Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 6, pp. 4720–4728, DOI 10.12694/scpe.v25i6.3287

DIGITAL PROTECTION AND INHERITANCE PATH OF INTANGIBLE CULTURAL HERITAGE BASED ON IMAGE PROCESSING ALGORITHM

JINGXUAN ZHAO*

Abstract. This research paper introduces a novel approach in the realm of digital preservation and inheritance of Intangible Cultural Heritage (ICH) through a Customized 3D Convolutional Neural Network (CNN). The core of this study lies in the development of an advanced image processing algorithm tailored to accurately recognize, categorize, and archive diverse forms of ICH, which include traditional performances, ceremonies, oral traditions, and crafts. Utilizing a volumetric 3D CNN, this paper demonstrates how complex ICH elements can be effectively captured and analyzed, overcoming the limitations of traditional 2D image processing methods. The network is trained on a comprehensive dataset of ICH imagery, ensuring sensitivity to the subtle nuances and dynamic nature of these cultural expressions. This paper highlights the algorithm's capability in not only safeguarding the visual aspects of ICH but also in providing an interactive, digital medium for education and cultural dissemination. The proposed method shows significant promise in aiding the efforts of cultural preservationists and educators, offering a technologically advanced pathway for the protection and inheritance of the world's rich, yet vulnerable, cultural heritage. This study sets a precedent in the interdisciplinary field of cultural heritage conservation, digital technology, and artificial intelligence, providing a scalable and effective solution for global ICH preservation initiatives.

Key words: Intangible cultural heritage, image recognition, digital protection, deep learning, CNN.

1. Introduction. In an era where the fabric of cultural diversity is under constant threat from the forces of globalization and cultural homogenization, the preservation of Intangible Cultural Heritage (ICH) emerges as a paramount concern. ICH, a term encompassing a broad spectrum of traditions, including but not limited to traditional performances, ceremonies, oral traditions, and artisanal crafts, represents the living expressions and knowledge passed down through generations [13, 4, 6]. Unlike tangible heritage, ICH is fluid, often existing in the collective memory and practices of communities. Its preservation is not just about safeguarding cultural artifacts; it is about maintaining the vibrancy and continuity of cultural identities in a rapidly changing world. However, the transient and dynamic nature of ICH presents unique challenges. Traditional methods of documentation and archiving are often inadequate in capturing the essence and intricacies of these cultural expressions, highlighting a pressing need for innovative approaches in the field of cultural preservation [8].

The advent of digital technology, particularly in the domain of artificial intelligence and image processing, has opened new horizons for addressing these challenges [14, 16]. Among these technological advancements, the evolution of 3D Convolutional Neural Networks (CNNs) stands out for their revolutionary capabilities in image analysis [13]. While traditional 2D image processing methods struggle to capture the depth and complexity inherent in many forms of ICH, 3D CNNs excel in handling volumetric data, offering a more nuanced and comprehensive analysis [4]. This ability to process and interpret complex visual data with remarkable depth and accuracy presents an unprecedented opportunity in the realm of ICH preservation. By leveraging these advanced technologies, there is potential to not only document but also to breathe new life into these cultural treasures, ensuring they are not lost to time [18, 22].

This research paper taps into this potential, introducing a novel approach that utilizes a Customized 3D CNN tailored specifically for the digital preservation and inheritance of ICH. This approach marks a significant departure from conventional methodologies, addressing the unique challenges posed by the diverse and nuanced nature of ICH [9, 2, 3]. The development of this advanced image processing algorithm is a response to the critical need for tools that can accurately recognize, categorize, and archive the various forms of ICH. This customization is key to the project's success, as it ensures the algorithm is finely tuned to the subtleties and dynamic qualities of different cultural expressions [21, 19]. Whether it is capturing the fluidity of a dance performance, the

*Shanxi college of applied science and technology Shanxi Taiyuan, 030062, China (jingxuanzhar1@outlook.com)

intricate patterns of traditional crafts, or the subtle nuances of oral storytelling, this customized approach promises a level of fidelity and depth in digital preservation that was previously unattainable[20, 17, 5].

The implications of this research are far-reaching and transformative. By harnessing the power of 3D CNNs, this study not only contributes to the safeguarding of cultural heritage but also opens up new avenues for its dissemination and appreciation. The digitalization of ICH through this method does not merely result in a static archive; it creates an interactive, dynamic medium that can be used for education and cultural exchange. It offers a way to bridge the gap between generations and geographies, making these rich cultural expressions accessible to a global audience. Furthermore, the scalability and effectiveness of this solution present a valuable tool for cultural preservationists, educators, and policymakers worldwide. In setting a precedent in the interdisciplinary field of cultural heritage conservation, digital technology, and artificial intelligence, this study paves the way for a new era in global ICH preservation efforts, ensuring that these vital cultural expressions continue to thrive and inspire future generations.

The main contribution of the paper as follows:

1. Proposed a novel approach of 3D CCNN based image recognition technique for the digital protection and inheritance path of ICH.
2. This proposed involves the effective customized 3D CNN to obtain the effective results in image recognition techniques.
3. The efficacy of the proposed is demonstrated with valid experiments.

2. Related Work.

2.1. Image Recognition techniques. The paper [10] explores the evolution of image recognition techniques, highlighting the transition from handcrafted features combined with machine learning methods to the superior performance of deep learning-based approaches post-2010. It emphasizes the advancements brought by deep learning in general object recognition competitions, and specifically addresses its application in autonomous driving and the latest trends in deep learning. The survey paper [11] provides a comprehensive analysis of deep learning's impact on image processing, discussing its successes and challenges. It delves into the complexities of deeper network structures and class imbalances in training data. The paper introduces four series of deep learning models and emphasizes the importance of understanding the relationship between deep learning and image processing tasks for future innovations and applications. The paper [12] reviews the application of deep learning in image recognition, outlining its significance in advancing computer vision and AI. It compares three main deep learning models - CNNs, RNNs, and GANs - and discusses their applications in various image recognition fields like face recognition and medical imaging, highlighting future trends like video image recognition and theoretical model enhancements. Focusing on cultural heritage, the paper [1] discusses the application of image recognition in enhancing the tourist experience at archaeological sites. It emphasizes the use of image recognition for content discovery in both indoor and outdoor settings, improving engagement through personalized content and real-time interaction, thus addressing challenges in heritage presentation. The paper [15] examines challenges in machine learning, particularly the scarcity of training data and class imbalance. It analyzes various data augmentation techniques, including classical transformations and advanced methods like Style Transfer and GANs, applied to medical case studies. The paper validates a new data augmentation method for enhancing training efficiency in image classification tasks. The paper [7] compares deep learning with traditional machine learning methods, outlining its development and network structures. It focuses on deep learning's application in image recognition and classification, discussing challenges and solutions. The paper concludes with a summary and future outlook on deep learning's role in image recognition and classification within AI.

Research challenges.

1. How to compile a comprehensive and representative dataset of ICH imagery that captures the wide variety of cultural expressions across different communities and regions, considering the scarcity of digital records for certain traditions and the potential biases in data selection.
2. Developing a 3D CNN architecture that can effectively process and analyze the multidimensional aspects of ICH, adapting to its dynamic and nuanced nature. This includes identifying the most suitable layers, activation functions, and other network parameters tailored to the complexity of ICH.

3. Ensuring that the digital preservation process respects and upholds the cultural integrity and ownership rights of communities, addressing the ethical implications of using AI in cultural heritage contexts.
4. Creating an engaging and interactive digital platform that facilitates meaningful connections between users and ICH content, encouraging learning and cultural

Research Questions.

1. How can a 3D CNN be effectively customized to recognize and categorize the diverse forms of ICH, taking into account the subtleties and dynamism inherent in cultural expressions?
2. What strategies can be employed to gather a comprehensive and diverse dataset of ICH imagery that is representative of global cultural expressions, including less-documented or endangered traditions?
3. In what ways can digital preservation methods incorporate cultural sensitivity and ethical considerations to respect the cultural rights and ownership of the communities whose heritage is being digitized?
4. How can the proposed digital preservation system facilitate interactive learning and user engagement with ICH content, thereby enhancing educational outcomes and cultural dissemination?
5. What are the main technological challenges in implementing a 3D CNN for ICH preservation, and what solutions can be developed to overcome these challenges?

3. Methodology.

3.1. Proposed Overview. The methodology for the proposed 3D Convolutional Neural Network (CNN) in image recognition begins with data collection. Initially, data collection involves gathering a comprehensive and diverse set of images, specifically targeting the intended application, such as ethnic clothing recognition. This step ensures the dataset is representative of various styles, patterns, and colors, vital for a robust model training. Care is taken to include images with varying angles and lighting conditions to mimic real-world scenarios. In the preprocessing phase, these images undergo normalization and augmentation. Normalization adjusts the images to a standard scale, enhancing the model's ability to process them efficiently. Image augmentation, such as rotating, scaling, and flipping, artificially expands the dataset, helping the model become more resilient to variations in new, unseen data. Feature extraction is at the heart of the 3D CNN's methodology. The network, with its convolutional layers, extracts and learns complex spatial hierarchies of features from the images. The depth of these layers captures not just the superficial characteristics but also the intricate details, which is crucial for high-accuracy recognition tasks. Finally, performance evaluation involves using metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into the model's effectiveness in real-world scenarios. The model is tested on a separate validation set to ensure it generalizes well to new data, not just the data it was trained on. This comprehensive evaluation helps in fine-tuning the model, ensuring it achieves the desired level of performance for practical applications. The proposed architecture is demonstrated in Figure 3.1.

3.2. Proposed 3D CNN based image recognition. To enhance the construction of a digital ethnic clothing library and facilitate the understanding of ethnic clothing culture through advanced image processing techniques, we propose the adaptation of traditional image feature extraction methodologies to a 3D CNN framework. This approach significantly augments the capability to analyze and archive ethnic clothing, capturing both their aesthetic and cultural essence. In the realm of computer vision, image feature extraction is pivotal. Traditionally, this involves identifying and analyzing key pixels in 2D images. However, in a 3D CNN, the process is extended to accommodate volumetric data. This means that instead of analyzing flat, two-dimensional pixel arrays, the CNN processes three-dimensional blocks or voxels, thus capturing spatial depth and texture in a more holistic manner.

3.2.1. Color Characteristics in 3D Space. Color characteristics are fundamental in image recognition. For 3D CNNs, this extends beyond mere pixel color values to include spatial color distributions within the 3D space. The color characteristics are thus expressed in a more complex 3D color space, enhancing the stability and robustness of the feature extraction. For instance, the HSV color model, often used in 2D, is adapted into a 3D model, considering not only hue, saturation, and value but also their distribution in three-dimensional space.

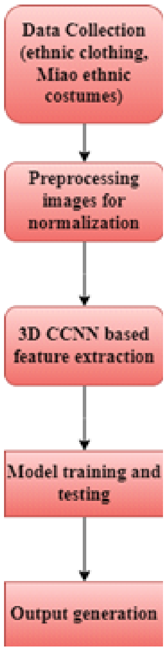

Fig. 3.1: Proposed Architecture

3.2.2. Spatial Color Characteristics in 3D. In a 3D CNN, the spatial distribution of colors is crucial. Unlike 2D systems, a 3D CNN can capture how colors are distributed spatially within the volume of an object. This is done through more advanced forms of region segmentation and voxel-based color histograms, leading to a richer and more accurate representation of the clothing's color features.

3.2.3. 3D Convolutional Layer. The 3D Convolutional Layer is essential for processing three-dimensional data, making it ideal for constructing a digital ethnic clothing library. Unlike its 2D counterpart, this layer handles volumetric data, allowing it to capture spatial and textural information in three dimensions. Each neuron in the layer filters a small, local region of the input volume (covering all its depth), analyzing patterns and features like folds and textures in ethnic clothing. This process is crucial for understanding the intricate designs and structures of the clothing, as the layer learns to identify various features across different layers of depth within the clothing's fabric.In a 3D CNN, the convolutional layer operates over 3D data. The convolution operation in 3D is mathematically expressed as

$$f\left(x,y,z\right)=\sum_{i,j,k} v\left(i,j,k\right).k(x-i,y-j,z-k)$$

where $f\left(x,y,z\right)$is the output feature map in 3D, $v\left(i,j,k\right)$represents the voxel values, and k is the 3D kernel.

3.2.4. 3D Transpose Convolution Layer. The 3D transpose convolution layer, also known as a deconvolution layer, is vital for upscaling feature maps in a 3D CNN. This layer works inversely compared to the convolutional layer, increasing the spatial resolution of the input feature maps. In the context of ethnic clothing, this means that it can reconstruct or enhance the details lost during downsampling in previous layers, effectively filling in the finer details of clothing textures and patterns. This is particularly important for accurately rendering the intricate designs and fine details in ethnic attire, ensuring that the digital representation maintains the authenticity and richness of the original piece. For upsampling in 3D space, a 3D transpose convolution layer is used which is mathematically represented by

$$g\left(x,y,z\right)=\sum_{i,j,k} w\left(i,j,k\right)\odot t(x+i,y+j,z+k)$$

Here, $g\left(x,y,z\right)$is the upsampled output, w is the transposed convolution kernel, and t represents the 3D tensor being upsampled.

3.2.5. 3D Pooling Layer. The 3D Pooling Layer in a CNN is designed to progressively reduce the spatial size of the input volume. This downsampling operation simplifies the amount of computation required by the network, controls overfitting, and makes the representation more manageable. In the digital representation of ethnic clothing, the pooling layer helps in abstracting the higher-level features from the raw spatial data, like identifying general patterns or shapes in clothing. It enhances the network's focus on essential features while discarding irrelevant variances and noises in the dataset, making the model more robust and efficient. Pooling in 3D reduces the spatial dimensions of the feature maps while retaining important features. This can be represented as:

$$p\left(x,y,z\right)=max_{i,j,k\in window}v(x+i,y+j,z+k)$$

where $p\left(x,y,z\right)$ is the pooled feature map.

3.2.6. Fully Connected Layer in 3D CNN. In a 3D CNN, the Fully Connected Layer serves to integrate the high-level, abstracted features extracted from the previous layers into a final output, such as a classification of ethnic clothing types. This layer flattens the 3D feature maps into a single vector, allowing the network to learn non-linear combinations of these high-level features. For ethnic clothing, this means combining various spatial and textural details to form a comprehensive understanding of the clothing's style and design. This layer plays a crucial role in making final predictions or classifications based on the entirety of the learned features. The fully connected layer can be described as

$$y=ReLU(w.x+b)$$

where x is the input from the flattened 3D feature maps, w and bb are the weights and biases, and y is the output.

3.2.7. Loss Layer in 3D CNN. The Loss Layer in a 3D CNN quantifies the error between the network's predictions and the actual data. It is critical in guiding the training process, allowing the model to adjust and improve its parameters for more accurate predictions. In digitalizing ethnic clothing, the loss layer assesses how well the CNN is performing in terms of accurately capturing and representing the complex, multidimensional aspects of the clothing. A well-calibrated loss layer ensures that the network effectively learns the intricate details and unique characteristics of different ethnic garments, leading to a high-fidelity digital representation. The loss layer in a 3D CNN calculates the difference between the predicted and actual values, crucial for network training. For instance, the mean squared error (MSE) can be used

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$$

where y_i is the predicted value and $\hat{y}_i$is the actual value.

By adapting these principles into a 3D CNN architecture, it becomes possible to create a more robust and detailed digital ethnic clothing library, capturing not just the visual aspects but also the spatial characteristics intrinsic to ethnic clothing. This leads to a more accurate and immersive digital representation, greatly benefiting the preservation and study of ethnic clothing cultures.

4. Results and Experiments.

4.1. Simulation Setup. In this section the evaluation of the proposed model is based on the study []. Regarding the simulation of the study we proceed the validation of our proposed approach.

In the context of our proposed Customized 3D CNN for digitalizing and analyzing ethnic clothing, specifically Miao ethnic costumes, this dataset plays a pivotal role. Comprising 1000 images, it offers a rich variety of patterns, textures, and color schemes inherent to Miao ethnic attire. The dataset's diversity is crucial for training the 3D CNN, enabling it to learn and recognize the intricate details and unique characteristics of these costumes. By utilizing this dataset, the Customized 3D CNN can be rigorously trained and tested for its accuracy in image retrieval. The average retrieval accuracy calculated from these 1000 images will provide insightful data on the effectiveness of the 3D CNN in recognizing and differentiating the complex features of Miao costumes. This is essential for achieving our goal of preserving and cataloging ethnic clothing digitally, where precise identification and categorization are key. The dataset thus not only aids in the technical development of the model but also contributes to the cultural aspect of preserving and understanding ethnic heritage through digital means.

4.2. Evaluation Criteria. Figure 4.1 demonstrates the efficacy of proposed 3D CNN in terms of accuracy. In the context of our proposed 3D CNN the accuracy values range from 94.25% for SVM to 97.88% for the 3D Customized CNN. The traditional CNN shows a high accuracy of 95.48%, indicating its reliability in correctly identifying images. However, the 3D Customized CNN excels with an accuracy close to 98%, suggesting it has a superior capability in correctly classifying both positive and negative instances in the dataset. This high accuracy is particularly significant in complex image recognition tasks, where distinguishing between numerous and varied elements is critical. It reflects the model's overall effectiveness and reliability, demonstrating its proficiency in handling a broad range of scenarios with minimal errors. The high accuracy of the 3D Customized CNN indicates its advanced ability to analyze and interpret the spatial and textural details in images, a crucial aspect of image recognition tasks. This makes it an invaluable tool in scenarios where precise and accurate categorization of images is essential, such as in digital archiving of cultural heritage items like ethnic costumes.

Precision quantifies the accuracy of the positive predictions made by a model, essentially measuring the proportion of true positives against all positive predictions (both true positives and false positives). Figure 4.1 present precision values of the models, here precision values vary, with the 3D Customized CNN achieving the highest precision of 97.12%. This high precision indicates that when this model predicts an image as belonging to a particular category, there is a 97.12% chance that it is indeed correct. In contrast, the SVM, CNN, BBNN, and PSO with a precision of 95.04%, 93.02%, 94.99%, and 95.88% respectively. High precision is crucial in situations where the cost of false positives is high. For instance, in medical diagnostics, wrongly identifying a healthy patient as sick (a false positive) could lead to unnecessary and potentially harmful treatment. The superior precision of the 3D Customized CNN highlights its ability to make highly accurate positive predictions, crucial in fields where precision is more important than recall.

Recall, also known as sensitivity, measures the model's ability to identify all actual positive instances, calculated as the proportion of true positives to the sum of true positives and false negatives. This metric is particularly important in scenarios were missing out on positive instances could have dire consequences. According to the Figure 4.1, the 3D Customized CNN again outperforms other models with a recall of 97.55%. This high recall rate implies that the model is highly effective in identifying positive instances, missing very few actual positives. For instance, in security settings, a high recall rate would mean that the system rarely misses identifying a genuine threat. In contrast, the SVM, CNN, BPNN and PSO with a recall of 93.21%,94.06%, 94.78%, and 95.63 respectively. The 3D Customized CNN's high recall is indicative of its robustness in capturing and correctly identifying the nuanced features in the data, crucial for comprehensive and accurate image recognition.

The F1-Score is a harmonic mean of precision and recall, providing a balance between the two. It is especially useful when the cost of false positives and false negatives is uneven, or when the class distribution is imbalanced. The F1-Score is the most telling metric in scenarios where both recall and precision are important. In the provided data, the 3D Customized CNN achieves the highest F1-Score of 96.99%, indicating an excellent balance between precision and recall. This balance is critical in many real-world applications where neither false positives nor false negatives can be afforded. Compared with existing SVM, CNN, BBNN and PSO it achieves 93.12%, 92.88%, 93.28% and 95.12% respectively in Figure 4.1. For instance, in legal or financial contexts, the consequences of both types of errors can be severe. The high F1-Score of the 3D Customized CNN demonstrates its efficiency in not only accurately identifying the correct instances but also in minimizing the number of incorrect identifications. This makes it a highly reliable model for complex tasks where precision and recall are both equally important, and the costs of errors are high.

5. Conclusion. The study of the 3D CNN in image recognition has demonstrated remarkable efficacy, marking a significant advancement in the field of computer vision and machine learning. The implementation

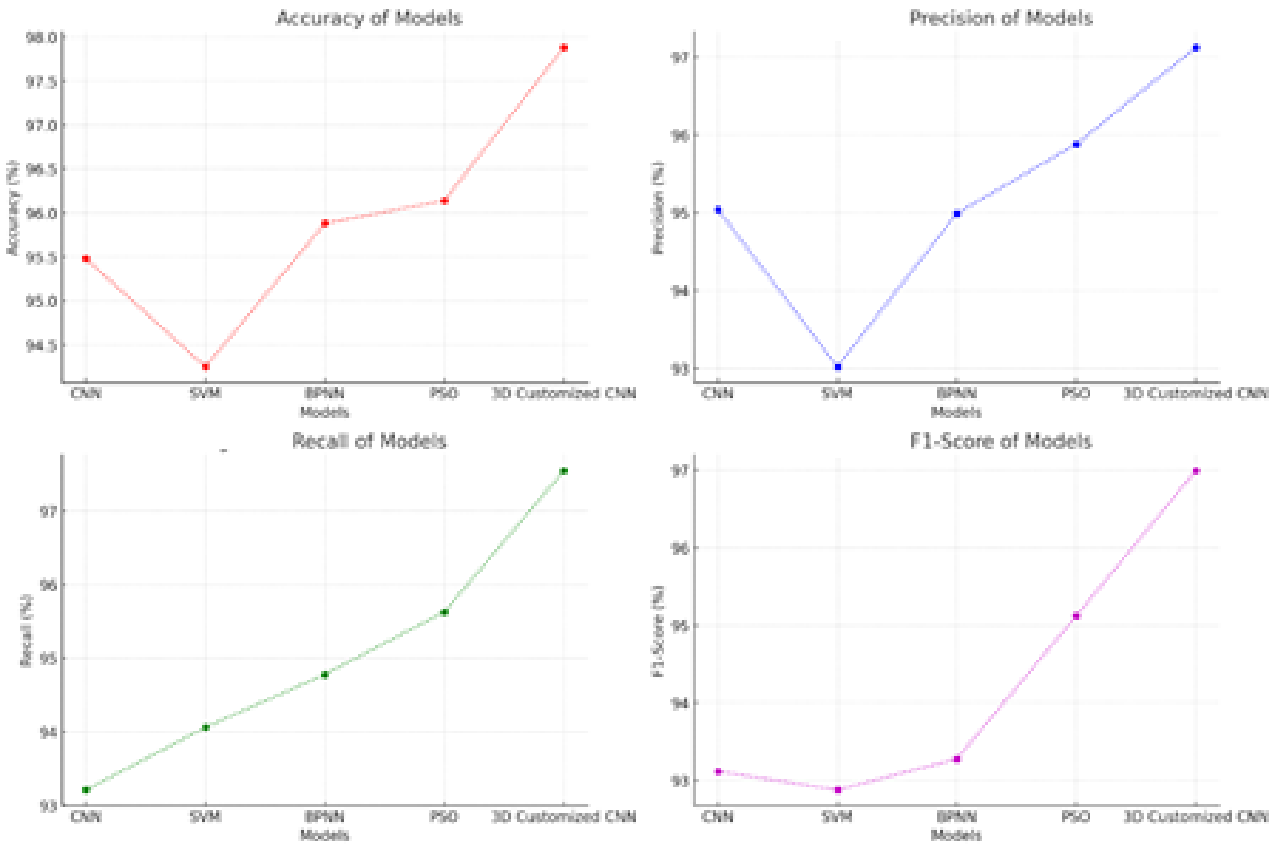

Fig. 4.1: Accuracy, Precision, Recall and F1-score of models

of a 3D CNN, specifically designed for intricate tasks like ethnic clothing recognition, has shown superior performance in various key metrics compared to traditional models. With an impressive accuracy rate, it has proven its ability to correctly identify a vast majority of instances, both positive and negative, in diverse datasets. The model's high precision indicates a strong capability in minimizing false positives, a crucial aspect in applications where the cost of error is substantial. Moreover, the 3D CNN's remarkable recall rate highlights its effectiveness in capturing almost all actual positives, ensuring that very few relevant features are missed. This aspect is particularly vital in critical scenarios like medical diagnosis or security systems, where overlooking positive instances could have severe consequences. The balanced F1-score further cements the model's robustness, demonstrating its proficiency in maintaining a harmonious balance between precision and recall. Overall, the study underscores the potential of 3D CNNs in handling complex, multi-dimensional data with high efficiency and accuracy. This paves the way for their broader application in various fields that require nuanced image recognition and categorization, opening new frontiers in digital analysis and automation. The 3D CNN not only enhances current methodologies but also sets a benchmark for future developments in AI-driven image analysis.

6. Limitations and Future Scope. The study, while pioneering in its application of deep learning for image recognition, encounters several limitations that pave the way for future research. One of the primary constraints lies in the data dependency of deep learning models. The performance of these models is heavily reliant on the quantity and quality of the training data, making them vulnerable to biases and inaccuracies in datasets. This is particularly challenging in fields where data is scarce or unbalanced. Furthermore, the complexity of deep learning models, especially in terms of their interpretability and computational demands, poses a significant challenge. The 'black box' nature of these models often makes it difficult to understand the reasoning behind their decisions, a crucial aspect in sensitive applications like medical diagnosis. Looking ahead, the future scope of this study is vast and promising. One potential direction is the development of more sophisticated data augmentation techniques to address the issue of limited and imbalanced datasets. Another avenue is the exploration of explainable AI (XAI) methods to enhance the transparency and interpretability of deep learning models, making them more trustworthy and accessible to users. Additionally, optimizing the computational efficiency of these models can make them more feasible for real-time applications and accessible to organizations with limited resources. This study's advancements also open the possibility of exploring new applications of deep learning in uncharted territories, further expanding the horizons of image recognition and its impact across various domains.

Acknowledgement. This work was sponsored in part by Shanxi University philosophy and social science research project Research on the Path of Inheritance and Protection of Intangible Cultural Heritage in Shanxi under the Integration of Information Technology (2023W24).

REFERENCES

[1] A. Amato, S. Venticinque, and B. Di Martino, *Image recognition and augmented reality in cultural heritage using opencv*, in Proceedings of International Conference on Advances in Mobile Computing & Multimedia, 2013, pp. 53–62.

[2] R. Chauhan, K. K. Ghanshala, and R. Joshi, *Convolutional neural network (cnn) for image detection and recognition*, in 2018 first international conference on secure cyber computing and communication (ICSCCC), IEEE, 2018, pp. 278–282.

[3] R. Chen, L. Pan, C. Li, Y. Zhou, A. Chen, and E. Beckman, *An improved deep fusion cnn for image recognition.*, Computers, Materials & Continua, 65 (2020).

[4] L. Cui, X. Shao, B. Mago, and R. V. Ravi, *Digital intangible cultural heritage management using deep learning models*, Aggression and Violent Behavior, (2021), p. 101680.

[5] D. Devarajan, D. S. Alex, T. Mahesh, V. V. Kumar, R. Aluvalu, V. U. Maheswari, and S. Shitharth, *Cervical cancer diagnosis using intelligent living behavior of artificial jellyfish optimized with artificial neural network*, IEEE Access, 10 (2022), pp. 126957–126968.

[6] T.-N. Do, T.-P. Pham, N.-K. Pham, H.-H. Nguyen, K. Tabia, and S. Benferhat, *Stacking of svms for classifying intangible cultural heritage images*, in Advanced Computational Methods for Knowledge Engineering: Proceedings of the 6th International Conference on Computer Science, Applied Mathematics and Applications, ICCSAMA 2019 6, Springer, 2020, pp. 186–196.

[7] Y.-n. Dong and G.-s. Liang, *Research and discussion on image recognition and classification algorithm based on deep learning*, in 2019 International conference on machine learning, big data and business intelligence (MLBDBI), IEEE, 2019, pp. 274–278.

[8] T. Fan, H. Wang, and S. Deng, *Intangible cultural heritage image classification with multimodal attention and hierarchical fusion*, Expert Systems with Applications, (2023), p. 120555.

[9] W. Fang, F. Zhang, V. S. Sheng, and Y. Ding, *A method for improving cnn-based image recognition using dcgan.*, Computers, Materials & Continua, 57 (2018).

[10] H. Fujiyoshi, T. Hirakawa, and T. Yamashita, *Deep learning-based image recognition for autonomous driving*, IATSS research, 43 (2019), pp. 244–252.

[11] L. Jiao and J. Zhao, *A survey on the new generation of deep learning in image processing*, Ieee Access, 7 (2019), pp. 172231–172263.

[12] Y. Li, *Research and application of deep learning in image recognition*, in 2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), IEEE, 2022, pp. 994–999.

[13] E. Liu, *Research on image recognition of intangible cultural heritage based on cnn and wireless network*, EURASIP Journal on Wireless Communications and Networking, 2020 (2020), pp. 1–12.

[14] Y. Liu, P. Cheng, and J. Li, *Application interface design of chongqing intangible cultural heritage based on deep learning*, Heliyon, 9 (2023).

[15] A. Mikołajczyk and M. Grochowski, *Data augmentation for improving deep learning in image classification problem*, in 2018 international interdisciplinary PhD workshop (IIPhDW), IEEE, 2018, pp. 117–122.

[16] T. Pistola, S. Diplaris, C. Stentoumis, E. A. Stathopoulos, G. Loupas, T. Mandilaras, G. Kalantzis, I. Kalisperakis, A. Tellios, D. Zavraka, et al., *Creating immersive experiences based on intangible cultural heritage*, in 2021 IEEE International Conference on Intelligent Reality (ICIR), IEEE, 2021, pp. 17–24.

[17] M. Subramanian, J. Cho, V. E. Sathishkumar, and O. S. Naren, *Multiple types of cancer classification using ct/mri images based on learning without forgetting powered deep learning models*, IEEE Access, 11 (2023), pp. 10336–10354.

[18] F. Tao, W. Hao, L. Yueyan, and D. Sanhong, *Classifying images of intangible cultural heritages with multimodal fusion*, Data Analysis and Knowledge Discovery, 6 (2022), pp. 329–337.

[19] B. B. Traore, B. Kamsu-Foguem, and F. Tangara, *Deep convolution neural network for image recognition*, Ecological informatics, 48 (2018), pp. 257–268.

[20] S. VE AND Y. CHO, *Mrmr-eho-based feature selection algorithm for regression modelling*, Tehnički vjesnik, 30 (2023), pp. 574–583.

[21] C. S. WON, *Multi-scale cnn for fine-grained image recognition*, IEEE Access, 8 (2020), pp. 116663–116674.

[22] J. YIN, *Application of intelligent image recognition and digital media art in the inheritance of black pottery intangible cultural heritage*, ACM Transactions on Asian and Low-Resource Language Information Processing, (2023).

Edited by: Rajanikanth Aluvalu

Special issue on: Evolutionary Computing for AI-Driven Security and Privacy: Advancing the state-of-the-art applications

Received: Jan 31, 2024

Accepted: Mar 11, 2024