International Journal of Applied Science; Vol. 8, No. 2; 2025
ISSN 2576-7240 E-ISSN 2576-7259
https://doi.org/10.30560/ijas.v8n2p14

Application of Machine Vision in Structural Deformation and Health Monitoring

Zhonglai Qin[1]

[1] Dalian University of Technology, China

Correspondence: Zhonglai Qin, Dalian University of Technology, Dalian, China. E-mail: 201953220@mail.dlut.edu.cn

Received: March 15, 2025; Accepted: April 3, 2025; Published: April 6, 2025

Abstract

There are two main methods of structural health monitoring. Among them, the traditional structure health monitoring technology is mainly artificial, which has some problems such as inflexibility, large error, high cost and difficult to adapt to the change of natural environment. The structural health monitoring method based on machine vision has the characteristics of flexible measuring point, high accuracy, fast speed and no contact. A variety of machine vision technologies, such as image acquisition, image processing, three-dimensional vision and deep learning technologies, have made great progress, and their application scenarios are constantly expanding. This paper is based on infrared image acquisition technology, 3D vision acquisition technology, image stitching technology, stereo vision technology four perspectives. The improvement of machine vision technology in the field of structural health monitoring is described in detail. The development trend of machine vision acquisition technology and image processing technology and the huge growth space of machine vision technology are analyzed. The application range of these key technologies is also introduced. The application results in building crack detection, seepage detection and fire detection are summarized, and the future development of this technology is prospected from the aspects of algorithm robustness.

Keywords: intelligent analysis, accurate monitoring,artificial intelligence, Computer Technology Applications, deep learning

1. Introduction

Monitoring the displacement and health status of built structures is of great significance. Bridges, as the arteries of the city, carry a large amount of traffic, and the number is increasing day by day. However, with the increase in the service life of the bridge, the safety problem is increasingly concerned.

Now, the main traditional means used in bridge structure monitoring and health status monitoring are still manual monitoring (Si et al., 2020), such as the use of vernier calipers (Si et al., 2020) and ultrasonic (Wulff et al., 1999). However, there are many problems with the above manual monitoring methods, such as high cost, long time, and large errors. As shown in Figure. 1, when the monitored object is a high-risk structure or in an extreme environment, the monitoring personnel need to bear greater work risks, and even threaten the safety of life.

Figure 1. The construction workers work in danger

Machine vision technology originated in the 1950s. With the continuous development of computer technology and the progress of artificial intelligence algorithms in recent years, machine vision technology has made remarkable progress. Thanks to deep learning, improved hardware performance and the popularity of big data, machine vision has made significant progress in terms of resolution, accuracy, three-dimensional vision, real-time and environmental adaptability. It has gradually been applied to more and more fields, such as building health monitoring (Spencer et al., 2019), medical imaging (Ker et al., 2018), intelligent robotics (Bai et al., 2023). Especially in the field of structural health monitoring, replacing traditional manual monitoring with machine vision-based measurement technology has become a hot spot and trend of current research.

However, current research is mainly focused on the research and optimization of a single machine vision technology, resulting in a lack of a comprehensive overview of the development of machine vision technology and future research trends. Therefore, it is necessary to conduct a comprehensive review of the application of machine vision in the monitoring of structural displacement and health status.

In this paper, the core technology of machine vision is introduced in detail from four aspects: infrared image acquisition, 3D vision acquisition, image stitching and stereo vision, and the application of these technologies in structural health monitoring is discussed. The paper also analyzes the development trend of machine vision acquisition technology and image processing technology, and points out the huge growth potential of this field. At the same time, the application range of the current key technologies is described, and the practical results achieved in building crack detection, seepage detection and fire detection are summarized. Finally, the future development direction of the technology, especially in the robustness of the algorithm, is prospected.

2. Machine Vision Equipment Development

With the continuous innovation and upgrading of machine vision equipment, machine vision cameras can be divided into the following main types according to different application requirements, technical characteristics, and imaging methods, as shown in Figure. 2. The following is a list of several major machine vision equipment combined with their evolution to discuss the development of machine vision equipment.

2.1 CCD and CMOS

Traditional machine vision cameras can be divided into CCD cameras and CMOS cameras according to the different sensors; CCD cameras have higher image quality and lower noise performance, and CMOS cameras have lower cost and power consumption than CCD cameras and have gradually replaced traditional CCD cameras (Fossum et al., 2014).

Figure 2. Classification of machine vision equipment

2.2 Infrared Camera

Based on the spectral range, it can be divided into visible light cameras, near-infrared cameras, and thermal imaging cameras, the difference between them is mainly the length of the perception band, respectively, 400 nm-700 nm (visible light), 700 nm-1400 nm (near infrared) and 3 µm-14 µm (infrared radiation). Visible light camera mainly captures the reflected visible light, suitable for daily shooting and video, the image effect is close to the human eye. Near-infrared cameras can sense infrared light that is invisible to the human eye and are mainly used for shooting in low-light environments and specialized spectral analysis, such as night vision and medical imaging.

Thermal imaging cameras mainly image through temperature differences, which are suitable for fire rescue, building inspection, and other fields to provide visualization of temperature information.

As shown in Figure 3, infrared imagers have become increasingly widely used in building structure monitoring, and special handheld imaging equipment has been developed for this application (Gade et al., 2014). In 2011, Sirmacek et al. (Sirmacek et al., 2011) proposed a method based on automatic detection of heat leakage in thermal imager images, which can detect heat leakage and damage only by using images taken by vehicles moving around relevant buildings. If a fire occurs in a building, the thermal imager can greatly reduce or eliminate the dependence on personnel, quickly find out the location of suspected hidden dangers, shorten the troubleshooting time, be simple and efficient, and has become a powerful tool in the field of building structure monitoring.

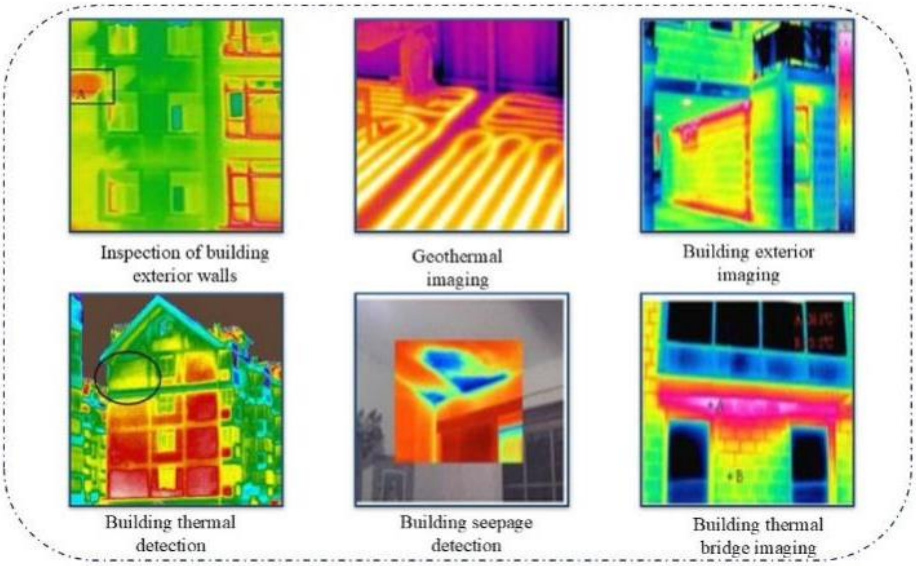

Figure 3. The application of thermal imaging renderings in the field of construction

2.3 3D Camera

It is worth noting that the machine vision system can be divided into 2D camera and 3D camera based on the imaging method, the traditional two-dimensional image sensor can only obtain the plane information of the object, and the emergence of 3D sensor technology enables the machine vision system to obtain the depth information of the object, greatly expanding the application field of machine vision. Common types of 3D cameras include: stereo vision cameras, structured light cameras, and LiDAR cameras. The comparison is shown below.

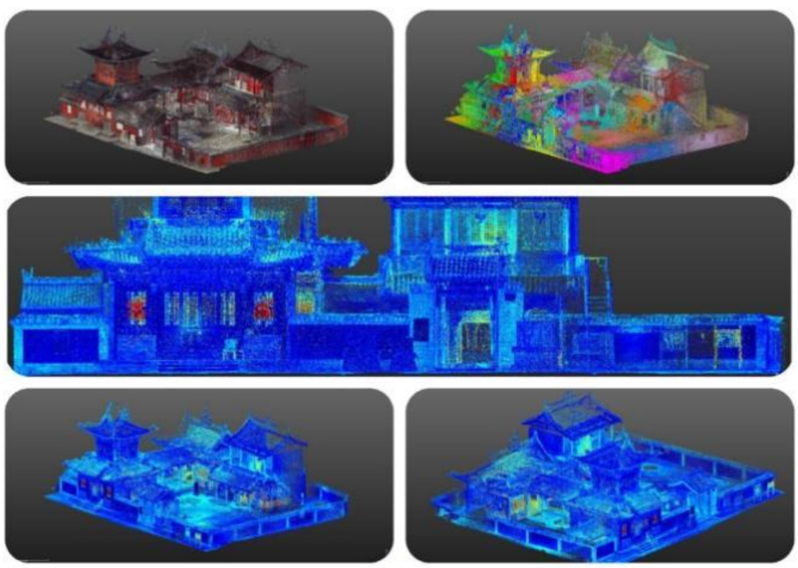

Figure 4. 3D depth image obtained from LiDAR sensor

Stereo vision is based on the principle of two eyes on an object, the greater the parallax, the closer the object is to the camera. By calculating the difference between the corresponding points in the image (parallax), the distance from the object to the camera can be deduced, thus generating three-dimensional information. Therefore, stereo vision is suitable for scenes with low equipment requirements and controllable lighting conditions (Laga et al., 2022).

A structured light camera is suitable for scanning objects with high precision and small volume. It works by projecting known patterns (usually stripes, grids, or dots) onto an object's surface, and then capturing the deformation of those patterns with a single camera. According to the projection pattern's deformation, the object's surface shape is calculated, and the depth information is obtained.

LiDAR cameras are suitable for high-precision 3D scanning over large areas by firing laser pulses and measuring their reflection time (similar to ToF), but unlike ToF, LiDAR typically uses a laser beam to scan the surrounding environment, generating dense point cloud data. The time it takes for these laser pulses to bounce back can calculate the exact position of each point, resulting in high-resolution three-dimensional data. LiDAR cameras are widely used in areas such as unmanned driving and geographic surveying. Figure. 4 shows a 3D depth image obtained from the LiDAR sensor (Zhao et al., 2019).

2.4 Trend

At this stage, although machine vision technology has made great progress, there are still many problems in the current development of machine vision, such as 1) the machine vision system is still insufficient to work in a variety of harsh environments, and the durability needs to be further improved. 2) The accuracy and data transmission speed of machine vision hardware can be further improved. 3) High-performance machine vision equipment is too large and energy consumption is too high. These key issues remain the focus of future machine vision hardware research.

3. Machine Vision Algorithm Development

Machine vision algorithms play a crucial role in modern technology and social life, which can not only imitate human visual systems, but also exceed human abilities in specific tasks. From the initial theoretical discussion to the current wide application, machine vision algorithms have experienced a long evolution process, and are moving towards a more intelligent and personalized direction.

3.1 Calibration

In machine vision, camera distortion can affect image authenticity, visual effects, and post-processing efficiency, especially in areas that require accurate representation of object shape, proportion, and detail. Camera calibration is the main method to solve camera distortion.

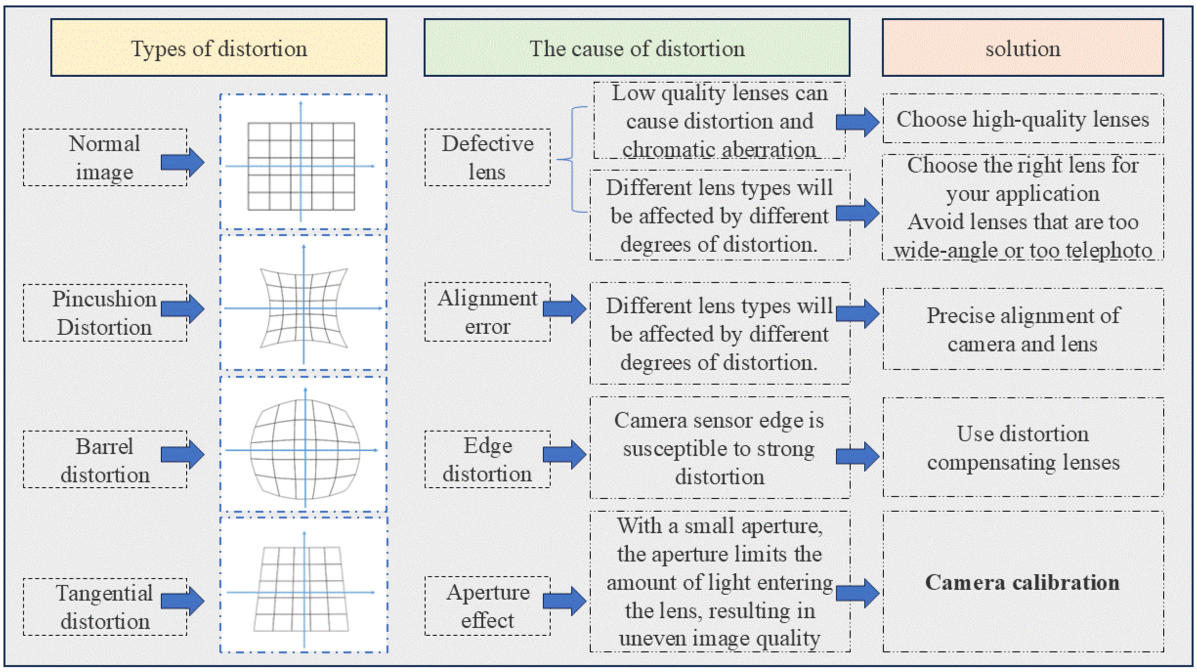

Figure 5. Types, causes, and solutions of camera distortion

The camera calibration work is to determine the internal and external parameters of the camera to establish the relationship between the three-dimensional world coordinates and the two-dimensional image coordinates to ensure the accurate measurement and positioning of the objects in the image. There are many methods for camera calibration. For example, Tsai et al. (Tsai et al., 1987) proposed a general two-step calibration method aimed at efficiently calculating the external position and orientation of the camera relative to the object's reference coordinate system, as well as effective focal length, radial lens distortion, and image scanning parameters. The most widely used is a flexible camera calibration technology proposed by Zhang (Zhang, 2000) in 2000, which is a method to obtain the internal and external parameters of the camera by analyzing the images of known geometric arrays at different angles. Figure. 5 summarizes the types, causes, and solutions of camera distortion.

3.2 Image Stitching

With the development of digital photography, people can easily take multiple images next to each other and hopefully synthesize them into a larger, more comprehensive image. The Image stitching algorithm mainly includes three key steps: image registration, image fusion, and Image stitching. Image stitching is an important research direction in the field of computer vision, and has great significance for the deformation and monitoring of structure panorama.

Image stitching refers to merging two or more images into one large image according to specific geometric transformation and alignment techniques, usually used to generate a panorama or expand the image field of view. Brown (Brown et al., 2007) uses the SIFT operator for automatic panoramic image stitching. This method is insensitive to the ordering, direction, proportion, and illumination of input images, and it is insensitive to noisy images that are not panoramic. More efficient methods such as multi-view geometry (Yu et al., 2024), RANSAC (Liu et al., 2022), and panoramic Image Stitching (Aziz et al., 2016) are available.

3.3 Stereo Vision

Binocular vision comes from human physiology: Humans have a pair of eyes and perceive the three-dimensional world through the parallax between the eyes. In computer vision, binocular vision simulates this physiological mechanism, obtaining different views of the same scene through two cameras (or two camera perspectives), and calculating the depth information of the scene by comparing different perspectives of the same scene in the two images.

With the insufficiency of traditional 2D images and the need for 3D images in practical engineering, binocular vision technology is also developed. Sidahmed et al. (Sidahmed et al., 1990) proposed a dual-camera calibration method for 3D machine vision metrology. Hosseinzadeh et al. (Hosseinzadeh et al., 2020) proposed a vision-based surveillance camera vibration monitoring method, which can restore the 3D movement of buildings by synchronizing vision-based measurement data of multiple floors.

3.4 Discussion

The improvement and success of machine vision algorithms have provided a great impetus for progress in the field of building structural health monitoring, and it is expected that these machine vision algorithms will gradually become popular in the field of building structural health in the future. Nevertheless, there is still room for further improvement of machine vision algorithms, mainly in three aspects: accuracy, multi-modal fusion, and real-time.

The improvement of the accuracy of target detection and segmentation is the focus of further improvement. For example, defects in building structures (such as cracks, corrosion, peeling, etc.) are often small and complex, and traditional machine vision algorithms may face the problem of insufficient detection accuracy, which can be combined with information from multiple perspectives, or combining depth cameras (such as LiDAR or Stereo Camera) with ordinary camera data.

It is very meaningful to promote the integration of visual information and environmental information. Relying on a single visual signal may not provide sufficient structural health information, especially in complex building structures. The resulting image will be affected by many factors (wind, light, heat, sound, time, and background changes), resulting in large errors. Therefore, image data can be combined with other sensor data (such as infrared thermal imaging, LiDAR, ultrasonic sensors, accelerometers, etc.) to improve the accuracy and reliability of structural health monitoring.

There are still challenges in the real-time performance and processing efficiency of machine vision structural monitoring systems. For example, building fire monitoring pays more attention to real-time, and strives to minimize fire hazards and eliminate fire hazards. Health monitoring of building structures often requires monitoring in large-scale buildings or over long periods, which requires machine vision systems to be able to process large amounts of data in real time. In this regard, computing tasks can be transferred from the cloud to edge devices (such as cameras, drones, smart sensors, etc.), using edge computing to accelerate image processing and defect analysis.

Although machine vision technology algorithms have made significant advances and are superior to human vision in many ways, there is still room for improvement. With the continuous progress of technology, these problems are expected to be gradually solved.

4. Survey of Applications

Machine vision technology was initially born as a computer data processing technology, but the combination of machine vision technology and other technologies has a wide range of important applications in the practical engineering of civil engineering and water conservancy engineering, mainly reflected in the following aspects: 1) crack detection; 2) Detection of seepage; 3) Detect fire.

4.1 Crack Detection

Now, the automatic monitoring of cracks is a hot research direction, which can bring the result of saving time and labor. As shown in Table 1, bridge safety accidents will cause a great blow to people's lives and property safety, and unsafe bridge structures are also a great hidden danger. People are used to regular health monitoring of Bridges, which often wastes too much time and money, and it is easy to miss the best period of crack repair. Moreover, as they are small and undetectable, cracks are easy to ignore in manual monitoring. Cardellicchio (Cardellicchio et al., 2023) studies that automatic identification of bridge cracks will greatly help bridge safety, and at the same time promote the ability to predict defects over time, which can more favorably ensure the safety of people's lives and property.

4.2 Detection of Seepage

Concrete structures often have a series of problems when facing surface damage such as water erosion. Seepage inspection is essential to delineate the damaged area and ensure the long-term safe operation of the dam. Therefore, Wang et al. (Wang et al., 2022) proposed a new convolutional neural network that can automatically identify dam surface seepage from heat maps collected by drones carrying thermal imaging cameras, providing a promising and cost-effective automatic dam surface seepage inspection method. Tan (Tani et al., 2024) used an image semantic segmentation method based on deep learning to automatically locate tunnel lining leakage. By studying the intelligent identification and surface damage detection technology of concrete damage, the concrete structure can be better monitored and maintained, and its service life and safety can be improved.

4.3 Fire Detection

Machine vision can achieve rapid and accurate identification of fires, avoiding the risk of forest fire damage. Abedi (Abedi et al., 2021) proposes a method that utilizes machine learning to identify vulnerable bridge fire hazards and is also able to pinpoint the bridge components that are vulnerable to fire and show confidence in their predictions. In future studies, we can further explore how to combine machine learning and deep learning technologies to develop more intelligent and efficient fire detection systems to deal with fire risks in different environments and scenarios.

5. Conclusion

This paper summarizes the latest research results of computer vision structural displacement monitoring technology in equipment development, algorithm progress and application scenarios from the perspective of structural health monitoring. In addition, the development process, classification, application scenarios, and future development trends of machine vision equipment and algorithms are introduced. The development status of machine vision for existing application scenarios including crack, seepage, and fire is introduced. Structural displacement monitoring based on computer vision still has a huge space for development. Improving accuracy and reducing errors caused by various factors are important research directions for the future. In addition, how to improve the efficiency and reliability of its application in structural health monitoring is still the focus of subsequent research.

References

[1] Aziz, M. U., Boulanger, P., & IEEE. (2016). Video-rate panorama for free-viewpoint TV. *In 13th Conference on Computer and Robot Vision (CRV)*, Victoria, Canada, 228–235. https://doi.org/10.1109/CRV.2016.46

[2] Abedi, M., & Naser, M. Z. (2021). Rapid, autonomous and intelligent machine learning approach to identify fire-vulnerable bridges. *Applied Soft Computing, 113*, 1514–1516. https://doi.org/10.1016/j.asoc.2021.107964

[3] Bai, Y., Zhang, B., Xu, N., Zhou, J., Shi, J., & Diao, Z. (2023). Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review. *Computers and Electronics in Agriculture, 61*(5), 673–681. https://doi.org/10.1016/j.compag.2023.107048

[4] Brown, M., & Lowe, D. G. (2007). Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision, 74*(1), 59–73. https://doi.org/10.1007/s11263-006-0002-3

[5] Cardellicchio, A., Ruggieri, S., Nettis, A., Reno, V., & Uva, G. (2023). Physical interpretation of machine learning-based recognition of defects for the risk management of existing bridge heritage. *Engineering Failure Analysis, 9*(2–3), 81–92. https://doi.org/10.1016/j.engfailanal.2023.106183

[6] Fossum, E. R., & Hondongwa, D. B. (2014). A review of the pinned photodiode for CCD and CMOS image sensors. *IEEE Journal of the Electron Devices Society, 2*(3), 33–43. https://doi.org/10.1109/JEDS.2014.2306412

[7] Gade, R., & Moeslund, T. B. (2014). Thermal cameras and applications: A survey. *Machine Vision and Applications, 25*(1), 245–262. https://doi.org/10.1007/s00138-013-0570-5

[8] Hosseinzadeh, A. Z., Tehrani, M. H., & Harvey, P. S., Jr. (2020). Modal identification of building structures using vision-based measurements from multiple interior surveillance cameras. *Engineering Structures, 99*(5), 752–757. https://doi.org/10.1016/j.engstruct.2020.109058

[9] Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep learning applications in medical image analysis. *IEEE Access, 6*, 9375–9389. https://doi.org/10.1109/ACCESS.2017.2788044

[10] Liu, Y. (2022). Improved feature point pair purification algorithm based on SIFT during endoscope image stitching. *Frontiers in Neurorobotics, 16*, 81–92. https://doi.org/10.3389/fnbot.2022.888627

[11] Laga, H., Jospin, L. V., Boussaid, F., & Bennamoun, M. (2022). A survey on deep learning techniques for stereo-based depth estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(4), 1738–1764. https://doi.org/10.1109/TPAMI.2021.3079303

[12] Spencer, B. F., Jr., Hoskere, V., & Narazaki, Y. (2019). Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering, 5*(2), 199–222. https://doi.org/10.1016/j.eng.2018.11.030

[13] Sirmacek, B., Hoegner, L., & Stilla, U. (2011). Detection of windows and doors from thermal images by grouping geometrical features. *2011 Proceedings of Joint Urban Remote Sensing Event (JURSE 2011)*, 133–136. https://doi.org/10.1109/JURSE.2011.5764752

[14] Si, Y., Rouse, J. P., & Hyde, C. J. (2020). Potential difference methods for measuring crack growth: A review. *International Journal of Fatigue, 28*, 59–63. https://doi.org/10.1016/j.ijfatigue.2019.105305

[15] Sidahmed, M. A., & Boraie, M. T. (1990). Dual camera calibration for 3-D machine vision metrology. *IEEE Transactions on Instrumentation and Measurement, 39*(3), 512–516. https://doi.org/10.1109/19.54860

[16] Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation, 3*(4), 323–344. https://doi.org/10.1109/JRA.1987.1087109

[17] Tan, Y., Li, X., Lai, J., & Ai, J. (2024). Real-time tunnel lining leakage image semantic segmentation via multiple attention mechanisms. *Measurement Science and Technology, 35*(7). https://doi.org/10.1088/1361-6501/ad174e

[18] Wulff, A. M., Hashida, T., Watanabe, K., & Takahashi, H. (1999). Attenuation behaviour of tuffaceous sandstone and granite during microfracturing. *Geophysical Journal International, 139*(2), 395–409. https://doi.org/10.1046/j.1365-246X.1999.00939.x

[19] Wang, Z. F., et al. (2022). Convolutional neural-network-based automatic dam-surface seepage defect identification from thermograms collected from UAV-mounted thermal imaging camera. *Construction and Building Materials, 323*(14), 59–63. https://doi.org/10.1016/j.conbuildmat.2022.126548

[20] Yu, Z., Li, L., Xie, J., Wang, C., Li, W., & Ning, X. (2024). Pedestrian 3D shape understanding for person re-identification via multi-view learning. *IEEE Transactions on Circuits and Systems for Video Technology, 34*(7), 5589–5602. https://doi.org/10.1109/TCSVT.2023.3327152

[21] Zhang, Z. Y. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(11), 1330–1334. https://doi.org/10.1109/34.888718

[22] Zhao, J. X., Xu, H., Liu, H. C., Wu, J. Q., Zheng, Y. C., & Wu, D. Y. (2019). Detection and tracking of

pedestrians and vehicles using roadside LiDAR sensors. *Transportation Research Part C: Emerging Technologies, 100*, 68–87. https://doi.org/10.1016/j.trc.2019.01.018

Copyrights