



Universiteit
Leiden
The Netherlands

Open-world continual learning via knowledge transfer

Li, Y.

Citation

Li, Y. (2026, January 27). *Open-world continual learning via knowledge transfer*. Retrieved from <https://hdl.handle.net/1887/4287955>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4287955>

Note: To cite this publication please use the final published version (if applicable).

Bibliography

- [1] Parisi, G.I., Kemker, R., Part, J.L., Kanan, C., Wermter, S.: Continual lifelong learning with neural networks: A review. *Neural Networks* **113** (2019) 54–71
- [2] Ke, Z., Liu, B., Huang, X.: Continual learning of a mixed sequence of similar and dissimilar tasks. *Advances in Neural Information Processing Systems* **33** (2020) 18493–18504
- [3] Han, R., Ren, X., Peng, N.: Econet: Effective continual pretraining of language models for event temporal reasoning. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. (2021) 5367–5380
- [4] Kudithipudi, D., Aguilar-Simon, M., Babb, J., Bazhenov, M., Blackiston, D., Bongard, J., Brna, A.P., Chakravarthi Raja, S., Cheney, N., Clune, J., et al.: Biological underpinnings for lifelong learning machines. *Nature Machine Intelligence* **4** (2022) 196–210
- [5] Wang, L., Zhang, X., Su, H., Zhu, J.: A comprehensive survey of continual learning: theory, method and application. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024)
- [6] Van de Ven, G.M., Tolias, A.S.: Three scenarios for continual learning. *arXiv:1904.07734* (2019)
- [7] Van de Ven, G.M., Siegelmann, H.T., Tolias, A.S.: Brain-inspired replay for continual learning with artificial neural networks (2020)
- [8] Liu, B., Mazumder, S., Robertson, E., Grigsby, S.: AI autonomy: Self-initiated open-world continual learning and adaptation. *AI Magazine* **44** (2023) 185–199
- [9] Chen, Z., Liu, B.: *Lifelong machine learning*. Morgan & Claypool Publishers (2018)
- [10] Yang, G., Fini, E., Xu, D., Rota, P., Ding, M., Nabi, M., Alameda-Pineda, X., Ricci, E.: Uncertainty-aware contrastive distillation for incremental semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2022) 2567–2581
- [11] McClelland, J.L., McNaughton, B.L., O’Reilly, R.C.: Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* **102** (1995) 419
- [12] McCloskey, M., Cohen, N.J.: Catastrophic interference in connectionist networks: The sequential learning problem. In: *Psychology of learning and motivation*. Volume 24. Elsevier (1989) 109–165
- [13] Wang, H., Liu, B., Wang, S., Ma, N., Yang, Y.: Forward and backward knowledge transfer for sentiment classification. In: *Proceedings of the Asian Conference on Machine Learning*, PMLR (2019) 457–472
- [14] Zhou, D.W., Wang, Q.W., Qi, Z.H., Ye, H.J., Zhan, D.C., Liu, Z.: Class-incremental learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024)
- [15] Bendale, A., Boulton, T.: Towards open world recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2015) 1893–1902
- [16] Fei, G., Wang, S., Liu, B.: Learning cumulatively to become more knowledgeable. In: *Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. (2016) 1565–1574

- [17] Tulving, E., Kroll, N.: Novelty assessment in the brain and long-term memory encoding. *Psychonomic bulletin & review* **2** (1995) 387–390
- [18] Murty, V.P., Ballard, I.C., Macduffie, K.E., Krebs, R.M., Adcock, R.A.: Hippocampal networks habituate as novelty accumulates. *Learning & Memory* **20** (2013) 229–235
- [19] Kim, G., Xiao, C., Konishi, T., Ke, Z., Liu, B.: A theoretical study on solving continual learning. *Advances in Neural Information Processing Systems* **35** (2022) 5065–5079
- [20] Scheirer, W.J., de Rezende Rocha, A., Sapkota, A., Boulton, T.E.: Toward open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35** (2012) 1757–1772
- [21] Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al.: Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences* **114** (2017) 3521–3526
- [22] Zenke, F., Poole, B., Ganguli, S.: Continual learning through synaptic intelligence. In Precup, D., Teh, Y.W., eds.: *Proceedings of the 34th International Conference on Machine Learning*. Volume 70 of PMLR., PMLR (2017) 3987–3995
- [23] Kishida, I., Chen, H., Baba, M., Jin, J., Amma, A., Nakayama, H.: Object recognition with continual open set domain adaptation for home robot. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. (2021) 1517–1526
- [24] Truong, T.D., Nguyen, H.Q., Raj, B., Luu, K.: Fairness continual learning approach to semantic scene understanding in open-world environments. *Advances in Neural Information Processing Systems* **36** (2023) 65456–65467
- [25] Kim, G., Xiao, C., Konishi, T., Ke, Z., Liu, B.: Open-world continual learning: Unifying novelty detection and continual learning. *Artificial Intelligence* **338** (2025) 104237
- [26] Dhar, P., Singh, R.V., Peng, K.C., Wu, Z., Chellappa, R.: Learning without memorizing. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE (2019) 5138–5146
- [27] Vijayan, P., Bhat, P., Zonooz, B., Arani, E.: Trire: A multi-mechanism learning paradigm for continual knowledge retention and promotion. In: *Advances in Neural Information Processing Systems*. (2023)
- [28] Masana, M., Liu, X., Twardowski, B., Menta, M., Bagdanov, A.D., Van De Weijer, J.: Class-incremental learning: survey and performance evaluation on image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2022) 5513–5533
- [29] Lo, S.Y., Oza, P., Patel, V.M.: Adversarially robust one-class novelty detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2022) 4167–4179
- [30] Li, Y., Yang, X., Wang, H., Wang, X., Li, T.: Learning to prompt knowledge transfer for open-world continual learning. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Volume 38(12). (2024) 13700–13708
- [31] Mundt, M., Pliushch, I., Majumder, S., Hong, Y., Ramesh, V.: Unified probabilistic deep continual learning through generative replay and open set recognition. *Journal of Imaging* **8** (2022) 93
- [32] Wang, Z., Zhang, Z., Lee, C.Y., Zhang, H., Sun, R., Ren, X., Su, G., Perot, V., Dy, J., Pfister, T.: Learning to prompt for continual learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (2022) 139–149
- [33] Housby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Gesmundo, A., Attariyan, M., Gelly, S.: Parameter-efficient transfer learning for nlp. In: *International conference on machine learning*, PMLR (2019) 2790–2799
- [34] Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., Neubig, G.: Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys* **55** (2023) 1–35
- [35] Li, Y., Wang, X., Yang, X., Bonsangue, M., Zhang, J., Li, T.: Improving open-world continual learning under the constraints of scarce labeled data. In: *Proceedings of the 31st*

- ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (2025)
- [36] Fei, N., Gao, Y., Lu, Z., Xiang, T.: Z-score normalization, hubness, and few-shot learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. (2021) 142–151
 - [37] Trosten, D.J., Chakraborty, R., Løkse, S., Wickstrøm, K.K., Jenssen, R., Kampffmeyer, M.C.: Hubs and hyperspheres: Reducing hubness and improving transductive few-shot learning with hyperspherical embeddings. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2023) 7527–7536
 - [38] Li, Y., Lai, G., Yang, X., Li, Y., Bonsangue, M., Li, T.: Exploring open-world continual learning with knowns-unknowns knowledge transfer. arXiv:2502.20124 (2025)
 - [39] Li, Y., Yang, Y., Gao, Q., Yang, X.: Cross-regional fraud detection via continual learning (special program). In: Proceedings of the AAAI Conference on Artificial Intelligence. Volume 37. (2023) 16260–16261
 - [40] Li, Y., Yang, X., Gao, Q., Wang, H., Zhang, J., Li, T.: Cross-regional fraud detection via continual learning with knowledge transfer. IEEE Transactions on Knowledge and Data Engineering (2024)
 - [41] Shin, H., Lee, J.K., Kim, J., Kim, J.: Continual learning with deep generative replay. Advances in Neural Information Processing Systems **30** (2017)
 - [42] Rebuffi, S.A., Kolesnikov, A., Sperl, G., Lampert, C.H.: icarl: Incremental classifier and representation learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2017) 2001–2010
 - [43] Serra, J., Suris, D., Miron, M., Karatzoglou, A.: Overcoming catastrophic forgetting with hard attention to the task. In: International conference on machine learning, PMLR (2018) 4548–4557
 - [44] Lopez-Paz, D., Ranzato, M.: Gradient episodic memory for continual learning. Advances in Neural Information Processing Systems **30** (2017)
 - [45] Mirzadeh, S.I., Farajtabar, M., Pascanu, R., Ghasemzadeh, H.: Understanding the role of training regimes in continual learning. Advances in Neural Information Processing Systems **33** (2020) 7308–7320
 - [46] Ahn, H., Cha, S., Lee, D., Moon, T.: Uncertainty-based continual learning with adaptive regularization. Advances in Neural Information Processing Systems **32** (2019)
 - [47] Chaudhry, A., Ranzato, M., Rohrbach, M., Elhoseiny, M.: Efficient lifelong learning with a-gem. In: International Conference on Learning Representations. (2018)
 - [48] Yang, R., Wang, S., Zhang, H., Xu, S., Guo, Y., Ye, X., Hou, B., Jiao, L.: Knowledge decomposition and replay: A novel cross-modal image-text retrieval continual learning method. In: Proceedings of the 31st ACM International Conference on Multimedia, New York, NY, USA, Association for Computing Machinery (2023) 6510–6519
 - [49] Li, Z., Hoiem, D.: Learning without forgetting. IEEE Transactions on Pattern Analysis and Machine Intelligence **40** (2017) 2935–2947
 - [50] Wu, Y., Chen, Y., Wang, L., Ye, Y., Liu, Z., Guo, Y., Fu, Y.: Large scale incremental learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2019) 374–382
 - [51] Lai, G., Li, Y., Wang, X., Zhang, J., Li, T., Yang, X.: Order-robust class incremental learning: Graph-driven dynamic similarity grouping. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. (2025)
 - [52] Xu, J., Zhu, Z.: Reinforced continual learning. Advances in Neural Information Processing Systems **31** (2018)
 - [53] Aljundi, R., Chakravarty, P., Tuytelaars, T.: Expert gate: Lifelong learning with a network of experts. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2017) 3366–3375
 - [54] Mallya, A., Lazebnik, S.: Packnet: Adding multiple tasks to a single network by iterative

- pruning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2018) 7765–7773
- [55] Joseph, K.J., Khan, S., Khan, F.S., Balasubramanian, V.N.: Towards open world object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2021) 5830–5840
- [56] Dhamija, A.R., Günther, M., Boulton, T.: Reducing network agnostophobia. *Advances in Neural Information Processing Systems* **31** (2018)
- [57] Bendale, A., Boulton, T.E.: Towards open set deep networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2016) 1563–1572
- [58] Shu, L., Xu, H., Liu, B.: Doc: Deep open classification of text documents. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. (2017) 2911–2916
- [59] Dang, S., Cao, Z., Cui, Z., Pi, Y., Liu, N.: Open set incremental learning for automatic target recognition. *IEEE Transactions on Geoscience and Remote Sensing* **57** (2019) 4445–4456
- [60] Zhu, X., Zhang, R., He, B., Guo, Z., Zeng, Z., Qin, Z., Zhang, S., Gao, P.: Pointclip v2: Prompting clip and gpt for powerful 3d open-world learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. (2023) 2639–2650
- [61] Zhao, X., Ma, Y., Wang, D., Shen, Y., Qiao, Y., Liu, X.: Revisiting open world object detection. *IEEE Transactions on Circuits and Systems for Video Technology* (2023)
- [62] Liang, S., Li, Y., Srikant, R.: Enhancing the reliability of out-of-distribution image detection in neural networks. In: International Conference on Learning Representations. (2018)
- [63] Liu, W., Wang, X., Owens, J., Li, Y.: Energy-based out-of-distribution detection. *Advances in Neural Information Processing Systems* **33** (2020) 21464–21475
- [64] Lin, Z., Roy, S.D., Li, Y.: Mood: Multi-level out-of-distribution detection. In: Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition. (2021) 15313–15323
- [65] Djuricic, A., Bozanic, N., Ashok, A., Liu, R.: Extremely simple activation shaping for out-of-distribution detection. In: The Eleventh International Conference on Learning Representations. (2023)
- [66] Morteza, P., Li, Y.: Provable guarantees for understanding out-of-distribution detection. In: Proceedings of the AAAI Conference on Artificial Intelligence. Volume 36. (2022) 7831–7840
- [67] Hsu, Y.C., Shen, Y., Jin, H., Kira, Z.: Generalized odin: Detecting out-of-distribution image without learning from out-of-distribution data. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. (2020) 10951–10960
- [68] Lee, K., Lee, K., Min, K., Zhang, Y., Shin, J., Lee, H.: Hierarchical novelty detection for visual object recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018) 1034–1042
- [69] Huang, R., Li, Y.: Mos: Towards scaling out-of-distribution detection for large semantic space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2021) 8710–8719
- [70] Linderman, R., Zhang, J., Inkawich, N., Li, H., Chen, Y.: Fine-grain inference on out-of-distribution data with hierarchical classification. In: Conference on Lifelong Learning Agents, PMLR (2023) 162–183
- [71] Liu, X., Zhai, J.T., Bagdanov, A.D., Li, K., Cheng, M.M.: Task-adaptive saliency guidance for exemplar-free class incremental learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2024) 23954–23963
- [72] Wang, Y., Ding, Y., He, X., Fan, X., Lin, C., Li, F., Wang, T., Luo, Z., Luo, J.: Novelty detection and online learning for chunk data streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43** (2020) 2400–2412
- [73] Yu, L., Liu, X., Van de Weijer, J.: Self-training for class-incremental semantic segmentation. *IEEE Transactions on Neural Networks and Learning Systems* **34** (2022) 9116–9127

- [74] Chan, R., Rottmann, M., Gottschalk, H.: Entropy maximization and meta classification for out-of-distribution detection in semantic segmentation. In: *Proceedings of the IEEE/CVF international conference on computer vision*. (2021) 5128–5137
- [75] Mazumder, S., Liu, B.: *Lifelong and Continual Learning Dialogue Systems*. Springer Nature (2024)
- [76] Schlachter, P., Liao, Y., Yang, B.: Deep open set recognition using dynamic intra-class splitting. *SN Computer Science* **1** (2020) 1–12
- [77] Aljundi, R., Reino, D.O., Chumerin, N., Turner, R.E.: Continual novelty detection. In: *Conference on Lifelong Learning Agents*, PMLR (2022) 1004–1025
- [78] Oza, P., Patel, V.M.: C2ae: Class conditioned auto-encoder for open-set recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2019) 2307–2316
- [79] Lee, K., Lee, K., Lee, H., Shin, J.: A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In: *Advances in Neural Information Processing Systems*. Volume 31. (2018) 7167–7177
- [80] Hendrycks, D., Gimpel, K.: A baseline for detecting misclassified and out-of-distribution examples in neural networks. *arXiv:1610.02136* (2016)
- [81] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv:2010.11929* (2020)
- [82] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. In: *Advances in Neural Information Processing Systems*. (2017) 5998–6008
- [83] Wang, Z., Zhang, Z., Ebrahimi, S., Sun, R., Zhang, H., Lee, C.Y., Ren, X., Su, G., Perot, V., Dy, J., et al.: Dualprompt: Complementary prompting for rehearsal-free continual learning. In: *European conference on computer vision*, Springer (2022) 631–648
- [84] Wang, L., Xie, J., Zhang, X., Huang, M., Su, H., Zhu, J.: Hierarchical decomposition of prompt-based continual learning: Rethinking obscured sub-optimality. *Advances in Neural Information Processing Systems* **36** (2023) 69054–69076
- [85] Lester, B., Al-Rfou, R., Constant, N.: The power of scale for parameter-efficient prompt tuning. *arXiv:2104.08691* (2021)
- [86] Li, X.L., Liang, P.: Prefix-tuning: Optimizing continuous prompts for generation. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. (2021) 4582–4597
- [87] Zhou, D.W., Sun, H.L., Ning, J., Ye, H.J., Zhan, D.C.: Continual learning with pre-trained models: a survey. In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*. (2024) 8363–8371
- [88] Smith, J.S., Karlinsky, L., Gutta, V., Cascante-Bonilla, P., Kim, D., Arbelle, A., Panda, R., Feris, R., Kira, Z.: Coda-prompt: Continual decomposed attention-based prompting for rehearsal-free continual learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (2023) 11909–11919
- [89] Jung, D., Han, D., Bang, J., Song, H.: Generating instance-level prompts for rehearsal-free continual learning. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (2023) 11847–11857
- [90] Wang, Y., Huang, Z., Hong, X.: S-prompts learning with pre-trained transformers: An occam’s razor for domain incremental learning. *Advances in Neural Information Processing Systems* **35** (2022) 5682–5695
- [91] Zhou, D.W., Cai, Z.W., Ye, H.J., Zhan, D.C., Liu, Z.: Revisiting class-incremental learning with pre-trained models: Generalizability and adaptivity are all you need. *International Journal of Computer Vision* **133** (2025) 1012–1032

- [92] Zhou, D.W., Zhang, Y., Wang, Y., Ning, J., Ye, H.J., Zhan, D.C., Liu, Z.: Learning without forgetting for vision-language models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2025)
- [93] Zhou, D.W., Sun, H.L., Ye, H.J., Zhan, D.C.: Expandable subspace ensemble for pre-trained model-based class-incremental learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2024) 23554–23564
- [94] Chen, S., Ge, C., Tong, Z., Wang, J., Song, Y., Wang, J., Luo, P.: Adaptformer: Adapting vision transformers for scalable visual recognition. *Advances in Neural Information Processing Systems* **35** (2022) 16664–16678
- [95] Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I.: Learning transferable visual models from natural language supervision. In: *Proceedings of the 38th International Conference on Machine Learning*. Volume 139 of *Proceedings of Machine Learning Research*. (2021) 8748–8763
- [96] Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., Tuytelaars, T.: Memory aware synapses: Learning what (not) to forget. In: *Proceedings of the European conference on computer vision (ECCV)*. (2018) 139–154
- [97] Quanz, B., Huan, J., Mishra, M.: Knowledge transfer with low-quality data: A feature extraction issue. *IEEE Transactions on Knowledge and Data Engineering* **24** (2012) 1789–1802
- [98] Krizhevsky, A., Hinton, G., et al.: Learning multiple layers of features from tiny images. Master’s thesis, University of Toronto (2009)
- [99] Ebrahimi, S., Meier, F., Calandra, R., Darrell, T., Rohrbach, M.: Adversarial continual learning. In: *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI 16*, Springer (2020) 386–402
- [100] Yu, Q., Aizawa, K.: Unsupervised out-of-distribution detection by maximum classifier discrepancy. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (2019) 9518–9526
- [101] Basart, S., Mantas, M., Mohammadreza, M., Jacob, S., Dawn, S.: Scaling out-of-distribution detection for real-world settings. In: *International Conference on Machine Learning*. (2022)
- [102] Buzzega, P., Boschini, M., Porrello, A., Abati, D., Calderara, S.: Dark experience for general continual learning: a strong, simple baseline. *Advances in Neural Information Processing Systems* **33** (2020) 15920–15930
- [103] Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. *Journal of Machine Learning Research* **9** (2008)
- [104] McInnes, L., Healy, J., Saul, N., Großberger, L.: Umap: Uniform manifold approximation and projection. *Journal of Open Source Software* **3** (2018)
- [105] Ahmad, T., Dhamija, A.R., Cruz, S., Rabinowitz, R., Li, C., Jafarzadeh, M., Boulton, T.E.: Few-shot class incremental learning leveraging self-supervised features. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2022) 3900–3910
- [106] Yoon, J., Madjid, S., Hwang, S.J., Yoo, C.D., et al.: On the soft-subnetwork for few-shot class incremental learning. In: *International Conference on Learning Representations, International Conference on Learning Representations* (2023)
- [107] Qin, C., Joty, S.: Lfpt5: A unified framework for lifelong few-shot language learning based on prompt tuning of t5. In: *International Conference on Learning Representations*. (2021)
- [108] Tao, X., Hong, X., Chang, X., Dong, S., Wei, X., Gong, Y.: Few-shot class-incremental learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2020) 12183–12192
- [109] Wah, C., Branson, S., Welinder, P., Perona, P., Belongie, S.: The caltech-ucsd birds-200-2011 dataset (2011)
- [110] Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al.: Matching networks for one shot

- learning. *Advances in Neural Information Processing Systems* **29** (2016)
- [111] Hendrycks, D., Basart, S., Mazeika, M., Zou, A., Kwon, J., Mostajabi, M., Steinhardt, J., Song, D.: Scaling out-of-distribution detection for real-world settings. In: *International Conference on Machine Learning*, PMLR (2022) 8759–8773
 - [112] Sehwag, V., Chiang, M., Mittal, P.: Ssd: A unified framework for self-supervised. In: *International Conference on Machine Learning*. (2021)
 - [113] Wang, H., Li, Z., Feng, L., Zhang, W.: Vim: Out-of-distribution with virtual-logit matching. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2022) 4921–4930
 - [114] Sun, Y., Ming, Y., Zhu, X., Li, Y.: Out-of-distribution detection with deep nearest neighbors. In: *International Conference on Machine Learning*, PMLR (2022) 20827–20840
 - [115] Park, J., Jung, Y.G., Teoh, A.B.J.: Nearest neighbor guidance for out-of-distribution detection. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (2023) 1686–1695
 - [116] Zhang, C., Song, N., Lin, G., Zheng, Y., Pan, P., Xu, Y.: Few-shot incremental learning with continually evolved classifiers. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2021) 12455–12464
 - [117] Chi, Z., Gu, L., Liu, H., Wang, Y., Yu, Y., Tang, J.: Metafscl: A meta-learning approach for few-shot class incremental learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2022) 14166–14175
 - [118] Peng, C., Zhao, K., Wang, T., Li, M., Lovell, B.C.: Few-shot class-incremental learning from an open-set perspective. In: *European Conference on Computer Vision*, Springer (2022) 382–397
 - [119] Zhou, D.W., Ye, H.J., Ma, L., Xie, D., Pu, S., Zhan, D.C.: Few-shot class-incremental learning by sampling multi-phase tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022)
 - [120] Yang, Y., Yuan, H., Li, X., Lin, Z., Torr, P., Tao, D.: Neural collapse inspired feature-classifier alignment for few-shot class-incremental learning. In: *The Eleventh International Conference on Learning Representations*. (2022)
 - [121] Ji, Z., Hou, Z., Liu, X., Pang, Y., Li, X.: Memorizing complementation network for few-shot class-incremental learning. *IEEE Transactions on Image Processing* **32** (2023) 937–948
 - [122] Wang, X., Liu, Y., Fan, J., Wen, W., Xue, H., Zhu, W.: Continual few-shot learning with transformer adaptation and knowledge regularization. In: *Proceedings of the ACM Web Conference*. Volume 2023. (2023)
 - [123] Li, J., Bai, Y., Lou, Y., Linghu, X., He, J., Xu, S., Bai, T.: Memory-based label-text tuning for few-shot class-incremental learning. *arXiv:2207.01036* (2022)
 - [124] Agarwal, A., Banerjee, B., Cuzzolin, F., Chaudhuri, S.: Semantics-driven generative replay for few-shot class incremental learning. In: *Proceedings of the 30th ACM International Conference on Multimedia*. (2022) 5246–5254
 - [125] D’Alessandro, M., Alonso, A., Calabrés, E., Galar, M.: Multimodal parameter-efficient few-shot class incremental learning. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (2023) 3393–3403
 - [126] Zhu, P., Zhang, W., Wang, Y., Hu, Q.: Multi-granularity inter-class correlation based contrastive learning for open set recognition. *Int. J. Softw. Informatics* **12** (2022) 157–175
 - [127] Huang, H., Wang, Y., Hu, Q., Cheng, M.M.: Class-specific semantic reconstruction for open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2022) 4214–4228
 - [128] Zhou, Y., Liu, P., Qiu, X.: Knn-contrastive learning for out-of-domain intent classification. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. (2022) 5129–5141
 - [129] van de Ven, G.M., Tolias, A.S.: Three continual learning scenarios. In: *Advances in Neural*

- Information Processing Systems. Volume 1(9). (2018)
- [130] McDonnell, M.D., Gong, D., Parvaneh, A., Abbasnejad, E., van den Hengel, A.: Ranpac: Random projections and pre-trained models for continual learning. *Advances in Neural Information Processing Systems* **36** (2024)
 - [131] Zhou, D.W., Ye, H.J., Zhan, D.C., Liu, Z.: Revisiting class-incremental learning with pre-trained models: Generalizability and adaptivity are all you need. *arXiv:2303.07338* (2023)
 - [132] Fang, Z., Lu, J., Liu, A., Liu, F., Zhang, G.: Learning bounds for open-set learning. In: *International conference on machine learning*, PMLR (2021) 3122–3132
 - [133] Shi, H., Wang, H.: A unified approach to domain incremental learning with memory: Theory and algorithm. *Advances in Neural Information Processing Systems* **36** (2024)
 - [134] Anthony, M., Bartlett, P.L., Bartlett, P.L., et al.: *Neural network learning: Theoretical foundations*. Volume 9. cambridge university press Cambridge (1999)
 - [135] Lian, D., Zhou, D., Feng, J., Wang, X.: Scaling & shifting your features: A new baseline for efficient model tuning. *Advances in Neural Information Processing Systems* **35** (2022) 109–123
 - [136] Jia, M., Tang, L., Chen, B.C., Cardie, C., Belongie, S., Hariharan, B., Lim, S.N.: Visual prompt tuning. In: *European Conference on Computer Vision*, Springer (2022) 709–727
 - [137] Chen, C.P.: A rapid supervised learning neural network for function interpolation and approximation. *IEEE Transactions on Neural Networks* **7** (1996) 1220–1230
 - [138] Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. *Neurocomputing* **70** (2006) 489–501
 - [139] Lomonaco, V., Maltoni, D.: Core50: a new dataset and benchmark for continuous object recognition. In: *Conference on robot learning*, PMLR (2017) 17–26
 - [140] Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J.: Lof: identifying density-based local outliers. In: *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*. (2000) 93–104
 - [141] Shi, H., Tayebi, M.A., Pei, J., Cao, J.: Cost-sensitive learning for medical insurance fraud detection with temporal information. *IEEE Transactions on Knowledge and Data Engineering* (2023) 1–14
 - [142] Kerpici, M., Ozkan, H., Kozat, S.S.: Online anomaly detection with bandwidth optimized hierarchical kernel density estimators. *IEEE Transactions on Neural Networks and Learning Systems* **32** (2021) 4253–4266
 - [143] Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q., Xiong, H., Akoglu, L.: A comprehensive survey on graph anomaly detection with deep learning. *IEEE Transactions on Knowledge and Data Engineering* (2021)
 - [144] Li, Z., Liu, G., Jiang, C.: Deep representation learning with full center loss for credit card fraud detection. *IEEE Transactions on Computational Social Systems* **7** (2020) 569–579
 - [145] Fanai, H., Abbasimehr, H.: A novel combined approach based on deep autoencoder and deep classifiers for credit card fraud detection. *Expert Systems with Applications* (2023) 119562
 - [146] Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. In: *NIPS 2014 Workshop on Deep Learning*, December 2014. (2014)
 - [147] Benchaji, I., Douzi, S., El Ouahidi, B., Jaafari, J.: Enhanced credit card fraud detection based on attention mechanism and lstm deep model. *Journal of Big Data* **8** (2021) 1–21
 - [148] Xie, Y., Liu, G., Yan, C., Jiang, C., Zhou, M.: Time-aware attention-based gated network for credit card fraud detection by extracting transactional behaviors. *IEEE Transactions on Computational Social Systems* (2022)
 - [149] Ergen, T., Kozat, S.S.: Unsupervised anomaly detection with lstm neural networks. *IEEE Transactions on Neural Networks and Learning Systems* **31** (2019) 3127–3141
 - [150] Wu, W., He, L., Lin, W., Su, Y., Cui, Y., Maple, C., Jarvis, S.: Developing an unsupervised

- real-time anomaly detection scheme for time series with multi-seasonality. *IEEE Transactions on Knowledge and Data Engineering* **34** (2020) 4147–4160
- [151] Wang, D., Lin, J., Cui, P., Jia, Q., Wang, Z., Fang, Y., Yu, Q., Zhou, J., Yang, S., Qi, Y.: A semi-supervised graph attentive network for financial fraud detection. In: *ICDM, IEEE* (2019) 598–607
 - [152] Kipf, T., Welling, M.: Semi-supervised classification with graph convolutional networks. In: *International Conference on Learning Representations*. (2016)
 - [153] Hamilton, W.L., Ying, R., Leskovec, J.: Inductive representation learning on large graphs. In: *NIPS*. (2017) 1025–1035
 - [154] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph attention networks. In: *International Conference on Learning Representations*. (2018)
 - [155] Peng, L., Mo, Y., Xu, J., Shen, J., Shi, X., Li, X., Shen, H.T., Zhu, X.: Grlc: Graph representation learning with constraints. *IEEE Transactions on Neural Networks and Learning Systems* (2023)
 - [156] Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., Yu, P.S.: Heterogeneous graph attention network. In: *The world wide web conference*. (2019) 2022–2032
 - [157] Cheng, D., Wang, X., Zhang, Y., Zhang, L.: Graph neural network for fraud detection via spatial-temporal attention. *IEEE Transactions on Knowledge and Data Engineering* (2020)
 - [158] Hei, Y., Yang, R., Peng, H., Wang, L., Xu, X., Liu, J., Liu, H., Xu, J., Sun, L.: Hawk: Rapid android malware detection through heterogeneous graph attention networks. *IEEE Transactions on Neural Networks and Learning Systems* (2021)
 - [159] Wang, C., Dou, Y., Chen, M., Chen, J., Liu, Z., Philip, S.Y.: Deep fraud detection on non-attributed graph. In: *Big Data, IEEE* (2021) 5470–5473
 - [160] Chen, Z., Sun, A.: Anomaly detection on dynamic bipartite graph with burstiness. In: *2020 IEEE International Conference on Data Mining (ICDM), IEEE* (2020) 966–971
 - [161] Gan, J., Hu, R., Mo, Y., Kang, Z., Peng, L., Zhu, Y., Zhu, X.: Multigraph fusion for dynamic graph convolutional network. *IEEE Transactions on Neural Networks and Learning Systems* **35** (2022) 196–207
 - [162] Zhang, X., Song, D., Tao, D.: Hierarchical prototype networks for continual graph representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45** (2023) 4622–4636
 - [163] Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., Yu, P.S.: Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In: *CIKM*. (2020) 315–324
 - [164] Lin, H., Liu, G., Wu, J., Zuo, Y., Wan, X., Li, H.: Fraud detection in dynamic interaction network. *IEEE Transactions on Knowledge and Data Engineering* **32** (2020) 1936–1950
 - [165] Wang, J., Song, G., Wu, Y., Wang, L.: Streaming graph neural networks via continual learning. In: *Proceedings of the 29th ACM international conference on information & knowledge management*. (2020) 1515–1524
 - [166] Zhou, F., Cao, C.: Overcoming catastrophic forgetting in graph neural networks with experience replay. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Number 5 (2021) 4714–4722
 - [167] Zhang, P., Yan, Y., Li, C., Wang, S., Xie, X., Song, G., Kim, S.: Continual learning on dynamic graphs via parameter isolation. In: *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. (2023) 601–611
 - [168] Zhu, Y., Ma, J., Yuan, C., Zhu, X.: Interpretable learning based dynamic graph convolutional networks for alzheimer’s disease analysis. *Information Fusion* **77** (2022) 53–61
 - [169] Liu, Z., Chen, C., Yang, X., Zhou, J., Li, X., Song, L.: Heterogeneous graph neural networks for malicious account detection. In: *Proceedings of the 27th ACM international conference on information and knowledge management*. (2018) 2077–2085
 - [170] Mo, Y., Chen, Y., Lei, Y., Peng, L., Shi, X., Yuan, C., Zhu, X.: Multiplex graph represen-

- tation learning via dual correlation reduction. *IEEE Transactions on Knowledge and Data Engineering* **35** (2023) 12814–12827
- [171] Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: online learning of social representations. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM (2014) 701–710
- [172] Zhang, J., Zhang, J., Ghosh, S., Li, D., Zhu, J., Zhang, H., Wang, Y.: Regularize, expand and compress: Nonexpansive continual learning. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. (2020) 854–862
- [173] Tang, B., Matteson, D.S.: Graph-based continual learning. In: *International Conference on Learning Representations*. (2020)

List of Figures

1.1	A conceptual framework of continual learning. a. Continual learning requires adapting to incremental tasks with dynamic data distributions. b. A desirable solution should ensure an appropriate trade-off between stability (orange arrow) and plasticity (green arrow), as well as a generalizability to intra-task (blue arrow) and inter-task (red arrow) distribution differences.	2
1.2	An illustration of the open-world continual learning problem and comparison between continual learning. In open-world continual learning, the test set of each task may include samples from classes unseen during training, deviating from the closed-world assumption of traditional continual learning. As shown in the figure, this setting imposes two key challenges: (1) preserving knowledge from previously learned tasks to mitigate catastrophic forgetting, and (2) performing robust open-set recognition to correctly identify and reject instances from unknown classes. While continual learning models tend to misclassify open samples as one of the known classes (e.g., the green pentagon and orange square), open-world continual learning models are explicitly designed to distinguish between known and unknown classes, thereby enabling more reliable and adaptive decision-making in dynamic environments.	4
3.1	In contrast to rehearsal-based continual learning approaches that sequentially adapt full or partial model parameters to new tasks and rely on a rehearsal buffer to alleviate catastrophic forgetting, prompt-based methods adopt a unified backbone model augmented by learned prompts [32, 83]. Task-specific information is encapsulated within the prompts, thereby eliminating the necessity of a rehearsal buffer. Furthermore, different methods have different ways of selecting and updating prompts dynamically, usually enabling task-agnostic inference at test time without explicit task identity.	32
3.2	Different kinds of prompt selection, including key-value matching, shared and task-specific retrieval, attention-based combination, and instance-specific prompt generation.	33

4.1	A simple illustration of OWCL.	37
4.2	Distributions of the outputted logits scores of test samples, using MAS (a typical continual learning baseline) and the proposed <i>Pro-KT</i> on Task 1 and Task 2, respectively. The x -axis is the unscaled class-maximum logits of samples, and the y -axis shows the number of samples.	38
4.3	Overview of the proposed <i>Pro-KT</i> framework. Left: Training phase. Right: Testing phase. During training, input samples are first projected using a query function. <i>Pro-KT</i> learns a new set of prompts and stores them in a prompt bank. Each sample is then enriched with selected prompts, and the resulting prompt-enhanced embeddings are processed by a frozen pre-trained backbone. The final representations are passed to a trainable classifier. Additionally, an adaptive threshold for a task-aware open-set boundary is learned based on the softmax entropy of these enhanced representations. In the <i>testing phase</i> , test samples are similarly projected and matched with prompts from the prompt bank using a sample-wise selection strategy, enabling transfer of both task-specific and generic knowledge. The enriched embeddings are fed into the unified classifier to obtain unscaled logit scores. Finally, <i>Pro-KT</i> classifies each sample as either <i>[known]</i> or <i>[unknown]</i> , using the learned adaptive threshold.	42
4.4	(a) and (b): Variation curve of A_N on Split CIFAR-100 and 5-Datasets, the x -axis indicates the task IDs. (c) and (d): Variation curve of F_N on Split CIFAR-100 and 5-Datasets, the x -axis indicates the task IDs.	53
4.5	Visualization of the prompt bank via T-SNE and UMAP after training on all tasks (Split CIFAR-100 dataset).	55
4.6	Left and middle are performance heatmaps (A_N) showing interaction effects between prompt length (L_P , y -axis) and selection size (K , x -axis) on both datasets, with fixed $M = 25$. Right is A_N variation across prompt counts (M) with optimal $L_P = 5$ and $K = 3$	56
5.1	Open-world few-shot continual learning framework overview. The model incrementally learns new tasks (task 1, ..., task n , etc.) with few-shot class examples (class 3, ..., class i , etc.) while simultaneously detecting unknown samples during testing and overcoming severe forgetting and overfitting challenges inherent in data-scarce continual learning scenarios.	60

5.2	The overall <i>PEAK</i> model consists of three key components: (1) ITA (Yellow): enhancing the samples by matching them with appropriate additional tokens; (2) MOB (Purple): constructing compact decision boundaries of knowns by margin-based loss function for open detection; (3) AKS (Red): incorporating knowledge learned from both knowns and unknowns, and facilitating the transition from unknowns to knowns.	64
5.3	Visualizations of model performance on known classification.	76
5.4	(a): ACC_N w.r.t token length L_P and additional token size K , given $l = 25$. (b): ACC_N w.r.t. l (i.e., the total number of additional tokens of each task) with $L_P = 5$ and $K = 5$ (take CUB200 for illustration).	78
5.5	(a) and (b): AUC_N and FPR_N w.r.t constraint governing deviations σ and margin m (CUB200 dataset).	78
5.6	Visual comparison of test sample representations: (left) baseline Transformer backbone versus (right) our proposed <i>PEAK</i> framework.	79
6.1	A toy example of the KIRO scenario in OWCL. During training Task 1 and Task t , Robots A and B learn <i>Golden Retrievers</i> and <i>Labradors</i> , both labeled as $[dog]$. In testing Task 1, both robots correctly classified $[dog]$ and $[bus]$ samples while identifying the unknown wolf without errors. However, in testing Task t , due to the emergence of <i>Labrador</i> , Robot A experienced distributional shifts when recognizing <i>Golden Retrievers</i> . This also led to the misclassification of <i>Wolves</i> as <i>Labradors</i> , likely due to visual similarity. In contrast, Robot B accurately classifies all known samples and consistently detects the wolf as an open sample via knowledge transfer.	82
6.2	From task t to task $t + 1$ in the KIRO scenario, we illustrate the embedding space using different baselines. Red circles and red triangles denote unknown samples encountered in tasks t and $t - 1$, respectively. Orange, blue, and green circles represent labeled training data learned during tasks $t - 1$, t , and $t + 1$, respectively. Dashed lines denote the decision boundaries inferred by the model during open-world continual learning. (a) The original embedding distributions for each task learned by the classical continual learning baseline EWC. (b) A current competitive open-world continual learning approach enforces tighter embedding cohesion for each task. (c) Our proposed method jointly addresses open detection and incrementally reduces known sample classification errors through knowledge transfer. [Best view in color]	84
6.3	Unscaled classification scores on the CINO scenario.	91
6.4	Unscaled classification scores on the CIRO scenario.	91
6.5	Unscaled classification scores on the KINO Scenario.	92
6.6	Unscaled classification scores on the KIRO Scenario.	93

6.7	Visualizations of main results.	109
6.8	Performance with varying NRP size.	111
6.9	The divergence between the actual embedding distributions and the <i>HoliTrans</i> 's embedding distributions with distribution-aware prototypes (The horizontal axis is the values of r , the vertical axis is ACC_t , and the dashed line indicates the optimal value of r). Orange represents the actual ones, blue represents ours.	112
7.1	Bottom: An illustration of cross-regional fraud detection. The orange arrows indicate a financial business expansion from Region 1 to Region 5, where the data characteristics vary across different regions with different patterns. Our proposed method aims to incrementally learn a unified fraud detection model that can perform well in each task through knowledge transfer. Top: The histograms in different colors denote each corresponding region's fraud detection performance (ACC) with training from scratch. The performance drops significantly after a model is trained using data from Region 2. Similarly, after training a model with data from Region 3, the performance of Region 1 and Region 2 fell further, respectively. . . .	117
7.2	Heat maps of temporal slices. On the horizontal axis is the amount of fraudulent transactions, and on the vertical axis are time slices. The darker the color, the more fraudulent transactions occurred during the current period.	124
7.3	The overall framework of the proposed CCL. Given a sequence of cross-regional fraud detection tasks, the proposed prototype-based replay method for knowledge transfer and the forgetting prevention approach are illustrated. Take the learning process of \mathcal{G}_2 as an example. In stage 1 of knowledge transfer, we randomly choose some target nodes from \mathcal{G}_1 and generate a prototype \mathbf{c}^1 . In stage 2, a Gaussian prior is added to perturb the prototype. In stage 3, nodes are sorted according to the distance from the perturbed prototype \mathbf{c}^1 , and we select the replay nodes according to the replay ratio r . Then, all the selected nodes and their neighbors are stored in the experience buffer \mathbb{B}_1 . Next, the replay buffer is appended to \mathcal{G}_2 for the training, with parameter smoothing for forgetting prevention. Finally, the final embedding of \mathbf{Z}_t is obtained by the GNN-based backbone accordingly.	125
7.4	Left: Variation curves of recall. Middle: Variation curves of AUC. Right: Variation curves of F1.	134
7.5	Sensitivity analysis of hyperparameters. We vary the hidden size $ \mathbf{h} $ as $\{8, 16, 24, 32\}$, the smoothing term λ as $\{1.5, 15, 150, 1500\}$, the smoothing term γ as $\{0.1, 0.01, 0.001, 0.0001\}$, the replay ratio r as $\{0.01, 0.05, 0.1, 0.15\}$	138

- 7.6 Node-type importance analysis via SHAP values. Each dot represents a sample's SHAP value for a specific node type, indicating the feature's impact on the model's output. Higher SHAP values (toward the right) denote stronger positive contributions to fraud detection, while lower values (toward the left) denote negative influence. Node types with greater vertical dispersion and right-shifted distributions are more influential in the model's decision-making. The color of each point corresponds to the masking level of that node type, where **red** indicates a higher feature value (i.e., lower masking), and **blue** indicates a lower feature value (i.e., more heavily masked). It can be observed that **Customer** nodes are the most impactful, followed by **Time Slice** and **Merchant**, underscoring the importance of temporal and behavioral semantics in cross-regional fraud detection. . . . 140

List of Tables

1.1	Overview of different OWCL scenarios.	5
2.1	Comparison of continual learning, open-set learning, and open-world continual learning.	21
3.1	Overview of the three CL scenarios.	25
3.2	Comparison between OSL and OOD Detection.	27
3.3	Notations used in the OWCL problem formulation.	29
4.1	Notation used in this chapter.	41
4.2	Evaluation results of unknown detection performance (%) across 10 sequential tasks (10 classes each) on Split CIFAR-100 and 5 tasks on 5-Datasets.	51
4.3	Classification accuracy (%) on known classes for Split CIFAR-100 and 5-Datasets benchmarks. Experiments evaluate continual learning performance across 10 sequential tasks (10 classes each) on Split CIFAR-100, and 5 sequential tasks for 5 datasets. We used the ResNet32 and ViT backbones in the experiments.	52
4.4	Ablation study on the variants of <i>Pro-KT</i>	54
5.1	Notations and explanations.	63
5.2	Results (%) regarding unknown detection. We report the results over 10 tasks for CUB200 (10-way 5-shot) and 8 tasks for MiniImageNet (5-way 5-shot).	74
5.3	Averaged classification accuracy (%) on CUB200 (10-way 5-shot). . .	75
5.4	Averaged classification accuracy (%) on MiniImageNet (5-way 5-shot). .	76
5.5	Ablation studies.	77
5.6	Performance under different γ configurations on the CUB200 dataset. .	79
6.1	Overview of the four OWCL scenarios.	83
6.2	Notations and explanations.	87
6.3	Dataset description.	107
6.4	Main Results. We evidence first , <u>second</u> , and <u>third</u> performances. . .	108
6.5	Main Results. We format first , <u>second</u> , and <u>third</u> performances. . .	109
6.6	Ablation Study.	110

6.7	Performance on Different NRP Size for CIRO (on Split-CIFAR100 dataset) and KIRO (on Open-CORE50 dataset).	111
7.1	Notations and explanations.	122
7.2	The statistics of datasets for each region.	133
7.3	Recall scores (%) across regions \mathcal{R}_1 to \mathcal{R}_5	134
7.4	AUC scores (%) across regions \mathcal{R}_1 to \mathcal{R}_5	135
7.5	F1 scores (%) across regions \mathcal{R}_1 to \mathcal{R}_5	135
7.6	Ablation study (Recall %) across regions \mathcal{R}_1 to \mathcal{R}_5	136
7.7	Ablation study (AUC %) across regions \mathcal{R}_1 to \mathcal{R}_5	136
7.8	Ablation study (F1 %) across regions \mathcal{R}_1 to \mathcal{R}_5	137