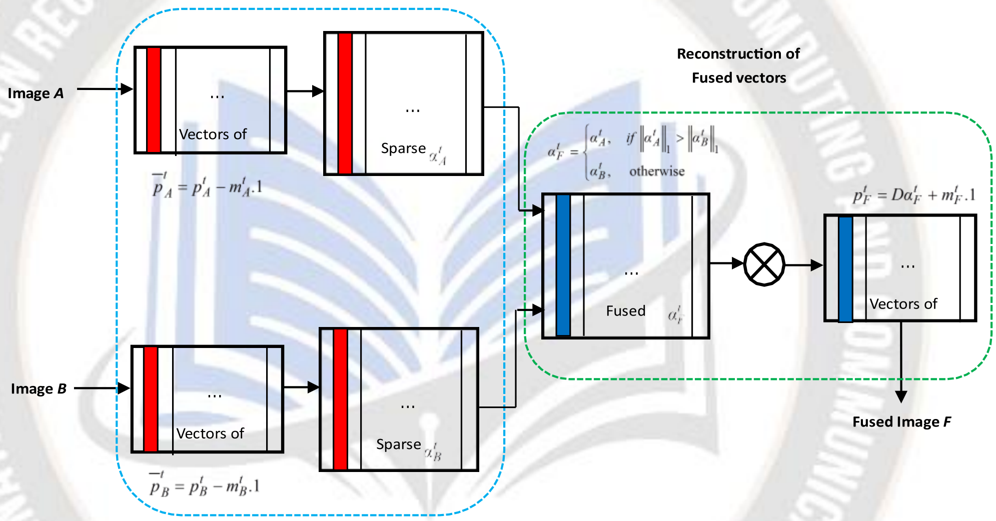
dictionaries is to employ sparse representation and K-means clustering, either with or without error restriction. According to Aharon et al. (2006), its convergence rate is quicker than MOD's.

Parameter Settings and Experimental Analysis

The training data set is built using 8x8 picture patches randomly selected from high-quality natural photographs obtained from http://decsai.ugr.es. The algorithm is subsequently trained using these patches. The global dictionary has 300 entries with an error bound of 0.01. There will be 40 iterations of the K-SVD method. Only 36.1% of the initial training data set—36,118 patches—made it into the final training sets. An analysis of the final datasets used for training led to this conclusion. You will be able to download and utilize the over-complete dictionary that was generated after 1463.099323 seconds. So, taking all of this into account considerably simplifies learning the lexicon.

Fusion Scheme

Figure 3, depicts a high-level image of the suggested fusion strategy.



Sparse Representation

Learned Dictionary *D*

Figure 3: GDLC Fusion

Due to its objectivity, subjective evaluation is a solid way to ascertain the ultimate picture quality. The scarcity of subject-matter specialists necessitates the development of objective fusion measures. The quality of the fusion algorithms, which consider both the original and fused images, is represented by this single numerical score. Several quality criteria have been used to confirm that these tactics are effective (Chen & Blum, 2009). The merging of pictures A and B yields the image F.

Edge-based quality metric (*QAB*/ *F*)

Edge information is transferred from the source images into the final fused images using the image fusion quality index (Xydeas & Petrovic 2000). Value ranges from 0 to 1. No information is lost between the source images and the final fused image with a value of 1.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 July 2024 **Revised:** 12 August 2024 **Accepted:** 30 September 2024*

Normalized Mutual Information (*QMI)*

For each input image A and B, the total information contained in the fused image F is measured. Having a higher value indicates that the fused image F has more information about A and B. As Qu et al. (2002) explain, QMI stands for Qualitative Multidimensional Interaction.

Information Symmetry or Fusion Symmetry (*FS*)

Fused image FS measures F symmetry with respect to input images A and B.

Piella Metric (*Q* ,*Qw* and *Qe)*

The weighted and edge-dependent quality indices of piella metrics (Piella & Heijmans 2003) are used to determine how good the final image F is in comparison to its source images A and B. Q is defined as the local quality index between A, B, and F.

$$Q(A, B, F) = (Q(A, F) + Q(B, F)) / 2$$

Peak signal to Noise Ratio (*PSNR*): A% signal to noise ratio is a measure of how well an image compares to its reference image. Assume I is the original image, and F is the merged version. In order to calculate the PSNR,

$$PSNR = 20*\log10(255 / \sqrt{MSE})$$

Image I and Image F each have a mean square error of

$$MSE = 1 / M * N \sum \sum (I(x, y) - F(x, y))2$$

Visual Information Fidelity (*VIFF*)

An image visual information content can be measured by the VIFF (Han et al. 2013). Definition refers to

$$VIFF(I1,..., In, IF) = \sum pkVIFFk(I1,..., In, IF)$$

Structural Similarity Index (*SSIM*)

Measurement of brightness and contrast, as well as structural similarities between two image vectors (patches) x and y, is done using SSIM.

RESULTS AND DISCUSSION

Both the global dictionary learning method and the suggested dictionary learning strategy produce comparable fusion performance. In order to learn the global vocabulary of the dictionary utilized in our suggested technique, traditional SR-based methods utilize the K-SVD algorithm. The size of the dictionary and the settings for error tolerance are identical in both kinds of dictionaries. To ensure that the efficacy of the fusion system is not solely reliant on the data utilized in the initial training phase, these training signals are considered. The suggested GDMC dictionary learning procedure outperforms the standard global dictionary learning approach, as demonstrated by the metrics Q AB / F and Q MI. A right-sided homonymous superior quadrant anopia, right-sided color blindness, and anomic fluency (occasional problem finding words) were also detected throughout the examination. There is a medial left occipital infarct indicated on computed tomography (CT) imaging, and an MRI scan shows that the infarct has extended into the left posterior cerebral artery region (see Table 1).

Table 1: Multi-modal image pairs assessment

Methods	*QAB/ F*	*FS*	*QMI*	*Qw*
NSCT	0.4598	0.0694	0.7147	0.4949
Sharp Fusion	0.4922	0.1104	**1.0768**	0.5439
GFF	0.6295	0.0775	0.6884	0.705
MST-SR	0.6188	0.1713	0.7525	0.7926
ASR-256	0.5701	0.0646	0.7208	0.676
ASR-128	0.5508	**0.0635**	0.7136	0.6539
GDMC	**0.6463**	0.1094	0.9644	**0.796**
	QY	***QCB***	***H***	***SD***
NSCT	0.7442	0.5682	4.8813	56.9217
Sharp Fusion	0.7178	0.5907	4.4101	65.1922
GFF	0.8716	0.6154	**5.0442**	64.0711
MST-SR	0.7995	0.645	4.8612	**76.6092**
ASR-256	0.8126	0.5976	4.8391	60.5299
ASR-128	0.8035	0.591	4.8422	60.0236
GDMC	**0.9337**	**0.6788**	4.7073	72.6299

COMPARITIVE ANALYSIS

For comparing the fusion performances of the GDFF and GDMC approaches, we have the following table (Table 2). Examining data sets utilised in the performance evaluation of different methodologies In total, we're looking at 15 pairs of medical images and 10 pairs of visible-infrared image pairs in the multi-modal medical data sets we're looking at. As can be seen from the table, the GDFF approach is better suited for the fusion of visible-infrared image pairings than the GDMC approach.

Table 2: Comparison between GDLC and GDMC approaches

Image Sets	GDLC				
	QAB/F	*QMI*	*Qw*	*QY*	*QCB*
Visible- Infrared Image Pairs (10)	**0.6083**	**0.6369**	**0.8434**	**0.8763**	**0.6083**
Medical Image Pairs (15)	0.635	0.8447	0.7751	0.9227	0.6494
Image Sets	**GDMC**				
	QAB/F	*QMI*	*Qw*	*QY*	*QCB*
Visible- Infrared Image Pairs (10)	0.5923	0.5369	0.8089	0.8579	0.5986
Medical Image Pairs (15)	**0.6463**	**0.9644**	**0.796**	**0.9337**	**0.6788**

More than a third of the initial training data set was used to build the global dictionary-based on focus features classification (GDFF). The impact of dictionary size on fusion performance has been studied, and an optimum size of 300 has been determined. With these parameters, the GDFF methodology provides promising results compared to existing methodologies, while the dictionary learning process is decreased by 92.4%.

CONCLUSION

The proposed Semantic Sparse Recoding framework successfully bridges the gap between low-level pixel data and high-level semantic understanding in visual content representation. By adopting a Global Dictionary Learning approach, the method achieves significant improvements in the quality of fused images while drastically reducing computational complexity. Comparative analyses with existing methods underscore the system's robustness and efficiency in diverse applications, such as medical imaging and surveillance. Future work will explore further integration of deep learning techniques to enhance the scalability and adaptability of the SSR approach to broader image processing challenges.

REFERENCES

1. Yang, M., Xu, Y., Wu, Z., & Wei, Z. (2022). Symmetrical irregular local features for fine-grained visual classification. *Neurocomputing, 505*, 304-314. https://doi.org/10.1016/j.neucom.2022.07.056
2. Preethi, S. J., & Krupa, N. B. (2023). Analyzing lower half facial gestures for lip reading applications: Survey on vision techniques. *Computer Vision and Image Understanding, 233*, 103738. https://doi.org/10.1016/j.cviu.2023.103738
3. Ngo, C. T., Buchberger, E. S., Nguyen, P. T. U., Newcombe, N. S., & Werkle-Bergner, M. (2024). Building a cumulative science of memory development. *Developmental Review, 72*, 101119. https://doi.org/10.1016/j.dr.2024.101119
4. Zhang, X., Li, J., & Hua, Z. (2022). MFFE: Multi-scale Feature Fusion Enhanced Net for image dehazing. *Signal Processing: Image Communication, 105*, 116719. https://doi.org/10.1016/j.image.2022.116719
5. Zhang, S., Liu, Y., Xiong, K., Tian, Y., Du, Y., Zhu, Z., Du, M., & Zhai, Z. (2024). A review of vision-based crop row detection methods: Focusing on field ground autonomous navigation operations. *Computers and Electronics in Agriculture, 222*, 109086. https://doi.org/10.1016/j.compag.2024.109086
6. Shamsi, A., Nezamabadi-Pour, H., & Saryazdi, S. A short-term learning approach based on similarity refinement in content-based image retrieval. International Journal of Multimedia tools and applications. 2014: 72(2): 2025-2039.
7. Shanmugapriya, N., & Nallusamy, R. A new content based image retrieval system using GMM and relevance feedback. Journal of Computer Science. 2014: 10(2): 330-339.
8. Shen, B, Hu, W, Zhang, Y & Zhang, Y 2009, ‘Image inpainting via sparse representation’, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 697-700.
9. Shreyamsha Kumar, BK 2013, ‘Multi-focus and

multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform', Signal, Image & Video Processing, vol. 7, no. 6, pp. 1125-1143.
10. Singh, R, Srivastava, R, Prakash, O, & Khare, A 2012, 'Multimodal Medical Image Fusion in Dual Tree Complex Wavelet Transform Domain Using Maximum and Average Fusion Rules', Journal of Medical imaging and Health Informatics, vol. 2, pp. 168-173.
11. Singhai, N. and Shandilya, S.K., 2010. A survey on: content based image retrieval systems. International Journal of Computer Applications. 2010: 4(2): 22-26.
12. Sollie, P 2003, Morphological principles and applications, Springer- Verlag, Berlin.
13. Stathaki, T 2008, 'Image Fusion: Algorithms and Applications', Academic Press.
14. Tatu, A., Lauze, F., Nielsen, M., & Kimia, B. (2011). Exploring the representation capabilities of the HOG descriptor. In *ICCV workshop*.
15. Tian, J & Chen, L 2012, 'Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure', Signal Processing, vol. 92, no. 9, pp. 2137-2146.
16. Tian, J, Chen, L, Ma, L & Yu, W 2011, 'Multi-focus image fusion using a bilateral gradient-based sharpness criterion', Optical Communication, vol. 284, no. 1, pp. 180-187.
17. Toet, A 1989, 'A morphological pyramidal image decomposition', Pattern Recogn. Letters, vol. 9, no. 4, pp. 255-226.
18. Toet, A, Ruyven, LJ & Valeton, JM 1989, 'Merging thermal and visual images by a contrast pyramid', Opt. Eng., vol. 28, no. 7, pp. 789-792.
19. Tropp, JA, Gilbert, AC & Strauss, MJ 2006, 'Strauss, algorithms for simultaneous sparse approximation', Part I: greedy pursuit, Signal Process, vol. 86, no. 3, pp. 572-588.
20. Velmurugan, K., & Baboo, L. D. S. S. Content-based image retrieval using SURF and colour moments. Global Journal of Computer Science and Technology. 2011: 11(10): 42-51.
21. Vogel, J. and Schiele, B. Semantic modeling of natural scenes for content- based image retrieval. International Journal of Computer Vision. 2007:72(2):133-157.
22. Wang, D., Hoi, S. C., He, Y., & Zhu, J. Mining weakly labeled web facial images for search-based face annotation. IEEE Transactions on Knowledge and Data Engineering. 2014: 26(1): 166-179.
23. Wang, J, Peng, J, Feng, X, He, G, Wu, J & Yan, K 2013, 'Image fusion with non-subsampled contourlet transform and sparse representation', Journal of Electron Imaging, vol. 22, no. 4, 043019, pp. 1-15.
24. Wright, J, Ma, Y, Mairal, J, Sapiro, G, Huang, TS & Yan, S 2010, 'Sparse representation for computer vision and pattern recognition', Proc. IEEE, vol. 98, no. 6, pp. 1031-1044.
25. Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feed forward neural networks. International conference on artificial Intelligence and statistics. 2010: 2(1):261-269.
26. Xu, D., Yan, S., Tao, D., Lin, S., & Zhang, H. J. Marginal fisher analysis and its variants for human gait recognition and content-based image retrieval. IEEE Transactions on Image Processing. 2007: 16(11): 2811-2821.
27. Xydeas, CS & Petrovic, V 2000, 'Objective image fusion performance measure', Electronic Letters, vol. 36, no. 4, pp. 308-309.
28. Yang, B & Li, 2012, 'Pixel-level image fusion with simultaneous orthogonal matching pursuit' Information Fusion, vol. 13, no. 11, pp. 10-19.
29. Yang, B & Li, S 2010, 'Multi-focus image fusion and restoration with sparse representation', IEEE Transactions on Instrumentation and Measurement, vol. 59, no. 4, pp. 884-892.
30. Bai, X, Zhang, Y, Zhou, F & Xue, B 2015, 'Quad tree-based multi- focus image fusion using a weighted focus measure', Information Fusion, vol. 22, pp. 105-118.
31. Aslantas, V, Bendes, E, Toprak, AN & Kurban, R 2011, 'A comparison of image fusion methods on visible, thermal and multi- focus images for surveillance applications', Proceedings of IET conference on imaging for crime detection and prevention, pp. 1-6.