Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **2**, June **2025**

CUFL: A novel Continuously Updated method for characterizing educational image recognition based on Few-shot Learning

Thien Le Quang

Faculty of Management Information Systems
Ho Chi Minh University of Banking
Ho Chi Minh City, Vietnam
thienlq@hub.edu.vn

Received February 2, 2025, revised April 5, 2025, accepted April 8, 2025.

Abstract. *Educational image recognition plays a crucial role in effectively managing classrooms. One of the methods to recognize educational images in classrooms is through few-shot deep learning. However, the drawback of this approach is the potential loss of distinctive features when integrating new features due to the limited data and few samples. To address this issue, we developed a continuously updated feature learning (CUFL) method to support educational image recognition. The continuous training process updates and generates feature matrices to facilitate future educational image recognition. We qualitatively and quantitatively evaluated the proposed model on two datasets: ImageNet-mini and T-Edu (a dataset we collected in classrooms at Thuyloi University). Experimental results demonstrate that CUFL outperforms current state-of-the-art methods.*

Keywords: Educational image recognition, few-shot learning, image processing

1. **Introduction.** Educational image recognition is a critical problem in the field of education in the era of Industry 4.0 [1]. It can support various tasks such as management, teaching assistance, training, and staff operations. Additionally, in recent years, educational systems have frequently implemented camera installations across school campuses, enabling the retrieval of footage and timely decision-making in case of incidents. As a result, the volume of educational image data is enormous and must be exploited efficiently using artificial intelligence (AI).

Machine learning and deep learning [2, 3, 4] models have been widely applied in every aspect of human life, including education, where deep learning models are indispensable. In recent years, a prominent model called few-shot deep learning [5] has been increasingly used to address problems involving limited and unlabeled data. However, a significant limitation of this approach is the storage and representation of features, as well as the model's susceptibility to degradation when new features are added. To address this issue, we propose a method capable of overcoming these challenges.

Furthermore, few-shot deep learning methods also require well-designed loss functions to address several challenges, such as: (1) avoiding overfitting, (2) training the model with limited data, and (3) accurately representing the characteristics of the educational labels to be recognized. We have refined the loss function to enhance the adaptability of few-shot deep learning models for educational image recognition tasks. Additionally, we have developed a process to ensure minimal changes during data updates and continuous training.

In summary, this paper accomplishes the following tasks:

- Develop a few-shot deep learning model capable of updating features during continuous retraining, named CUFL.
- Design a loss function suitable for the educational image recognition problem.
- Build a dataset specifically for educational image recognition, named T-Edu.
- Evaluate the proposed model using both qualitative and quantitative methods.

2. Related Works.

2.1. Baseline convolutional neural network for the educational image recognition task. Convolutional neural networks [6, 7, 8] trained for the educational image recognition task [9, 10] typically rely on supervised [10], unsupervised, and semi-supervised [11] learning approaches. These methods require training on large, labeled datasets, posing significant challenges in data preparation and processing.

Approaches for addressing educational image recognition problems using convolutional neural networks often involve repeated training on various datasets [10], including the optimization of hyperparameters [12], the development of advanced loss functions [13], or leveraging modern backbones [14] tailored to the educational image recognition task. However, these methods primarily address isolated subproblems and fail to comprehensively solve broader challenges related to data and flexibility in educational image recognition tasks.

2.2. Few-shot learning in the educational image recognition task. Few-shot deep learning [15, 16, 17] is approached as the reuse of data [17, 18] or limited data [19] in educational image recognition tasks [20]. This method requires the base model to be trainable on small datasets while still being applicable for future predictions. Most approaches focus on constructing feature spaces [16] or loss functions that can be learned with minimal data [19]. However, these methods have limitations, and retraining the model often results in the loss of previously learned features, effectively resembling a complete retraining from scratch.

Some models can fine-tune parameters [21, 22] to adapt to previously trained models. Liu [23] proposed a ranking-based model to retain features during prior training processes. Additionally, Wanyan et al. [24] introduced an attention mechanism for few-shot deep learning to reuse features during training. Building upon this foundation, we aim to develop a few-shot deep learning approach that enables the continuous updating of features after multiple training cycles, ensuring adaptability and compatibility with diverse educational image datasets.

3. Proposed Method.

3.1. Problem Definition. In the context of educational image recognition, the Continuously Updated based on Few-Shot Learning (CUFL) problem addresses the challenge of training a deep learning model to incrementally learn new classes across multiple sessions $\{D^{(0)}, D^{(1)}, \ldots, D^{(T)}\}$ without degrading the quality of previously learned feature representations. Each session $D^{(t)}$ contains a training set $\{(x_i, y_i)\}_{i=1}^{|D^{(t)}|}$, where x_i is an input image and y_i is its corresponding label. The framework operates under the following constraints: (1) Base Session ($D^{(0)}$): Provides a comprehensive label set $C^{(0)}$ with abundant training samples per class, serving as the foundation for initial feature learning; (2) Incremental Sessions ($D^{(t)}, t > 0$): Introduce new classes $C^{(t)}$ under a strict p-way q-shot

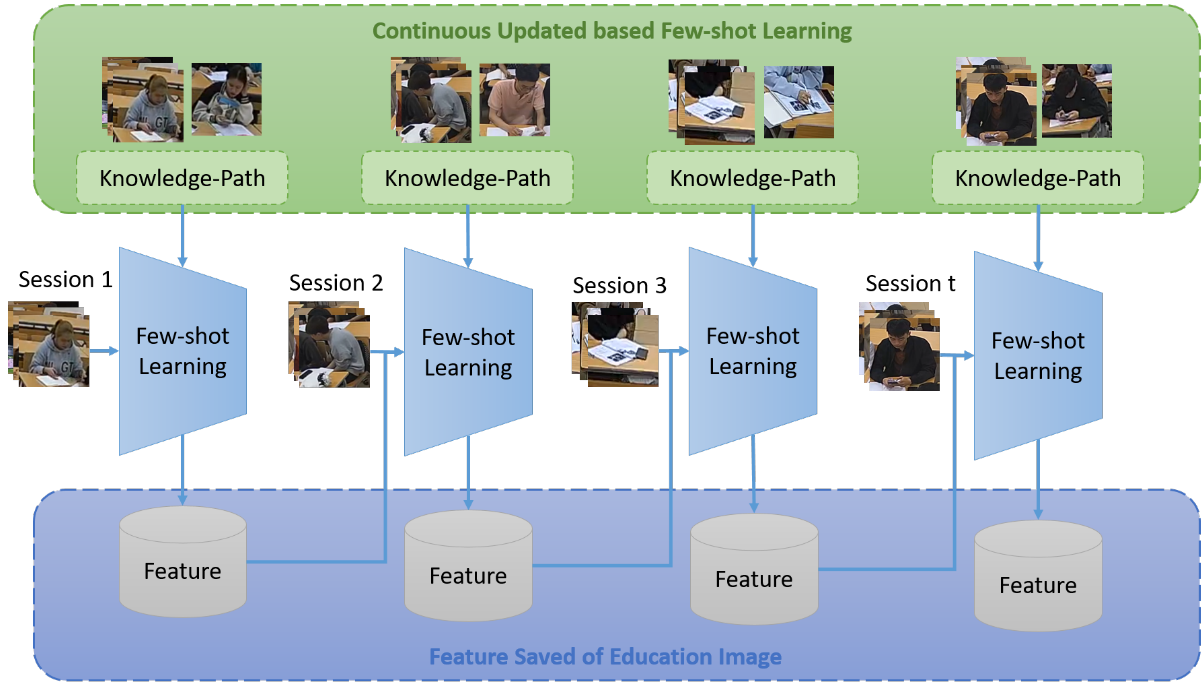

FIGURE 1. The framework of the proposed CUFL

protocol (i.e., p novel classes with only q labeled examples per class). Class sets across sessions are disjoint: $C^{(t)} \cap C^{(t')} = \emptyset, \forall t \neq t'$; and (3) Data Constraints: Training data from prior sessions are inaccessible in subsequent sessions, except for limited exemplars used for replay to mitigate forgetting.

The primary challenges include:

- Catastrophic Forgetting: The model risks losing discriminative power for old classes when adapting to new classes with minimal data.
- Feature Representation Stability: Educational images (e.g., diagrams, formulas, or handwritten notes) require robust and stable feature embeddings that remain discriminative across incremental updates.
- Accuracy-Compatibility Trade-off: The model must maintain high classification accuracy over all encountered classes $\bigcup_{i=0}^{t} C^{(i)}$ during evaluation at session t, balancing the integration of new knowledge with the preservation of old knowledge.

The objective of CUFL in this domain is to design a dynamic feature update mechanism that enables the model to (1) assimilate novel classes from few-shot examples without distorting the feature space of old classes, and (2) sustain high performance across all sessions, even after repeated retraining. This mechanism must prioritize the retention of critical features for educational content recognition, such as fine-grained visual patterns in STEM diagrams or symbolic representations, while efficiently adapting to emerging classes with limited supervision.

3.2. **Model Architecture.** Figure 1 illustrates a Continuous Updated based Few-shot Learning (CUFL), designed to process images in the field of education. The system consists of three main components: Knowledge-Path, Few-shot Learning, and Feature Saved. Firstly, the Knowledge-Path, presented at the top of the image, represents the input datasets comprising images of educational activities (e.g., students taking notes or working on tasks). These images serve as input data for the system to initiate the learning process.

In each step, the Few-shot Learning modules, depicted as blue cone-shaped blocks, process and analyze the images from each Knowledge-Path. The Few-shot Learning approach enables the system to acquire new knowledge from a minimal number of samples, effectively extracting significant features from the input images without requiring a large dataset. The extracted features are then stored in databases (represented by gray cylindrical blocks) within the Feature Saved component. This stored knowledge can be reused, allowing the system to enhance its performance in subsequent learning steps.

The system operates under a continuous update mechanism, where knowledge learned from different Knowledge-Paths is integrated and refined over time. This ensures continuous learning and analysis, enabling the system to improve its effectiveness progressively. In practical applications, the system can be used to monitor and evaluate students' learning activities through image analysis, automatically recognize learning behaviors, or support machine learning in scenarios with limited data. This approach not only reduces training costs but also optimizes the ability to process educational data in various environments.

3.3. **Loss function.** The loss function of the CUFL architecture is designed based on the Adaptive Layer Selection mechanism using Adaptive Sparse Rate (ASR) to balance adaptive learning on new layers (adapter training) and knowledge preservation from frozen layers. The overall loss function is defined as:

$$L_{\text{total}} = L_{\text{task}} + \lambda \cdot L_{\text{preserve}} \tag{1}$$

where L_{task} is the standard task-specific loss (e.g., Cross-Entropy Loss for image classification), computed from the model's output after training the adapters in the K selected layers (with the smallest ASR) on new data. Meanwhile, L_{preserve} ensures the stability of the unselected layers ($N - K$ layers), preventing large changes that might lead to knowledge forgetting. This preservation loss is defined using the singular values from the covariance matrix decomposition ($C(t)$) as:

$$L_{\text{preserve}} = \sum_{l \notin L_{\text{selected}}} \|\Sigma_l(t) - \Sigma_l(t-1)\|_2^2, \tag{2}$$

where $\Sigma_l(t)$ represents the singular values for layer l at session t. By minimizing fluctuations in singular values across sessions, the regularization retains the "redundant capacity" of the frozen layers. The trade-off between learning new tasks and preserving previous knowledge is controlled by the hyperparameter λ, which adjusts the emphasis on L_{preserve} relative to L_{task}. At each session $t + 1$, ASR is recalculated based on the updated replay data ($D_{\text{replay}}(t + 1)$) to determine the set of layers $L_{\text{selected}}(t + 1)$ for adaptation. The loss function dynamically focuses training on these K layers while imposing stability constraints on the remaining $N - K$ layers through L_{preserve}.

4. **Experiments.**

4.1. **Dataset.** MiniImageNet [25] is a compact dataset designed for testing Few-shot Learning models, derived from the renowned ImageNet dataset. This dataset comprises a total of 100 classes, each containing 600 low-resolution images (84×84 pixels), which helps reduce computational costs compared to the full ImageNet dataset. MiniImageNet is typically divided into three subsets: Train (64 classes), Validation (16 classes), and Test (20 classes). This division ensures that the classes in the training set do not overlap with those in the test set, creating a more realistic Few-shot Learning testing environment where models need to generalize to unseen classes.

The T-Edu dataset consists of multiple '.mp4' video files, each approximately 1.015 MB in size, recorded in classroom environments. We collected 10 videos from the T-EDU dataset. Each video represents a specific camera angle within a classroom and has an average duration of 60 minutes. These videos are labeled systematically using formats such as ‘DXX_YYYYMMDDHHMMSS.mp4‘, where ‘DXX‘ represents different cameras or classroom positions. The dataset contains 30 object labels, such as male students, female students, desks, chairs, books, and laptops, and 28 action labels, such as using a phone, opening a notebook, turning sideways, and writing. The data is continuous, capturing dynamic classroom activities with objects and short- or long-term actions. This presents a dual challenge for models to process both spatial features (objects) and temporal features (actions).

Each object or action in the dataset is annotated using LabelMe [26], a tool that allows precise bounding box annotation and label assignment. Objects like students, desks, and books are marked with bounding boxes, while actions such as using a phone or opening a notebook are annotated with temporal and spatial context. This ensures the dataset is well-prepared for object detection and action recognition tasks, with high-quality annotations to support robust model training.

To train the CUFL (Continuous Updated Few-shot Learning) model, the dataset can be divided into incremental training sessions $(t_1, t_2, ..., t_n)$. Each session uses data from specific cameras or locations to simulate continuous learning. For example, session t_1 could include videos from ‘D01‘, focusing on familiar objects (students, desks) and actions (writing, turning sideways). Subsequent sessions (t_2, t_3) introduce new cameras (‘D02‘, ‘D03‘) with unseen objects or actions to encourage adaptive learning. To mitigate catastrophic forgetting, replay data (D_{replay}) from earlier sessions is selected (e.g., 10% of prior samples) and reintroduced during training.

Adaptive Sparse Rate (ASR) mechanisms dynamically identify layers or features requiring updates for new tasks, while freezing others to preserve prior knowledge. Task Loss (L_{task}) optimizes performance on new labels, while Preservation Loss (L_{preserve}) stabilizes previously learned features by minimizing significant changes in frozen layers. Each session recalculates ASR to prioritize adaptation to new data while maintaining knowledge integrity.

This strategy supports efficient incremental learning for CUFL in a classroom setting, allowing the model to continuously adapt to new objects and behaviors while retaining past knowledge. Practical applications include automated classroom monitoring, detecting learning behaviors (e.g., writing, discussing, using phones), and analyzing relevant objects (e.g., desks, books, notebooks) to enhance educational analytics.

4.2. **Experiment setup.** We implement the CUFL framework using a Vision Transformer (ViT) as the feature extraction backbone and a Mamba-based Few-Shot Learning architecture for incremental classification. The model is trained with Stochastic Gradient Descent (SGD) at a learning rate of 0.005 for 58 epochs, incorporating early stopping to prevent overfitting. For few-shot incremental learning, the data is partitioned into 06 sequential training sessions $\{D^{(0)}, D^{(1)}, \ldots, D^{(5)}\}$, where the base session $D^{(0)}$ includes a sufficiently labeled class set $C^{(0)}$ for initial feature learning. Subsequent sessions ($t > 0$) adhere to a 5-way 5-shot protocol: each session introduces 5 novel classes with 5 labeled examples per class. To ensure disjoint class sets across sessions ($C^{(t)} \cap C^{(t')} = \emptyset, \forall t \neq t'$), classes are evenly partitioned and assigned to distinct sessions. we set the trade-off hyperparameter λ to 0.5, which we found to provide a good balance between the task-specific loss (L_{task}) and the preservation loss (L_{preserve}). This configuration enables CUFL to incrementally integrate new classes from limited data while maintaining feature space stability, thereby optimizing the balance between *knowledge acquisition* and *knowledge retention* in few-shot class-incremental learning.

TABLE 1. CUFL performance comparison on miniImageNet dataset.

Method	Accuracy in each session						Average accuracy
	0	1	2	3	4	5	
TEEN [27]	0.7055	0.6323	0.6053	0.5524	0.5324	0.5208	0.5915
GFSL [28]	0.6650	0.6038	0.5653	0.4996	0.4870	0.4532	0.5457
FACT [29]	0.7590	0.7084	0.6556	0.6174	0.5841	0.5694	0.6490
LIMIT [30]	0.7589	0.7199	0.6741	0.6135	0.5866	0.5741	0.6545
FeSSSS [31]	0.8150	0.7704	0.6956	0.6434	0.6055	0.5887	0.6864
CUFL (ours)	**0.8842**	**0.8526**	**0.8168**	**0.7526**	**0.7263**	**0.7025**	**0.7892**

TABLE 2. CUFL performance comparison on T-Edu dataset (ours).

Method	Accuracy in each session						Average accuracy
	0	1	2	3	4	5	
TEEN [27]	0.7364	0.7035	0.6824	0.6624	0.6397	0.6025	0.6712
GFSL [28]	0.6927	0.6694	0.6397	0.6026	0.5826	0.5584	0.6243
FACT [29]	0.7822	0.7472	0.7194	0.6812	0.6521	0.6335	0.7026
LIMIT [30]	0.7952	0.7635	0.7395	0.7025	0.6876	0.6635	0.7253
FeSSSS [31]	0.8526	0.8375	0.8034	0.7843	0.7524	0.7264	0.7927
CUFL (ours)	**0.9045**	**0.8852**	**0.8364**	**0.8054**	**0.7842**	**0.7581**	**0.8289**

We formulate the following research questions (RQs) to guide our investigation:

- **RQ1:** How does the CUFL framework perform compared to state-of-the-art few-shot learning methods in terms of classification accuracy and feature stability?
- **RQ2:** What is the predictive performance of CUFL when applied to real-world educational image recognition tasks, particularly with fine-grained or domain-specific visual content?

4.3. **Performance Compare (RQ1).** The performance comparison tables of CUFL on the miniImageNet dataset (table 1) and the T-Edu dataset (table 2) demonstrate consistency in maintaining a superior advantage over existing Few-Shot Class-Incremental Learning (FSCIL) methods. On the T-Edu dataset (a real-world educational image dataset constructed in this study), CUFL achieves an average accuracy of 82.89% over six learning sessions, outperforming the previous strongest method (FeSSSS [31]: 79.27%) by approximately 3.6% and older methods such as TEEN [27] (67.12%) and GFSL (62.43%) by around 10%. This confirms CUFL's adaptability to domain-specific data like educational images, which require fine-grained recognition (e.g., object, student activity).

In the base session (session 0), CUFL attains 90.45% accuracy on T-Edu, significantly higher than its performance on miniImageNet (88.42%), reflecting the advantage of the Vision Transformer (ViT) architecture in learning features from structured and less noisy data. As sessions progress, CUFL's accuracy decline on T-Edu is also slower compared to miniImageNet: from session 0 to session 5, accuracy drops by 14.64% on T-Edu (90.45% to 75.81%), whereas on miniImageNet, the drop is 18.17% (88.42% to 70.25%). This difference suggests that CUFL better leverages the stability of educational data, which often exhibits repetitive structures and lower variability compared to general datasets.

Comparing the two datasets, baseline methods (e.g., FeSSSS [31], LIMIT [30]) consistently perform better on T-Edu than on miniImageNet, likely due to the lower noise level and higher homogeneity of educational images. However, the performance gap between CUFL and other methods on T-Edu remains around 3–5%, highlighting the superiority

FIGURE 2. Prediction of CUFL model

of the Mamba mechanism in addressing FSL-specific challenges. Notably, in the final session (session 5), CUFL achieves 75.81% on T-Edu, surpassing FeSSSS (72.64%) by 5% and TEEN (60.25%) by 15%, demonstrating effective resistance to catastrophic forgetting even as the number of learning sessions increases.

4.4. **Qualitative Study (RQ2).** Figure 2 illustrates the prediction results of the CUFL model throughout different training sessions. In session 1, the model is trained with only a single label, "girl," along with other simple labels, resulting in very high accuracy. In the second training phase (session 2), additional images of male students are introduced, and the model maintains relatively high accuracy.

As more categories are added, such as books, notebooks, pens, and other classroom-related objects, accuracy gradually declines. Some misclassifications appear, such as predicting "boy" for an image of a book. Misclassified images are highlighted with red borders, whereas correctly classified images have green borders.

After six training sessions, the model has been exposed to an increasing number of labels and datasets. However, its overall performance does not degrade significantly. CUFL demonstrates the ability to recognize basic classroom behaviors (e.g., talking, using a phone), making it well-suited for educational applications. Nevertheless, some ambiguous or irrelevant labels indicate potential noise in the input data or gaps in the training process, which could affect the model's reliability in real-world applications.

5. **Conclusions.** In this paper, we propose CUFL, an efficient solution to address the challenges of Few-shot learning by decomposing model weights into components that store essential features while maintaining adaptability. By freezing key feature components that retain prior knowledge and adjusting redundant capacities, our framework strikes a balance between preserving learned knowledge and acquiring new tasks, particularly in the domain of educational image recognition. Our adaptive layer selection feature further enhances this balance by dynamically allocating adapters based on their sensitivity. Unlike existing methods that rely on attention mechanisms or prompt-based modules, CUFL focuses purely on label features throughout multiple training sessions. Extensive experiments on multiple benchmark datasets demonstrate that CUFL outperforms state-of-the-art methods, proving its effectiveness in mitigating feature forgetting across multiple incremental training stages while maintaining data balance within the Few-shot learning paradigm. Future work can explore integrating semantic guidance and

optimizing the sensitivity estimation process to further enhance CUFL's scalability and generalization.

REFERENCES

[1] Tri, Nguyen Minh, Pham Duy Hoang, and Nguyen Trung Dung. "Impact of the industrial revolution 4.0 on higher education in Vietnam: challenges and opportunities." Linguistics and Culture Review 5.S3 (2021): 1-15.

[2] Long Duong Phi, Binh Pham Nguyen Thanh, and Quang Tran Van "A Classification Method based on Concatenation Features for Diagnosing Skin Diseases" Journal of Information Hiding and Multimedia Signal Processing, Vol. 16, No. 1, pp. 401-413, March 2025.

[3] Marwa Hussien Mohamed, Mohamed Elkholy, Marwa. A.Marzouk "Innovative Machine Learning Approaches for Identifying Pre-diabetes in Patients" Journal of Information Hiding and Multimedia Signal Processing, Vol. 16, No. 1, pp. 365-378, March 2025.

[4] Tran-Anh, Dat, et al. "LmGa: Combining label mapping method with graph attention network for agricultural recognition." Knowledge and Information Systems 67.1 (2025): 789-810.

[5] Gharoun, Hassan, et al. "Meta-learning approaches for few-shot learning: A survey of recent advances." ACM Computing Surveys 56.12 (2024): 1-41.

[6] Li, Zewen, et al. "A survey of convolutional neural networks: analysis, applications, and prospects." IEEE transactions on neural networks and learning systems 33.12 (2021): 6999-7019.

[7] Cong, Shuang, and Yang Zhou. "A review of convolutional neural network architectures and their optimizations." Artificial Intelligence Review 56.3 (2023): 1905-1969.

[8] Derry, Alexander, Martin Krzywinski, and Naomi Altman. "Convolutional neural networks." Nature Methods 20.9 (2023): 1269-1270.

[9] Eliyas, Sherin. "Classifying Emotions and Engagement in Online Learning Based on Single Facial Recognition using CNN and Mobilenet-v2." 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0. IEEE, 2024.

[10] Sun, Bo, et al. "Student Class Behavior Dataset: a video dataset for recognizing, detecting, and captioning students' behaviors in classroom scenes." Neural Computing and Applications 33 (2021): 8335-8354.

[11] Xiao, Junfei, et al. "Learning from temporal gradient for semi-supervised action recognition." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

[12] Lu, Jin. "Research on classroom behavior analysis and optimization strategies based on convolutional neural network in smart learning environment." Frontiers in Educational Research 6.12 (2023).

[13] Liu, Tingting, et al. "Facial expression recognition method with multi-label distribution learning for non-verbal behavior understanding in the classroom." Infrared Physics & Technology 112 (2021): 103594.

[14] Aly, Mohammed. "Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model." Multimedia Tools and Applications (2024): 1-40.

[15] Wang, Huimin, and Min Chen. "Application of the Flipped Classroom Mode under Few-Shot Learning in the Teaching of Health Physical Education in Colleges and Universities." Computational Intelligence and Neuroscience 2022.1 (2022): 1465613.

[16] Hatano, Masashi, et al. "Multimodal cross-domain few-shot learning for egocentric action recognition." European Conference on Computer Vision. Springer, Cham, 2025.

[17] Samarasinghe, Sarinda, et al. "CDFSL-V: Cross-domain few-shot learning for videos." Proceedings of the IEEE/CVF international conference on computer vision. 2023.

[18] Song, Yisheng, et al. "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities." ACM Computing Surveys 55.13s (2023): 1-40.

[19] Qian, Yilei, et al. "Few-Shot Continual Learning for Activity Recognition in Classroom Surveillance Images." arXiv preprint arXiv:2409.03354 (2024).

[20] Yu, Zefang, et al. "From Raw Video to Pedagogical Insights: A Unified Framework for Student Behavior Analysis." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 21. 2024.

[21] Moratelli, Nicholas, et al. "Are Learnable Prompts the Right Way of Prompting? Adapting Vision-and-Language Models with Memory Optimization." IEEE Intelligent Systems (2024).

[22] Su, Jianpo, et al. "Few-shot domain-adaptive anomaly detection for cross-site brain images." IEEE Transactions on Pattern Analysis and Machine Intelligence 46.3 (2021): 1819-1835.

[23] Liu, Yu, et al. "Model behavior preserving for class-incremental learning." IEEE Transactions on Neural Networks and Learning Systems 34.10 (2022): 7529-7540.
[24] Wanyan, Yuyang, et al. "Active exploration of multimodal complementarity for few-shot action recognition." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
[25] Lee, Eugene, Cheng-Han Huang, and Chen-Yi Lee. "Few-shot and continual learning with attentive independent mechanisms." Proceedings of the IEEE/CVF international conference on computer vision. 2021.
[26] Yeleussinov, Arman, Talgat Islamgozhayev, and Zhanibek Kozhirbayev. "Creation and Annotation of a Handwritten Text Database for the Kazakh Language: Methodologies and Preliminary Results." 2024 International Conference on Electrical, Communication and Computer Engineering (ICECCE). IEEE, 2024.
[27] Wang, Qi-Wei, et al. "Few-shot class-incremental learning via training-free prototype calibration." Advances in Neural Information Processing Systems 36 (2024).
[28] Ye, Han-Jia, Hexiang Hu, and De-Chuan Zhan. "Learning adaptive classifiers synthesis for generalized few-shot learning." International Journal of Computer Vision 129.6 (2021): 1930-1953.
[29] Zhou, Da-Wei, et al. "Forward compatible few-shot class-incremental learning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
[30] Zhou, Da-Wei, et al. "Few-shot class-incremental learning by sampling multi-phase tasks." IEEE Transactions on Pattern Analysis and Machine Intelligence 45.11 (2022): 12816-12831.
[31] Ahmad, Touqeer, et al. "Few-shot class incremental learning leveraging self-supervised features." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.