Models for Enhancing Autonomous Driving Accuracy

Subjects: Computer Science, Artificial Intelligence

Contributor: Henry Alexander Ignatious , Hesham El-Sayed , Manzoor Ahmed Khan , Parag Kulkarni

Higher-level autonomous driving necessitates the best possible execution of important moves under all conditions. Most of the accidents caused by the AVs launched by leading automobile manufacturers are due to inadequate decision-making, which is a result of their poor perceivance of environmental information. In today's technology-bound scenarios, versatile sensors are used by AVs to collect environmental information. Due to various technical and natural calamities, the environmental information acquired by the sensors may not be complete and clear, due to which the AVs may misinterpret the information in a different context, leading to inadequate decision-making, which may then lead to fatal accidents. To overcome this drawback, effective preprocessing of raw sensory data is a mandatory task. Pre-processing the sensory data involves two vital tasks, namely data cleaning and data fusion. Since the raw sensory data are complex and exhibit multimodal characteristics, more emphasis is given to data preprocessing.

autonomous vehicles (AVs)   data fusion   situation awareness   data-preprocessing

machine learning (ML)

1. Introduction

Autonomous cars have the ability to cut traffic accidents significantly. This hypothesis is based on the fact that autonomous cars reduce the influence of a human component on the likelihood of a traffic collision. Autonomous vehicles are being tested in real-life traffic circumstances all around the world [1]. According to [1], if AVs become accessible in 2025, the market in the United States of America might reach 8 million in ten years. The widespread deployment of autonomous vehicles is intended to minimize collisions, relieve traffic congestion, increase fuel efficiency, reduce parking demands, and provide mobility to individuals who are unable to drive [2]. The same survey predicted that by 2040 about 33 million autonomous vehicles will be on the road. This just means that people will have to become acclimated to self-driving automobiles. This is supported by facts and data, yet further testing is needed to ensure everyone's safety.

In order to avoid accidents caused by the AVs, the decisions taken by the AVs to overcome roadside events must be instant and accurate [3]. For accurate decision-making, the environmental data perceived by the AVs must be clear and accurate. Perception systems must be exact in order to provide a precise comprehension of the world. They must be able to work in adverse conditions and even when particular sensors are broken or defective. In order to collect environmental data, as well as data relating to autonomous vehicle characteristics, sensor systems

must be functional. However, data received from a variety of devices, including sensors, thermal cameras, radars, and so on, have heterogeneous multimodal features, making it difficult to achieve the correct perception. In order to enhance the situation awareness of the AV, there has to be more emphasis on the data preprocessing tasks of the AVs. Data preprocessing involves data cleaning and multimodal fusion. The upcoming paragraphs highlight the relationship between situation awareness and data preprocessing.

According to [4], situation awareness (SA) is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". Situational awareness acts as a bridge connecting our comprehension of the environment, ongoing real-time occurrences, and potential future developments. To establish the situation awareness of their perceived environment, AVs collect environmental data using external and internal sensors, namely LiDAR, radar, and ultrasonic, followed by stereo and thermal cameras. The vehicular information collected can exist in different formats, such as textual, image, video, and audio, and also exhibit many discrepancies. The discrepancies exhibited by the vehicular data might be in the form of irrelevant data, missing data, outliers, and duplicate data. In addition, images may contain missing data in the form of blurred images and audio data that contains noisy irrelevant data. Hence, preprocessing the sensory data is essential to improve the accuracy. Since data cleaning has been well studied, this proposal focuses on the second task of data preprocessing, namely data fusion. The sensory data exhibit multimodal characteristics, hence an effective data fusion model is mandatory to convert the heterogeneous data into a unique format and also to enhance their accuracy for further effective processing.

Most of the studies fail to propose a generic multimodal fusion methodology to handle the diversity existing among different datasets. The relevant research literature also does not clearly explain key operations such as feature selection and dimensionality reduction of multimodal data, the mechanisms for 2D to 3D multimodal data transformation and storage, and the methodology for converting multimodal data to a single unique data format. Further very few contributions have been carried out to fuse multispectral environment data collected from sensors and satellites. Versatile fusion models with advanced image processing and machine learning techniques are required to fuse the multispectral high-resolution data. The accuracy level reached by most of the referred decision-making frameworks and models is around 85%. This statement proves that more emphasis must be given to preprocessing, especially in data fusion tasks to improve the data accuracy, which has an impact on enhancing the situation awareness of AVs to improve their accuracy of decision-making.

2. Hybrid Image Fusion Models

B. Shahian Jahromi et al. [5] have proposed a novel hybrid multi-sensor fusion pipeline configuration for autonomous cars that handles environment perception tasks such as road segmentation, obstacle identification, and tracking. A suggested encoder–decoder-based fully convolutional neural network (FCNx) and a standard extended Kalman filter (EKF) nonlinear state estimator approach are used in this fusion framework. It also employs optimal camera, LiDAR, and radar sensor configurations for each fusion approach. The purpose of this hybrid architecture is to create a fusion system that is cost-effective, lightweight, adaptable, and resilient (in the event of a sensor failure). It employs the FCNx algorithm, which improves road identification accuracy above benchmark

models while preserving real-time efficiency in an embedded computer for autonomous vehicles. D. Jia et al. [6] have presented a hybrid spatiotemporal fusion (STF) technique based on a deep learning model called the hybrid deep-learning-based spatiotemporal fusion model (HDLSFM). With a minimum amount of input, the method develops a hybrid framework for the reliable fusion of morphological and physiological data that explains the physical material at the surface of the earth. To handle radiation discrepancies across various types of satellite pictures, the suggested method combines a regressive deep-learning-based related radiometric normalization, a deep-learning-based super-resolution, and a linear-based fusion. Using Fit-FC as a benchmark, the HDLSFM's propensity to predict phenological and land-cover change has been demonstrated. Meanwhile, HDLSFM is immune to changes in radiation across different types of satellite images as well as the time interval between the forecast and base dates, assuring its usefulness in the synthesis of fused time-series data.

Y. Wang et al. [7] have proposed a hybrid fusion strategy that takes into consideration the geographical and semantic properties of sensor inputs concerning occurrences. To achieve this, the authors have used Cmage, an image-based representation for both physical and social sensor data that describes the situation of certain visual notions (e.g., "crowdedness" and "people marching"). The authors have proposed a fusion model that describes sparse sensor information using a Gaussian process based on the acquired Cmage representation, which combines multimodal event signals with a Bayesian method and integrates spatial links between the sensor and social data. A. V. Malawade et al. [8] have proposed a selective sensor fusion framework, namely HydraFusion, which learns to recognize the present driving environment and then combines the appropriate mix of sensors to enhance robustness without sacrificing efficiency. HydraFusion is the first method to suggest dynamically shifting between early fusion, late fusion, and combinations in between, so modifying both how and when fusion is used. On the industry-standard Nvidia Drive PX2 AV hardware platform, the authors show that HydraFusion outperforms early and late fusion techniques by 13.66 percent and 14.54 percent, respectively, without increasing computing complexity or energy consumption. Both static- and deep-learning-based context identification algorithms are proposed and evaluated by the authors.

Y. Zhao et al. [9] have proposed a hybrid spatial-temporal-spectral image fusion model (HSTSFM) for simultaneously generating synthetic satellite data with high spatial, temporal, and spectral resolution (STSR), which blends the high spatial resolution from the panchromatic image of the Landsat-8 Operational Land Imager (OLI), the high temporal resolution from the multispectral image of the Moderate Resolution Imaging Spectroradiometer (MODIS), and the high spectral resolution from the hyper-spectral image of Hyperion to produce high spatial–spectral image fusion, high spatial–temporal image fusion, and high temporal–spectral image fusion, which are the three fusion modules included in the proposed HSTSFM. To show the performance of the proposed technique, a set of test data containing both phenological and land cover type changes in Beijing suburbs, China, are used. B. Latreche et al.[10] have suggested an effective hybrid image fusion approach based on the integer lifting wavelet transform (ILWT) and the discrete cosine transformer (DCT) that are suited for video streaming networks (VSNs). There are two phases in the proposed fusion algorithm. To begin, the ILWT approximation coefficients (low frequencies) are fused by using the variance as an activity level measure in the DCT domain. Second, the high-frequency detail coefficients are fused using the best-weighted average based on the correlation between coefficients in the ILWT domain. The suggested solution addresses information loss, computational complexity,

time and energy consumption, and memory space due to integer operations in the ILWT domain. Extensive tests have been carried out to show that the suggested method outperforms other picture fusion algorithms in the literature, both intuitively and numerically.

X. Zhang et al. [11] have proposed a multi-focus image fusion benchmark (MFIFB), that includes a test set of 105 picture pairings, a code library of 30 MFIF algorithms, and 20 evaluation measures. MFIFB is the first MFIF benchmark, providing a forum for the community to assess MFIF algorithms thoroughly. To understand the performance of these algorithms further, extensive tests have been carried out utilizing the suggested MFIFB. Effective MFIF algorithms are found by examining the experimental findings. More significantly, some remarks on the current state of the MFIF field are provided, which might aid in a better understanding of this topic. D. Kaimaris and A. Kandylas [12] have suggested an innovative mechanism to obtain multispectral image data using UAVs and fuse them to improve the accuracy of the data. The photos from Parrot's tiny multispectral (MS) camera Sequoia+ are examined at two ancient sites: a Byzantine wall (ground application) in Thessaloniki, Greece, and a mosaic floor (aerial application) at the archaeological site of Dion, Greece. The camera obtains RGB and MS pictures at the same time, which prevents image fusion, as is the case with the conventional use of panchromatic (PAN) and MS images in satellite passive systems. Using the image fusion methods of satellite PAN and MS pictures, this research shows that effective digital processing of the images (RGB and MS) of tiny MS cameras may result in a fused image with a high spatial resolution that maintains a considerable proportion of the original MS image's spectral information. The great spectrum fidelity of the fused pictures allows for high-precision digital measurements in ancient sites, such as precise digital item separation, area measurements, and recovery of information not apparent with standard RGB sensors using MS and RGB data from tiny MS sensors. Reference [13] has proposed a versatile hybrid fusion model to fuse infrared and visible image fusion models. They have used the combined concepts of visibility enhancement and multiscale decomposition to fuse the images. Initially, the authors proposed an effective preprocessing model followed by a decomposition model to decompose the information into the layers of their customized CNN model. Further, they have integrated the concepts of a visual saliency illumination map (VSIM) to retain the contrast information and enhance the fusion process. Reference [14] has proposed a hybrid image fusion model to fuse medical images that exhibit multimodal characteristics. They have used the dual combination of nonsubsampled contourlet transform (NSCT) and discrete wavelet transform (DTCWT) approaches to fuse the images. The authors have used an advanced CNN model to create weight maps to monitor the pixel movement of the images. Further, the authors have included an advanced comparison-based method to convert the fusion mode to the appropriate coefficients required for the CNN model.

Since the research has extended its work by proposing a versatile GAN model to fuse all types of advanced image data, the research has explored the contributions of researchers in image fusion using GAN models. Reference [15] has proposed a novel hybrid image fusion model using GAN techniques called PAN-GAN. The author's model is used to fuse panchromatic images. The PAN-GAN model uses a separate adversarial mechanism that establishes a bond with the discriminators to preserve the spectral and spatial information of the fused images. Similarly, reference [16] has proposed an innovative fusion model, namely GAN-FM, which uses the GAN principle to fuse infrared and visible images. The authors have designed a full-scale skip-connected generator along with discriminators based on Markovian principles for extracting features at different scales and establishing a link with

the generators to retain the contrast of the fused images. In yet another interesting study, reference [17] has proposed an innovative hybrid image fusion model, namely THFuse, which uses GAN approaches to fuse infrared and visible images. The authors have used advanced fusion strategies such as transformer and hybrid feature extraction concepts to process both global and local image information. Reference [18] has proposed a versatile image fusion model called mask deep fusion network for visible and infrared image fusion (MDFN). The authors have proposed a novel mechanism to compute the weight score for every pixel to estimate the contributions of the two input source images. This operation transfers valuable information from source to fused images, helping them to retain their contrast. Reference [19] has suggested a hybrid image fusion model, namely the pair feature difference guided network (FDGNet), to fuse multimodal medical images. The authors have proposed a weight-guided mechanism to extract the features from complex medical images efficiently. Further, the authors have introduced a factor, namely hybrid loss, composed of weight fidelity loss and feature difference loss to train the network effectively.

Since this research focuses on the feature extraction process, detailed literature related to some of the proposed image feature extraction methods is analyzed. This proposed research plans to extract four important image features, namely color, edge, height, and width. In their recent publication, Li et al. [20] introduced a cutting-edge generative adversarial network named MSAt-GAN. This novel model incorporates multiscale feature extraction and deep attention techniques to merge infrared and visible images seamlessly. By utilizing three distinct fields for feature extraction, the model enhances the accuracy of data fusion. Moreover, the deep attention mechanism facilitates the extraction of multi-level features through spatial and channel attention, thus enabling effective data fusion. Reference [21] in their work, have introduced a versatile fusion model called multi-exposure image fusion on generative adversarial networks (MEF-GAN) with the aim of effectively fusing image data. The proposed model consists of two components: a generator and a discriminator network, which are trained concurrently to form an adversarial network. The generator is responsible for producing synthesized fused images that resemble the source image, while the discriminator is trained to differentiate between the source image and the fake fused images generated by the generator. This adversarial relationship helps to preserve data integrity and prevents information loss in the fused image, ultimately leading to a fused image probability distribution that closely approximates reality.

In their recent study, reference [22] proposed a robust fusion model called correlation-driven feature decomposition fusion (CDDFuse) as an effective solution. The authors employed Restormer blocks to extract cross-modality image features and seamlessly integrate them with an advanced convolutional neural network (CNN) model. Additionally, Lite Transformer (LT) blocks were incorporated to extract low-level features. To establish the correlation between low-frequency and high-frequency features, the authors introduced a correlation-based loss factor. By leveraging the proposed LT model and invertible neural networks (INN), the authors successfully fused the low- and high-frequency features, resulting in the generation of the fused image. Reference [23] in their recent publication, introduced a multi-focus image fusion model that combines the principles of Transformers and an advanced CNN model to fuse multimodal image data effectively. By incorporating both local information from the CNN model and global information from the transformers, the accuracy of fusion is significantly improved.

Furthermore, the authors proposed a feedback mechanism that maximizes the utilization of features, thereby enhancing the performance of the networks in feature extraction.

3. Feature Extraction Models (Image Data)

P. Tiede et al.[24] have proposed a novel universal image feature extraction approach called variational image domain analysis, which is used for a wide range of very long baseline interferometry (VLBI) image reconstructions. Variational image domain analysis, unlike earlier methods, may be used for any image reconstruction, independent of its structure. The authors' approach gives clear ideas on how to extract salient image features such as color and edge. Y. Liu, H et al.[25] have customized a CNN model to extract deep features of images related to food. The CNN model, when paired with nondestructive detection techniques and a computer vision system, has great potential for identifying and analyzing complex food matrices. CNN-based features outperform handmade or machine-learning-based features. N. Liang et al. [26] have proposed a multi-view structural feature extraction approach to provide a thorough characterization of spectral–spatial structures of various objects, which consists mostly of the stages below. First, the original image's spectral number is reduced using the minimum noise fraction (MNF) approach, and then the local structural feature is extracted from the dimension-reduced data using a relative total variation. The nonlocal structural characteristics from intra-view and inter-view are then produced using a superpixel segmentation approach that takes into account the intra- and inter-similarities of superpixels. The final picture characteristics for classification are formed by combining the local and nonlocal structural features. S. Barburiceanu et al. [27] have presented a texture feature extraction approach with increased discriminating power for volumetric pictures. The technique is used to classify textured volumetric data. The authors employ feature vectors obtained from local binary patterns (LBP) and the gray-level co-occurrence matrix-based approach to combine two complementing types of information. R. Ahmed Bhuiyan et al.[28] have provided a feature extraction methodology for human activity recognition that is both efficient and low in dimension. The enveloped power spectrum (EPS) is employed in this feature extraction approach to recover impulse components of the signal utilizing frequency domain analysis, which is more robust and noise intolerant. For human activity recognition, linear discriminant analysis (LDA) is utilized as a dimensionality reduction approach to extract the smallest amount of discriminant characteristics from the envelope spectrum, human activity recognition (HAR). A multi-class support vector machine (MCSVM) is used to recognize human activities using the derived characteristics. To extract robust features, Bo do et al. [29] have used a stacked convolutional denoising autoencoder (SCDAE), which reduces susceptibility to partially damaged data, or input data that are partially missing. Trial-and-error experiments were used to optimize SCDAE parameters such as network depth, number of convolution layers, number of convolution kernels, and convolution kernel size.

The analysis identified some gaps in both the explored image fusion models and feature extraction models. Regarding image fusion, most of the referred image models do not produce better accuracy. A generic fusion model to fuse all formats of data is missing. Most referred studies use the minimum dataset to evaluate their models. Complicated operations such as image transformation (2D to 3D) and other image functionalities are not transparent in many studies. Moreover, there are minimum contributions related to multispectral image fusion. Most

of the referred models are complex and require advanced algorithms and techniques. There is a need for developing computationally efficient fusion algorithms that can operate in real-time or near real-time scenarios without sacrificing the quality of the fused images. Exploring techniques such as model compression, hardware acceleration, and parallel processing can help bridge this gap.

Implementing and fine-tuning these models can be challenging, requiring significant computational resources and expertise. Image fusion is a subjective task, and the quality of the fused image can vary depending on individual preferences and application requirements. Most referred models involve multiple parameters and design choices, making it difficult to determine an optimal fusion result that satisfies everyone. There is still room for exploring more efficient and effective deep-learning architectures specifically designed for hybrid image fusion. Research should focus on developing novel network architectures, attention mechanisms, and loss functions that can capture complementary information from multiple input images and improve the fusion quality. Most models often lack interpretability and explainability. It is challenging to understand the decision-making process and the contribution of different input images in the fusion result. Further research is needed to develop techniques that can provide insights into the fusion process, visualize the information fusion at different stages, and offer explanations for the final fusion outcome.

Regarding the referred feature extraction models, most of the referred models are complicated and require high computational costs. Most models depend on CPU utilization time and memory. Further, the feature detection models depend on the experiences of the designer. Having the specified gaps as the point of motivation, this research proposes a generic data fusion engine to fuse all formats of data and also proposes innovative strategies to extract the salient features of image and audio data. Many feature extraction models are trained and optimized for specific datasets or domains. However, there is a need for models that can generalize well across different domains, such as medical imaging, natural images, satellite imagery, and more. Developing domain-agnostic feature extraction models that can capture and represent diverse types of data effectively remains a challenge. With the increasing demand for real-time and large-scale applications, there is a need for feature extraction models that are efficient and scalable. Developing lightweight architectures and techniques for efficient feature extraction, model compression, and hardware acceleration is an ongoing research direction to enable faster and more resource-efficient feature extraction.

To address the identified gaps, this research introduces effective models for feature extraction that can extract image features from all types of image data. Additionally, a hybrid image fusion model is proposed to fuse 2D and 3D multispectral image data. The research presents advanced projection and image transformation formulas to enhance the efficiency of the image fusion process. However, the performance of the proposed hybrid image fusion model is found to be unsatisfactory when applied to 3D point cloud data and when dealing with large image datasets. To overcome these limitations, the research expands its scope by proposing an innovative image fusion model that incorporates advanced concepts from the GAN (Generative Adversarial Network) model. This new model performs various tasks, including advanced feature extraction to capture both spatial and spectral information, as well as generator and discriminator modules to facilitate fusion tasks while preserving image quality.

Customized kernel functions are introduced for the CNN (Convolutional Neural Network) layers to execute the specified tasks.

References

1. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning . National Academies. Retrieved 2023-10-11

2. Daniel J. Fagnant; Kara Kockelman; Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp. Res. Part A: Policy Pr.* **2015**, *77*, 167-181.

3. Hitesh Mohapatra; Amiya Kumar Rath; Nibedan Panda; IoT infrastructure for the accident avoidance: an approach of smart transportation. *Int. J. Inf. Technol.* **2022**, *14*, 761-768.

4. Mica R. Endsley; Design and Evaluation for Situation Awareness Enhancement. *null* **1988**, *32*, 97-101.

5. Babak Shahian Jahromi; Theja Tulabandhula; Sabri Cetin; Real-Time Hybrid Multi-Sensor Fusion Framework for Perception in Autonomous Vehicles. *Sensors* **2019**, *19*, 4357.

6. Terzic J, Terzic E, Nagarajah R, Alamgir M, Terzic J, Terzic E,Nagarajah R, Alamgir M Ultrasonic sensing technology. Ultrasonic Fluid Quantity Measurement in Dynamic Vehicular Applications: A Support Vector Machine Approach.. *Springer International* **2013**, *nil*, 11-35.

7. Yuhui Wang; Christian von der Weth; Yehong Zhang; Kian Hsiang Low; Vivek K. Singh; Mohan Kankanhalli. Concept Based Hybrid Fusion of Multimodal Event Signals; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, United States, 2016; pp. 14-19.

8. Arnav Vaibhav Malawade; Trier Mortlock; Mohammad Abdullah Al Faruque. HydraFusion: Context-Aware Selective Sensor Fusion for Robust and Efficient Autonomous Vehicle Perception; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, United States, 2022; pp. 68-79.

9. Zhao Y, Huang B A hybrid image fusion model for generating high spatial-temporal-spectral resolution data using OLI-MODIS-Hyperion satellite imagery. *International Journal of Geological and Environmental Engineering* **2017**, *11*, 869-74.

10. Boubakeur Latreche; Slami Saadi; Mecheri Kious; Ali Benziane; A novel hybrid image fusion method based on integer lifting wavelet and discrete cosine transformer for visual sensor networks. *Multimedia Tools Appl.* **2018**, *78*, 10865-10887.

11. Xingchen Zhang. Multi-focus Image Fusion: A Benchmark; https://arxiv.org/pdf/2005.01116.pdf, Eds.; null: null, 2020; pp. null.

12. Dimitris Kaimaris; Aristoteles Kandylas; Small Multispectral UAV Sensor and Its Image Fusion Capability in Cultural Heritage Applications. *Heritage* **2020**, *3*, 1046-1062.

13. Yueying Luo; Kangjian He; Dan Xu; Wenxia Yin; Wenbo Liu; Infrared and visible image fusion based on visibility enhancement and hybrid multiscale decomposition. *Opt.* **2022**, *258*, 168914.

14. Nawar Alseelawi; Hussein Tuama Hazim; Haider Th.Salim Alrikabi; A Novel Method of Multimodal Medical Image Fusion Based on Hybrid Approach of NSCT and DTCWT. *Int. J. Online Biomed. Eng. (iJOE)* **2022**, *18*, 114-133.

15. Jiayi Ma; Wei Yu; Chen Chen; Pengwei Liang; Xiaojie Guo; Junjun Jiang; Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Inf. Fusion* **2020**, *62*, 110-120.

16. Hao Zhang; Jiteng Yuan; Xin Tian; Jiayi Ma; GAN-FM: Infrared and Visible Image Fusion Using GAN With Full-Scale Skip Connection and Dual Markovian Discriminators. *IEEE Trans. Comput. Imaging* **2021**, *7*, 1134-1147.

17. Jun Chen; Jianfeng Ding; Yang Yu; Wenping Gong; THFuse: An infrared and visible image fusion network using transformer and hybrid feature extractor. *Neurocomputing* **2023**, *527*, 71-82.

18. Chaoxun Guo; Dandan Fan; Zhixing Jiang; David Zhang; MDFN: Mask deep fusion network for visible and infrared image fusion without reference ground-truth. *Expert Syst. Appl.* **2023**, *211*, 118631.

19. Gucheng Zhang; Rencan Nie; Jinde Cao; Luping Chen; Ya Zhu; FDGNet: A pair feature difference guided network for multimodal medical image fusion. *Biomed. Signal Process. Control.* **2023**, *81*, 104545.

20. Junwu Li; Binhua Li; Yaoxi Jiang; Weiwei Cai; MSAt-GAN: a generative adversarial network based on multi-scale and deep attention mechanism for infrared and visible light image fusion. *Complex Intell. Syst.* **2022**, *8*, 4753-4781.

21. Han Xu; Jiayi Ma; Xiao-Ping Zhang; MEF-GAN: Multi-Exposure Image Fusion via Generative Adversarial Networks. *IEEE Trans. Image Process.* **2020**, *29*, 7203-7216.

22. Zhao Z, Bai H, Zhang J, Zhang Y, Xu S, Lin Z, Timofte R, Van Gool L. Cddfuse: Correlation-driven dual-branch feature decomposition for multi-modality image fusion. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023 (pp. 5906-5916).

23. Xuejiao Wang; Zhen Hua; Jinjiang Li; Multi-focus image fusion framework based on transformer and feedback mechanism. *Ain Shams Eng. J.* **2023**, *14*, 101978.

24. Paul Tiede; Avery E. Broderick; Daniel C. M. Palumbo; Variational Image Feature Extraction for the Event Horizon Telescope. *Astrophys. J.* **2022**, *925*, 122.

25. Yao Liu; Hongbin Pu; Da-Wen Sun; Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends Food Sci. Technol.* **2021**, *113*, 193-204.

26. Nannan Liang; Puhong Duan; Haifeng Xu; Lin Cui; Multi-View Structural Feature Extraction for Hyperspectral Image Classification. *Remote. Sens.* **2022**, *14*, 1971.

27. Stefania Barburiceanu; Romulus Terebes; Serban Meza; 3D Texture Feature Extraction and Classification Using GLCM and LBP-Based Descriptors. *Appl. Sci.* **2021**, *11*, 2332.

28. Rasel Ahmed Bhuiyan; Nadeem Ahmed; Amiruzzaman; Rashedul Islam; A Robust Feature Extraction Model for Human Activity Characterization Using 3-Axis Accelerometer and Gyroscope Data. *Sensors* **2020**, *20*, 6990.

29. Bo Du; Wei Xiong; Jia Wu; Lefei Zhang; Liangpei Zhang; Dacheng Tao; Stacked Convolutional Denoising Auto-Encoders for Feature Representation. *IEEE Trans. Cybern.* **2016**, *47*, 1017-1027.

Retrieved from https://encyclopedia.pub/entry/history/show/113440