Enhancement of Sonar Detection in Karst Caves Through Advanced Target Location and Image Fusion Algorithms

Hongquan Lei[1], Diquan Li[1], Haidong Jiang[2*]

[1] School of Geosciences and Info-Physics, Central South University, Changsha 410083, China
[2] Institute of Resources and Environmental Engineering, Guizhou Institute of Technology, Guiyang 550003, China

Corresponding Author Email: 20170783@git.edu.cn

ABSTRACT

Over extended periods, nature's forces have sculpted and solidified unique cave formations. The scientific community's increasing interest in studying these karst caves has highlighted the significance of advanced detection methods. Sonar caving, a prevalent technique, still grapples with comprehensive detection within these intricate structures. This study delves into the amalgamation of target location and image fusion algorithms to bridge this gap. Target location elucidates an object's position, dimensions, and intrinsic characteristics within a designated temporal and spatial domain. Image fusion, a leading topic within image processing, leverages the consolidation of diverse data types and formats, thereby playing a pivotal role in image enhancement, compression, and recognition. From a deep learning perspective, neural network output values were calculated, facilitating an improvement in positioning accuracy. Further, through matrix decomposition and wavelet transforms, the fusion efficacy was scrutinized, aiming to broaden the sonar detection ambit. Comparative experimentation underscored the efficacy of integrating these two algorithms in sonar-based karst cave detection. It was observed that not only did the image clarity amplify, but there was also a notable 7.93% augmentation in positioning accuracy and a surge in detection speed. Cumulatively, a 7.64% boost in the overall efficiency of sonar detection in karst caves was achieved, underscoring the imperative nature of these technological advancements for cave surveillance.

1. INTRODUCTION

Sonar technology, having evolved over several centuries, was initially recognized for its military implications. As technological advancements progressed, its applications were expanded to include aviation, terrestrial, and marine domains, with reconnaissance operations becoming a standard procedure. In the realm of cave detection, an increasing emphasis has been placed on the proficient processing of acoustic inputs. However, traditional sonar instruments were observed to primarily capture sound characteristics from the cave's exterior within a limited range. In more intricate scenarios, the requirements for in-depth identification and localisation within caves were not met, leading to questions about optimizing sonar's role in cave exploration.

With socio-economic progression and the rise of scientific prowess, the significance of sonar detection technology was underscored. A multitude of scholars embarked on exhaustive studies in this field. Zhong et al. [1] incorporated a minute detector signal into the sonar signal processor. Through quantitative evaluations of the collected data, it was used for the surveying of 14 karst caves in railway tunnels, thereby fostering the growth of the western region and accelerating the erection of high-speed railways. In a distinct study by Mallios et al. [2], data was collected in an unstructured underwater cave environment using an autonomous underwater vehicle, with the intention of site-specific validation of authentic ground points. This method was found to bolster the

proficiency of underwater assessments and advance aquatic intervention mechanisms. Abu and Diamant [3], in their work, introduced an unsupervised, statistical-based algorithm tailored for the detection of submerged entities in synthetic aperture sonar imagery, ensuring a harmonious equilibrium between detection probability and false positive rates. A remarkable endeavour by Argotea et al. [4] integrated the ERT (Emergency Response Team) survey with sonar technology, unearthing compelling evidence of a cavern beneath the Moon Pyramid, which spurred dialogues regarding the pyramid's symbolic implications and the inception of the site's urban blueprint. Lu et al. [5] scrutinized background statistics derived from multiple nodes and, contingent on these statistics, ascertained suitable algorithms. It was discerned through computational simulations and evaluation of documented sonar data that centralized detection apparatuses exhibited heightened resilience in a plethora of diverse marine environments. Lastly, Yu et al. [6] harnessed seabed backscatter side-scan sonar to procure detailed seabed topographies, subsequently crafting high-definition seabed imagery by employing super-resolution techniques, ensuring the preciseness and efficacy of object identification. The gravitas of sonar, especially in cave exploration, cannot be understated, and rigorous scientific inquiry can form the cornerstone for ensuing endeavours in this sphere.

Target localization and image fusion algorithms have emerged as prominent research foci in recent times. A surge of academic interest in these domains has been observed, with

numerous scholars advancing this frontier of knowledge. In the realm of spine surgery, target localization methodologies were employed by Manbachi et al. [7]. Through rigorous experimentation, it was confirmed that this technology could autonomously annotate vertebrae in intraoperative radiographs. Such innovations were found to diminish the error margins in machine-independent inspections and bolstered active assistance, leading to augmented confidence during surgical procedures.

Mirzaei Hotkani et al. [8], in their work, harnessed broadband noise data from varied ship typologies to adeptly pinpoint underwater targets. The utilization of the MFP (Matched-Field Processing) algorithm, coupled with aquatic environmental model estimators, was observed to be pivotal in enhancing maritime navigational safety by optimizing the alignment between received signals. In a distinct investigation, Tomic et al. [9] reformulated the target localization dilemma into a generalized trust region subproblem framework. A novel robust estimator was subsequently introduced, circumventing the traditionally slow convergence rate challenges observed with legacy algorithms. As a result, the quandary of target localization in stringent indoor terrains, contingent on distance gauging, was addressed.

A novel image fusion algorithm was put forth by Huang et al. [10], marrying the nonlinear approximations of the contourlet transform with image region features. This synergy, which drew inspiration from the merits and demerits of the contourlet transform and multimodal medical imagery, was determined to elevate the visual appeal and quality of medical image fusion. Concurrently, enhanced image denoising outcomes were documented. Sun et al. [11] introduced a Laplacian pyramid-rooted region stitching approach to amalgamate multifocal images secured through microscopy. It was discerned that this strategy retained a more substantial portion of the original data compared to pixel-level counterparts and minimized the color aberrations in the resultant fused images. Anandhi and Valli [12] articulated a fusion algorithm within the domain of the non-subsampled contourlet transform, premised on synthetic aperture radar and panchromatic visuals. The mechanism was recognized for its adeptness at offering malleable multi-scale, multi-directional image scaling, ensuring maximal retention of edge and texture data.

While sonar holds an unparalleled position amongst acoustic detection modalities, advancements in knowledge and technological capabilities have rendered passive sonar detection less effective against contemporary silent submarines. Hence, refining sonar detection efficacy has become imperative. Addressing this lacuna, the present study endeavours to introduce a target localization algorithm in tandem with an image fusion procedure. The comprehensive processing of multi-sensor data is harnessed to amplify the precision of target discernment.

2. SONAR-BASED TARGET LOCATION AND IMAGE FUSION ALGORITHMS IN KARST CAVE DETECTION

(1) Underlying principles of karst cave detection via sonar

Sonar detection has perennially stood as a focal point of research. Its employment in karst cave exploration is recognized as a standard method. This technique is predominantly anchored in the identification of entities within cavities. Sound data are transmuted directly into electrical signals or alternative communication forms, aiming to furnish more lucid outcomes for operators. While this mode of detection has been shown to effectively map the cave's interior milieu and geological layout, it is not devoid of specific inaccuracies. A rigorous evaluation of its precision is mandated [13]. The intricate principles of karst cave detection via sonar are delineated in Figure 1.

As illustrated in Figure 1, sonar encompasses an array of acoustic sensors and mechanized instruments. Within this assembly, electrical signals, generated by the transmitter housed in the electronic cabinet, are relayed to the cavern post interfacing with the transducer. The strategic organization of these transducers constitutes the array, facilitating the mutual conversion of acoustic and electrical energies. When acoustic pulses navigate aquatic mediums, reflections arise upon target encounter. These reverberated sound waves are subsequently captured by the transducer, undergoing conversion into electrical impulses, which are then relayed to the electronic cabinet. Consequently, variations in echo waveforms and intensities across distinct frequency spectrums become visible on display panels.

In the context of karst cave detection using sonar, underground regions might encompass an amalgam of gaseous, liquid, and solid phases. Cave walls, perforated with orifices of varied dimensions and morphologies, exert differential impacts on acoustic signals. Hence, accurate interpretations necessitate the consideration of myriad external conditions and inherent challenges. The potential strengths and limitations of this technology in exploration pursuits will be analyzed. Furthermore, emphasis will be placed on elucidating the contributions of target localization algorithms and image fusion methodologies to the sonar-mediated identification of karst formations.
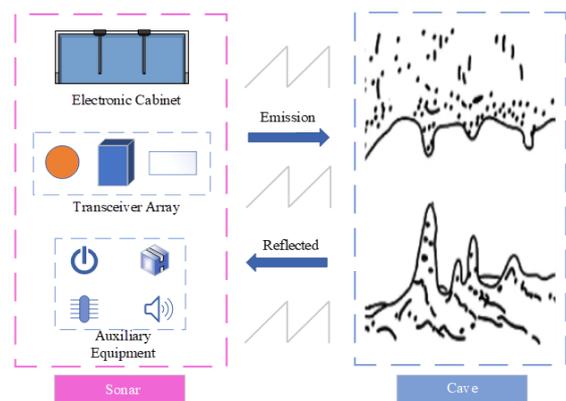

Figure 1. The principle of sonar detection in caves

(2) Underpinnings of target localization techniques

Target localization serves as a cornerstone within radar systems. This process entails the derivation of the discrepancy between measured and actual observed values through the formulation of a mathematical model. Among the myriad steps involved, identifying the optimal solution stands out as the most intricate. With the burgeoning advancements in modern electronic information technology coupled with satellite navigation, technical prowess in this domain is observed to be increasingly refined. Its utility, especially in multi-sensor information fusion and real-time tracking, has been underscored by promising prospects [14]. The algorithmic principles guiding target localization are depicted in Figure 2.
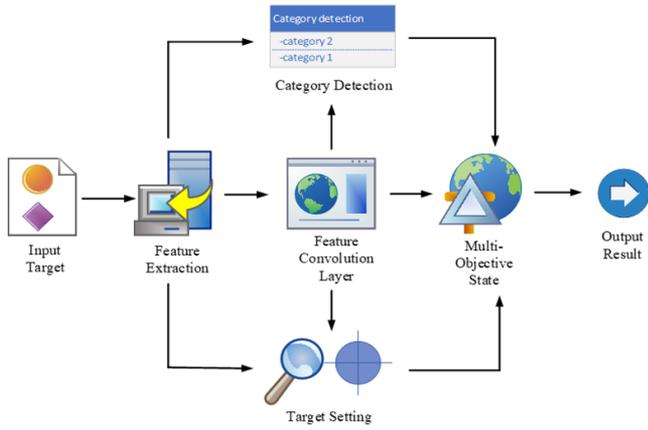
Figure 2. Algorithmic framework of target localization

Object localization is recognized as an adaptive learning algorithm predicated on fuzzy logic principles. Initially, upon input target perception, its location is promptly estimated by the sensor network. Subsequent outputs are deduced based on the prevailing state, thereby refining the prospective target domain. In the ensuing phase, visual characteristics are extracted from the procured image and supplementary convolutional layer features are garnered. A training set is constructed leveraging these features, where filtered pertinent data undergoes utilization for category detection and locational discernment. Ultimately, through a combination of image segmentation and feature extraction, classification outcomes are assessed and the real-time status of the target is ascertained, ensuring precision in target tracking and recognition. Owing to its efficacy in enhancing identification accuracy, the target localization algorithm has been found to be of significant merit in practical applications.

(3) Elucidation of the image fusion procedure

The progression of computer technology and the proliferation of internet usage have spurred heightened expectations concerning visual perception capabilities. Image fusion encompasses the amalgamation of data derived from disparate sources, typically gathered via a range of sensors. Once relayed for computational processing via networks, a coherent, precise, and intuitive aggregate assessment is rendered [15]. Generally initiated at the source, image processing assimilates salient environmental information, potentially mutable along its pathway, capitalizing on the interrelation among multi-source images to actualize target identification or detection. This technique is acknowledged for its superior real-time responsiveness and voluminous information capture. The intricate steps underpinning image fusion are delineated in Figure 3.

The fusion methodology can be segregated into four distinct stages. Initially, 'n' raw datasets or images are procured from external environments. Subsequent to this acquisition, unification and specialized treatment are conducted, aligning with data layer standards—a process colloquially termed as 'image registration'. Predominantly reliant on a reference image, the auxiliary images undergo intricate processing to synchronize temporally and spatially with the entirety of the image collection. This is followed by image preprocessing, a phase marked by novel representation of object interrelations within the backdrop and encapsulation of their intrinsic content. Concluding the sequence, the processed images are amalgamated to produce comprehensive visuals, which are

subsequently presented to end-users or diverse application frameworks. Grounded in multi-scale image recognition paradigms, image fusion methodologies have emerged as avant-garde avenues for data acquisition. Their recurrent incorporation within an array of intelligent systems underscores their adaptability and resilience in multifaceted environments.
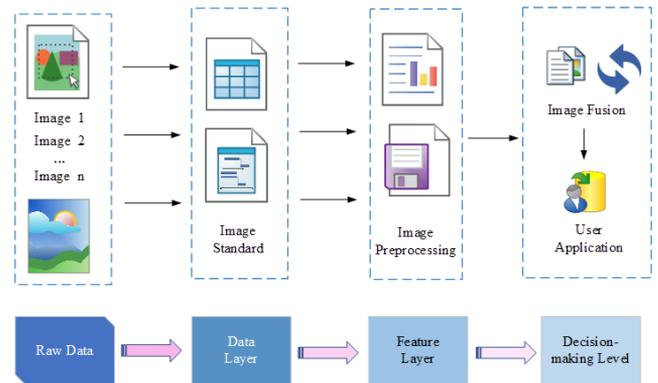

Figure 3. Schematic representation of the image fusion procedure

2.1 Formulation of the target localization algorithm and image fusion algorithm

1) Target Localization Algorithm

Object localization has been established as a pivotal axis of research within deep learning algorithms. By fractionating intricate issues into several rudimentary sub-processes, this adaptive tracking technique, founded on the least squares principle, is manifested. It has been observed to surmount certain limitations inherent to conventional PID (Proportional Integral Derivative) control [16]. Furthermore, dynamic parameter adjustments, tailored to specific prerequisites, can be undertaken to achieve the anticipated outcome.

1. Elucidation of the Least Squares Matrix Solution
Let the matrix representation of the function $f(n)$ be given as:

$$f_a(\mathrm{n}) = Na \tag{1}$$

Wherein N represents the matrix of dimension $i \times j$, with i denoting the number of samples and j indicating the number of sample features. a is identified as the feature vector.

Subsequently, the formulation of the loss function is represented as:

$$C(a) = \frac{1}{2}(Na - M)^q (Na - M) \tag{2}$$

On differentiating Eq. (2) concerning the eigenvector and equating to zero, the following is derived:

$$\frac{\partial}{\partial a}C(a) = N^q(Na - M) = 0 \tag{3}$$

Upon sequential arrangement, the expression becomes:

$$NqNa = NqM \tag{4}$$

Multiplication on both sides by the inverse of (NqN) yields:

$$a = (NqN) - 1N^q M \qquad (5)$$

From Eq. (5), the vector a can be computed. It has been discerned that sonar detection systems, leveraging the least square method, exhibit proficiency in discerning diverse hole types, thus bolstering classification accuracy.

2) Deep Neural Network

At the heart of deep learning lies object localization. It is posited that deep learning serves as the essential component of the deep neural network, recognized for its robust spatial reasoning capabilities and elevated prediction accuracy. Such attributes enable it to facilitate swift analyses of intricate challenges or unstructured data sets [17].

Within the perceptron of a neural network, the input layer's neurons are harnessed for their learning capacities during training sessions. Conversely, artificial neural network algorithms are employed by the output layer for data processing. A linear association between the output and the input is observed:

$$s = \sum_{a=1}^{i} p_a q_a + c \qquad (6)$$

where, p stands as the coefficient delineating the linear relationship, while c' acts as the bias coefficient.

Given the activation function $\lambda(s)$ and deeming the output value of both the hidden and output layers as e, the subsequent output from the hidden layer is defined as:

$$\begin{aligned} e_1 2 &= \lambda(s_1 2) \\ &= \lambda(p_{11}2q_1 + p_{12}2q_2 + p_{13}2q_3 + c_1 2) \end{aligned} \qquad (7)$$

Furthermore, the output layer's value is represented by:

$$\begin{aligned} e_1 3 &= \lambda(s_1 3) \\ &= \lambda(p_{11}3q_1 2 + p_{12}3q_2 2 + p_{13}3q_3 2 + c_3 3) \end{aligned} \qquad (8)$$

Upon analysis of Eqs. (7) and (8), a pervasive pattern is discerned:

$$e_m n = \lambda(s_m n) = \lambda(\sum_{k=1}^{} d p_{mk} n e_k n - 1 + c_m^n) \qquad (9)$$

When expressed in matrix notation, the nth layer's output is articulated as:

$$en = \lambda(sn) = \lambda(Anen - 1 + cn) \qquad (10)$$

where, A denotes a linear coefficient matrix.

Through this methodology, the accurate identification of surface features is efficiently achieved. For enhancement of target localization precision, diverse variants of the Kalman filter are routinely utilized.

3) Algorithm for image fusion

The algorithm for image fusion draws its foundation from multi-scale image recognition technology. In intricate environments, this approach is proficient in distinguishing between varying targets and background data. The transition from 2D to 3D processing is facilitated with heightened efficiency and speed. Such enhancement not only elevates

monitoring accuracy but also fortifies adaptability and resilience in multifaceted scenes [18].

a. Local Standard Deviation

To some degree, the local standard deviation mirrors the relative stability of a region. This deviation is pivotal in gauging the trajectory of alterations in image contrast. Furthermore, the span of variation serves as a crucial yardstick in assessing the modulating patterns of gray values.

For an observational window with dimensions $a \times a$, the equations for its standard deviation are articulated as:

$$\partial = \sqrt{\frac{1}{a^2} de(e_{ij} - \bar{e})} \qquad (11)$$

$$de = \sum_{i=m-\frac{a-1}{2}}^{m+\frac{a-1}{2}} \sum_{j=n-\frac{a-1}{2}}^{n+\frac{a-1}{2}} \qquad (12)$$

where, e_{ij} is indicative of the pixel value within the given vicinity, juxtaposed against the mean value derived from all neighboring pixels.

b. Non-Negative Matrix Factorization

Non-negative Matrix FactorizationIn recent times, the introduction of non-negative matrix factorization has been noted as a contemporary method. This technique has not only showcased its aptitude to address a multitude of challenges that elude traditional algorithms but also boasts attributes such as simplicity, efficacy, and ease of programming. The utility of this method in computing decoupling and stabilizing nonlinear systems has been extensively recognized [19].

Given a matrix C, composed entirely of non-negative components, it can be represented by the multiplication of two distinct matrices as:

$$C_{a \times b} = A_{a \times r} B_{r \times b} \qquad (13)$$

where, matrix C exhibits dimensions of $a \times b$, while matrix A has dimensions of $a \times r$, and B holds the dimensions $r \times b$. The factor r delineates the dimensionality pertinent to the non-negative matrix factorization.

The range of R is constrained by:

$$r < \frac{ab}{a+b} \qquad (14)$$

At the instance when $r = 0$, the feature base matrix, A, becomes synonymous with all the image attributes.

Eq. (14) can be depicted in the form of a vector product:

$$C_n = \sum_{m=1}^{r} A_m B_{mn} \qquad (15)$$

where, C_n signifies the nth column of matrix C, while A_m characterizes the mth column of matrix A.

From the inference drawn from Eq. (15), each individual column of matrix C can be mirrored by the pertinent components found in matrices A and B. Consequently, matrix A is perceived as the foundational matrix, with matrix B taking on the role of the coefficient matrix. In scenarios where A embodies the intrinsic image characteristics, the following

relationship emerges:

$$C \approx AB \qquad (16)$$

c. Wavelet Transform

The wavelet transform has been observed to effectively distil detailed information from images. Particularly in instances marked by substantial noise, wavelet decomposition proves superior in disentangling texture and colour attributes, consequently achieving an augmented compression ratio. Such capabilities underscore its pivotal role in image enhancement [20].

The continuous function characterizing the wavelet transform is delineated by:

$$G(i,j) = \langle g, \lambda_{i,j} \rangle = \frac{1}{\sqrt{i}} \int_{-\infty}^{\infty} g(e) \lambda\left(\frac{e-j}{i}\right) ce \qquad (17)$$

$$g(e) = \frac{1}{D_\lambda} \int_{-\infty}^{\infty} G(i,j) \lambda_{i,j}(e) \frac{cicj}{i^2} \qquad (18)$$

In this context, $\lambda_{i,j}(e)$ epitomizes the shift and dilation of the foundational basis. The factor i, acting as the dilation factor, prescribes the support length of the wavelet function, ensuring i remains greater than 0. On the other hand, 'b' stands as the shift factor, concurrently serving as the wavelet's temporal parameter.

Upon setting constraints wherein i_0 exceeds 1 and j_0 is greater than 0, the expression for the discrete wavelet transform emerges as:

$$\lambda_{x,y}(e) = i_0^{-\frac{x}{2}} \lambda(i_0^{-x} e - y j_0) \qquad (19)$$

where, both x and y are integral in nature.

In the realm of image processing, the unidimensional wavelet transform is extrapolated to a bidimensional transformation space:

$$G(i, j_m, j_n) = \iint g(m,n) \lambda_{i, j_m, j_n}(m,n) cmcn \qquad (20)$$

Within this equation, m and n denote the translation of the image across their specific dimensions.

Subsequently, the inverse operation associated with the two-dimensional wavelet transform is conceptualized as:

$$\iint g(m,n)$$
$$= \frac{1}{D_\lambda} \iiint G(i, j_m, j_n) \lambda_{i, j_m, j_n}(m,n) cj_m cj_n \frac{ci}{i^3} \qquad (21)$$

3. COMPARATIVE ANALYSIS OF TARGET LOCALIZATION AND IMAGE FUSION ALGORITHMS IN SONAR DETECTION WITHIN KARST CAVES

3.1 Experimental procedure

To assess the efficacy of the target localization and image fusion algorithms in sonar detection within karst caves, an experiment was designed involving 20 cave explorers. Participants were arbitrarily divided into two equal groups: Group A and Group B. For the detection task, Group A was furnished with the traditional sonar detection system, whereas Group B utilized a system augmented by the target localization and image fusion algorithms. Metrics of comparison included accuracy, detection speed, and clarity of the resultant images.

Subsequently, participants were requested to evaluate the detection systems based on their experiences. Ratings were given on a scale ranging from 1 to 5, with 5 being the most favourable. Feedback was assimilated through questionnaires, the results of which were meticulously collated. Observations and recordings pertaining to the experimental data were meticulously documented.

3.2 Analysis of experimental data

1) Accuracy Assessment

Upon the incorporation of the target positioning algorithm, a comparative analysis was undertaken with both groups, Group A and Group B, situated in an identical cave detection environment. The comparative results are illustrated in Figure 4.
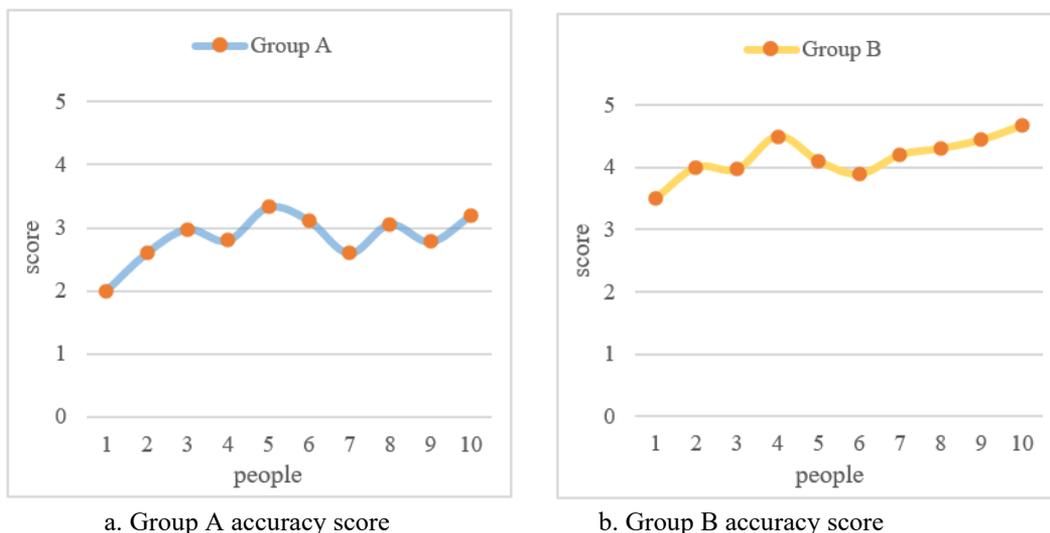

a. Group A accuracy score   b. Group B accuracy score

Figure 4. Comparative analysis of accuracy scores for both groups

a. Group A image resolution size
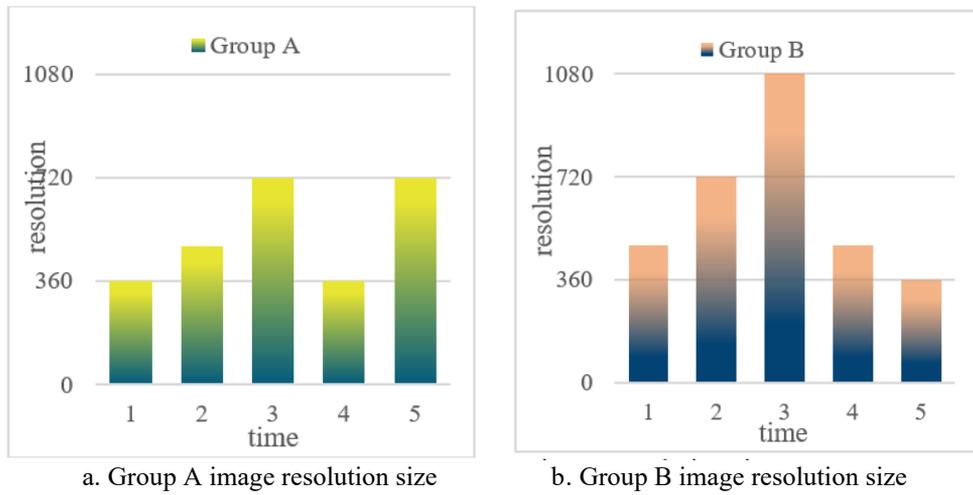b. Group B image resolution size

Figure 5. Resolution comparison between the two groups' systems

In Figure 4a, the detection accuracy scores garnered by Group A are depicted. Conversely, Figure 4b delineates the precision ratings of Group B. An evident disparity in scores is observable with Group B exhibiting superior accuracy scores compared to Group A. Scores for Group A fluctuated significantly within a range of 2-3.5 points, placing them in a lower middle tier. Such marked variability in scores suggests a pronounced divergence in perceptions concerning the system's accuracy within this group. On the other hand, Group B scores manifested a slight yet consistent ascendancy, underpinned by minimal variations. This conveys a consensus among Group B participants, emphasizing their confidence in the enhanced accuracy of their detection technology. A juxtaposition of the data sets from the two groups infers that, under congruent environmental constraints, the incorporation of the target positioning algorithm potentially augments the accuracy of sonar detection within karst caves, culminating in a more precise localization of targets.

2) Resolution and Clarity Assessment

Five consecutive detection sessions were undertaken by both groups to procure detected imagery. It was established that the display screen size remained consistent across both groups. A comparison of the resolution between the original system and the imagery posts the image fusion algorithm

implementation revealed results depicted in Figure 5.

Figure 5a illustrates the resolution metrics of Group A's images after the five successive detections. In contrast, Figure 5b delineates the resolution metrics for Group B's images. From the data, it was observed that the peak resolution achieved by Group A across the five detections stood at 720P, with the imagery quality being relatively high-definition and the lowest resolution recorded at 270P. Despite this, clarity remained in the images even at this reduced resolution. Conversely, the images detected by Group B exhibited a maximum resolution of 1080P, characterized by enhanced vibrancy and ultra-clear quality. Subsequent calculations established the average resolution for Group A at 528P, while Group B exhibited an average resolution of 624P. Given the fixed screen size, it can be inferred that higher resolutions result in enhanced image clarity, indicating that the sonar detection system, post the image fusion algorithm's integration, provides superior definition.

3) Assessment of Detection Speed

To evaluate the efficiency of each detection system, each participant was instructed to undertake a complete cave detection session, with the time duration from initiation to conclusion being recorded. Figure 6 offers a comparative visual representation of this data.
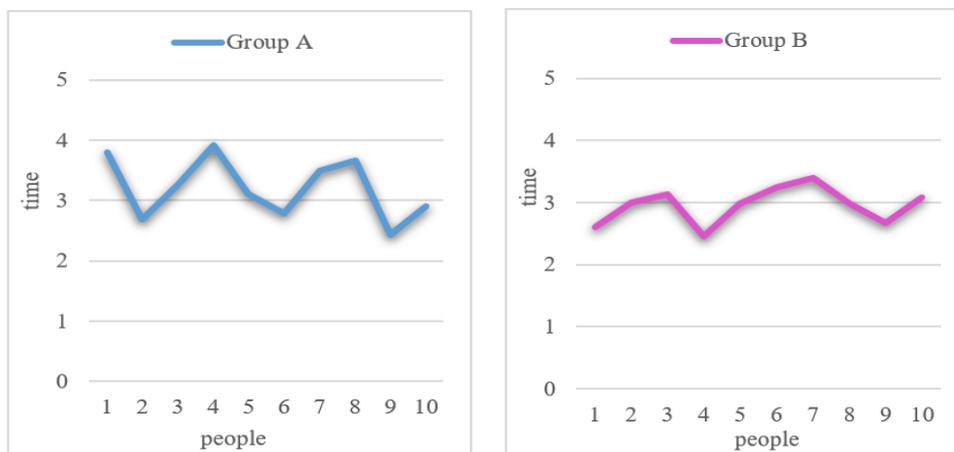
a. Time required for Group A to complete a test experiment   b. Time required for Group B to complete a test experiment

Figure 6. Detection speed comparison of the two groups of systems

Figure 6a provides a temporal account of Group A's detection duration, while Figure 6b delineates that of Group B. It was observed from the figure that the temporal data points for Group B generated a notably gentler curve, with an approximate uniform detection duration of 3 minutes for all members. Conversely, for Group A, time durations spanned between 2.5 to 4 minutes, exhibiting significant variations, indicative of disparities among the Group A participants. Based on the compiled data, it can be inferred that the sonar cave detection system, when integrated with the target positioning and image fusion algorithms, displayed enhanced detection speed and efficiency.

The subsequent analytical procedure entailed a juxtaposition of the original sonar detection process for karst caves against the system post the application of the target positioning and image fusion algorithms. For clarity in representation, resolutions were normalized within a range of 10, as represented in Table 1 and Figure 7.
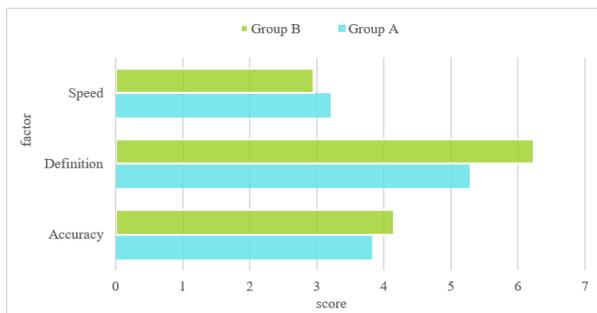

Figure 7. Comparative analysis of detection efficacies of both groups

Table 1. Comparative evaluation of detection metrics

	Group A	Group B
Accuracy	3.827	4.157
Definition	5.28	6.24
Speed	3.204	2.954

From the data presented in Figure 7, it was discerned that Group A's weighted average across accuracy, definition, and detection speed stood at 4.11, while that of Group B was found to be 4.45. It was noted that Group A operated with the original sonar detection system, whereas Group B utilized the enhanced system integrated with target localization and image fusion algorithms. Subsequent calculations revealed that the overall detection efficacy of Group B surpassed that of Group A by 7.64%. Additionally, improvements were seen in accuracy (by 7.93%) and detection speed (by 7.81%). A pronounced enhancement was also observed in image sharpness.

Through this research, the application of the target localization algorithm in the realm of sonar detection was empirically validated. This algorithm, as deduced from the experiment, offers invaluable assistance to personnel, aiding in the accurate determination of underwater target positions and their movement trajectories, particularly in intricate environments. This not only enhances the precision in identifying cave depths and entrance orientations but also substantially reduces risks associated with blind cave exploration. Moreover, the image fusion technique, by augmenting image processing, substantially enhances image quality, a pivotal factor for increasing image retrieval efficiency. The integration of these computational tools ensures the sonar sensor achieves a more nuanced and accurate analysis of cave structures.

3.3 Discussion

From the accumulated data and the employed methods for sonar detection of karst caves, along with the specific findings of this study, several key observations and outcomes were deduced:

(1) A comprehensive review of the extant literature was conducted, intertwining it with both contemporary empirical findings and theoretical studies. Through this analysis, the present developmental trajectory of sonar systems was discerned. Its pivotal significance, as well as the inherent challenges in the domain of karst engineering exploration, was illuminated. Notably, the implications of the target positioning algorithm and image fusion algorithm in the realm of sonar detection of karst caves were accentuated.

(2) By pivoting on target positioning and image fusion, the feature vector was computed using the least square method, facilitating the differentiation of varied karst cave types. Viewing this through the lens of deep learning, a surge in positioning accuracy was perceived. The computation of the local standard deviation depicted the evolving trends in image contrast. Furthermore, wavelet transformation was employed to segregate features, which in turn, expedited the extraction of actionable information, thereby amplifying the image's overall efficacy.

(3) For a more empirical understanding, 20 adept karst cave detectors were enlisted for an investigative study, wherein comparative experiments were orchestrated. From this initiative, it was observed that sonar detection, post the incorporation of the target positioning and image fusion algorithms, often yielded more streamlined results. Such detections were marked by heightened accuracy and reliability, leading to images that were both sharper in clarity and more intuitive in interpretation.

4. CONCLUSION

From the accumulated data and the employed methods for sonar detection of karst caves, along with the specific findings of this study, several key observations and outcomes were deduced:

(1) A comprehensive review of the extant literature was conducted, intertwining it with both contemporary empirical findings and theoretical studies. Through this analysis, the present developmental trajectory of sonar systems was discerned. Its pivotal significance, as well as the inherent challenges in the domain of karst engineering exploration, was illuminated. Notably, the implications of the target positioning algorithm and image fusion algorithm in the realm of sonar detection of karst caves were accentuated.

(2) By pivoting on target positioning and image fusion, the feature vector was computed using the least square method, facilitating the differentiation of varied karst cave types. Viewing this through the lens of deep learning, a surge in positioning accuracy was perceived. The computation of the local standard deviation depicted the evolving trends in image contrast. Furthermore, wavelet transformation was employed to segregate features, which in turn, expedited the extraction of actionable information, thereby amplifying the image's

overall efficacy.

(3) For a more empirical understanding, 20 adept karst cave detectors were enlisted for an investigative study, wherein comparative experiments were orchestrated. From this initiative, it was observed that sonar detection, post the incorporation of the target positioning and image fusion algorithms, often yielded more streamlined results. Such detections were marked by heightened accuracy and reliability, leading to images that were both sharper in clarity and more intuitive in interpretation.

REFERENCES

[1] Zhong, S., Wang, R., Wang, Z. (2017). Surveying Karst Caves Under 14 Tunnels Bottom on Gui-Guang High-Speed Railway. In Technology and Application of Environmental and Engineering Geophysics: Selected Papers of the 7th International Conference on Environmental and Engineering Geophysics, ICEEG-Beijing 2016, pp. 279-283. https://doi.org/10.1007/978-981-10-3244-8_33

[2] Mallios, A., Vidal, E., Campos, R., Carreras, M. (2017). Underwater caves sonar data set. The International Journal of Robotics Research, 36(12): 1247-1251. https://doi.org/10.1177/0278364917732838

[3] Abu, A., Diamant, R. (2019). A statistically-based method for the detection of underwater objects in sonar imagery. IEEE Sensors Journal, 19(16): 6858-6871. https://doi.org/10.1109/JSEN.2019.2912325

[4] Argote, D.L., Tejero-Andrade, A., Cárdenas-Soto, M., et al. (2020). Designing the underworld in Teotihuacan: Cave detection beneath the moon pyramid by ERT and ANT surveys. Journal of Archaeological Science, 118: 105141. https://doi.org/10.1016/j.jas.2020.105141

[5] Lu, S.P., Ding, F., Li, R.W. (2021). Robust centralized CFAR detection for multistatic sonar systems. Chinese Journal of Electronics, 30(2): 322-330. https://doi.org/10.1049/cje.2021.02.003

[6] Yu, F., Zhu, Y., Wang, Q., et al. (2019). Segmentation of side scan sonar images on AUV. In 2019 IEEE Underwater Technology (UT), Kaohsiung, Taiwan, pp. 1-4. https://doi.org/10.1109/UT.2019.8734433

[7] Manbachi, A., De Silva, T., Uneri, A., et al. (2018). Clinical translation of the levelcheck decision support algorithm for target localization in spine surgery. Annals of Biomedical Engineering, 46: 1548-1557. https://doi.org/10.1007/s10439-018-2099-2

[8] Mirzaei Hotkani, M., Bousquet, J.F., Seyedin, S.A., Martin, B., Malekshahi, E. (2021). Underwater target localization using opportunistic ship noise recorded on a compact hydrophone array. Acoustics, 3(4): 611-629. https://doi.org/10.3390/acoustics3040039

[9] Tomic, S., Beko, M., Dinis, R., Montezuma, P. (2017). A robust bisection-based estimator for TOA-based target localization in NLOS environments. IEEE Communications Letters, 21(11): 2488-2491. https://doi.org/10.1109/LCOMM.2017.2737985

[10] Huang, H., Feng, X.A., Jiang, J. (2017). Medical image fusion algorithm based on nonlinear approximation of contourlet transform and regional features. Journal of Electrical and Computer Engineering, 2017. https://doi.org/10.1155/2017/6807473

[11] Sun, J., Han, Q., Kou, L., Zhang, L., Zhang, K., Jin, Z. (2018). Multi-focus image fusion algorithm based on Laplacian pyramids. Journal of the Optical Society of America A, 35(3): 480-490. https://doi.org/10.1364/JOSAA.35.000480

[12] Anandhi, D., Valli, S. (2018). An algorithm for multi-sensor image fusion using maximum a posteriori and nonsubsampled contourlet transform. Computers & Electrical Engineering, 65: 139-152. https://doi.org/10.1016/j.compeleceng.2017.04.002

[13] Ma, Y., Mao, Z.Y., Qin, J., Ding, W.J., Xiao, Y.J., Meng, X.Y., Luo, R. (2019). A quick deployment method for sonar buoy detection under the overview situation of underwater cluster targets. IEEE Access, 8: 11-25. https://doi.org/10.1109/ACCESS.2019.2961555

[14] Mavrommati, A., Tzorakoleftherakis, E., Abraham, I., Murphey, T.D. (2017). Real-time area coverage and target localization using receding-horizon ergodic exploration. IEEE Transactions on Robotics, 34(1): 62-80. https://doi.org/10.1109/TRO.2017.2766265

[15] Jiao, L., Deng, J., Zhang, L., Lin, Q. (2021). Multi-focus image fusion algorithm based on grey relation of similarity in contourlet domain. In Advanced Information Networking and Applications: Proceedings of the 35th International Conference on Advanced Information Networking and Applications (AINA-2021), Toronto, Canada, 3: 280-288. https://doi.org/10.1007/978-3-030-75078-7

[16] Caspermeyer, J. (2017). Blind as bats: bat echolocation study reveals key evolutionary trade-offs with other senses. Molecular Biology and Evolution, 34(1): 241-242. https://doi.org/10.1093/molbev/msw257

[17] Qin, S., Zhang, Y.D., Amin, M.G., Gini, F. (2016). Frequency diverse coprime arrays with coprime frequency offsets for multitarget localization. IEEE Journal of Selected Topics in Signal Processing, 11(2): 321-335. 10.1109/JSTSP.2016.2627184

[18] Xiao, D., Wang, L., Xiang, T., Wang, Y. (2017). Multi-focus image fusion and robust encryption algorithm based on compressive sensing. Optics & Laser Technology, 91: 212-225. https://doi.org/10.1016/j.optlastec.2016.12.024

[19] Zuo, J., Zhao, W., Chen, L., et al. (2022). Multi-focus image fusion algorithm based on random features embedding and ensemble learning. Optics Express, 30(5): 8234-8247. https://doi.org/10.1364/OE.452081

[20] Argoul, F., Arneodo, A., Elezgaray, J., Grasseau, G., Murenzi, R. (1989). Wavelet transform of fractal aggregates. Physics Letters A, 135(6-7): 327-336. https://doi.org/10.1016/0375-9601(89)90003-0