Title

Accelerating ecosystem monitoring through computer vision with deep metric learning

Author names and affiliations

Yurika Oba[1], Hideyuki Doi[1]

1 Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan.

Corresponding author

Yurika Oba

E-mail address: oba.yurika.8a@kyoto-u.ac.jp

Abstract

Feature extraction from environmental observation data based on deep learning models has made significant progress. However, the current methods may not be optimal because of the increasing volume of data, complexity of data characteristics, and labeled data limitations. In this study, we focused on deep metric learning as a new application for environmental observation data to overcome these challenges. The extraction of features such as patterns and changes from large and complex environmental observation data using a deep metric learning approach may provide new opportunities for monitoring ecosystems experiencing unprecedented loads from climate change and human activities. We expect that deep metric learning will be a powerful tool for various ecosystem monitoring systems, from remote sensing of wide-area data to ecological data obtained through field surveys.

Keywords

ecosystem monitoring, deep metric learning, few-shot learning, zero-shot learning, remote sensing, field observation data

Introduction

Understanding environmental changes through continuous and long-term monitoring is an important process that can provide a warning of biodiversity loss and adverse environmental changes (Pasher et al., 2014). Monitoring and assessing the impacts of ongoing conservation programs are important for sustainable and effective ecosystem management. Monitoring through images and videos, which observes in a non-contact or non-destructive manner can reduce the stress on ecosystems and living organisms (Marini et al., 2018); they are one of the main methods of ecosystem monitoring. Recently, computer vision-based ecosystem monitoring has significantly progressed with the development of deep learning (Franceschini et al., 2023; Jamali et al., 2022; Zhang et al., 2020). Recent developments in observation equipment technology have enabled data accumulation at higher frequencies and resolutions over a wider area. Because of the great compatibility between these backgrounds and the technology provided by deep learning, ecosystem monitoring employing deep learning models will be a promising field in the future. However, with the development of various methods, challenges specific to ecosystem monitoring have emerged. It is often difficult to prepare large labeled datasets, and there may be cases in which the observed data characteristics belong to untrained classification classes. To address these challenges, new methods that use deep metric learning, which is a combination of deep learning and metric learning, have been proposed.

In general, metric learning projects a smaller distance between samples belonging to the same class, while increasing the distance between samples belonging to different classes. Deep metric learning adopts the metric learning concept and learns the distance or similarity between data (Kaya & Bilge, 2019; Lu et al., 2017a) (Fig. 1 and 2). This advancement provides a solution to the challenges associated with few-shot and zero-shot learning methods. Few-shot learning is a method designed to handle various tasks by efficiently learning from a few data sets (Li, X. X. et al., 2023), and zero-shot learning is a learning method that handles unknown classes not present in the training data by adding supplemental information to existing models (Zabihzadeh & Masoudifar, 2023). With the deep learning approach, the classification accuracy tends to depend on the amount of labeled training data, and it has been difficult to achieve sufficient effectiveness for small amounts of data or unknown classes. In contrast, deep metric learning techniques, which can be applied to small datasets or datasets with unknown classes, are often used for face recognition and verification (Golwalkar & Mehendale, 2022; Hu et al., 2014; Lu et al., 2017b) and anomaly detection (Kosuge et al., 2023). Recently, it has also been applied in the field of natural science such as remote sensing, agriculture, and wildlife identification. In this study, we introduce some examples of research that have used deep metric learning and discuss the potential of this technology in the field of ecosystem monitoring. Such approaches towards increasingly large and complex data streams may lead to solutions for more effective utilization of spatiotemporal imbalances in ecosystem monitoring data.

Deep metric learning application in remote sensing

Improvements in Earth observation technology have provided the opportunity to access vast amounts of data (Vance et al., 2024). At the same time, there has been growing interest in developing techniques for processing, analyzing, searching, and clustering the obtained data. Methods based on deep learning models have significantly contributed to the performance of these technologies. However, current deep-learning models require preparing large labeled datasets for training, and the significant cost incurred in this process is a challenge. Therefore, methods that use a small or limited number of labeled datasets have been developed. The few-shot object detection method developed for aerial image analysis uses deep metric learning and knowledge inheritance to successfully improve the detection performance in novel categories (Li, W. Z. et al., 2023). A deep metric learning approach with generative adversarial network regularization was designed to achieve more accurate high spatial resolution remote sensing image retrieval with small training samples. It has been demonstrated to outperform state-of-the-art methods (Cao et al., 2020). Various characterization methods for remote sensing images have been proposed to address the increasing complexity of observational data associated with recent developments in remote-sensing technology. Characterization methods that incorporate deep metric learning are promising in this area; however, the limitations of labeled data may prevent their application to various opportunities. To address this issue, a semi-supervised deep metric learning model (specifically, a metric space that jointly preserves the discrimination capability for labeled and unlabeled remote sensing scenes trained using convolutional neural networks (CNNs)) was proposed, which can be implemented with a small amount of labeled data (Kang et al., 2020). The impact of annotation errors on image characterization was also investigated, and a new loss function was employed to augment label noise tolerance in the remote sensing image characterization framework based on deep metric learning (Kang et al., 2021). Sphere loss has also been proposed to simultaneously reduce the intra-class distance and increase the interclass distance (Wang et al., 2021). A similarity retention loss based on deep metric learning (Zhao et al., 2020) has also been proposed to improve the efficiency of the retrieval of large amounts of remote sensing images, and a new architecture for deep metric learning based on residual attention has been developed (Cheng et al., 2021). In this architecture, the residual attention is improved in terms of number and position, as the residual attention branch obtains more distinctive features, and improvements in search tasks have been demonstrated.

Deep metric learning has also been applied to images obtained using various sensors and to synthetic images. Remote-sensing images obtained using hyperspectral imaging comprise tens to hundreds of spectral bands. This vast amount of spectral information provides better object discrimination than multispectral imagery and has been applied to environmental monitoring and precision agriculture. Synthetic aperture radar (SAR), an all-weather sensor, observes the physical properties of the ground surface, such as bumps, slopes, and undulations. Polarimetric SAR observations obtained using SAR enable object identification by receiving multiple pieces of information. Because of these characteristics, SAR sensors are used in variety of applications, such as monitoring ground changes (uplift and subsidence) (Zhang et al., 2021) and deforestation (Bouvet et al., 2018). As with the remote sensing images described above, the limitation of labeled training data has been an issue for images obtained using either sensor. To address this problem, hyperspectral image classification, for example, proposed a deep metric learning-based feature embedding model that can address both same-scene and cross-scene classification tasks under the constraint of a few labeled samples (Deng et al., 2019). New CNNs have been developed using both multiscale convolution and determinantal point process-based diversity-promoting deep metrics (Gong et al., 2019), and the application of online hard mining concepts to deep metric learning has significantly improved the classification accuracy of limited labeled data (Dong et al., 2021). A new deep semi-supervised metric learning network was proposed to reduce labeled data limitations in polarimetric SAR classification, and the use of discriminative information obtained from metric learning improved the classification performance (Liu et al., 2020). A pan-sharpened image, which is a high-resolution multispectral image obtained by combining high-resolution panchromatic and low-resolution multispectral images (Vivone et al., 2015), was designed to overcome the physical limitations of spatial and spectral resolutions in optical imaging. Despite advancements in various pan-sharpening techniques, the inherent biases introduced during image synthesis continue to present challenges, highlighting the need for further refinement in this field. A combined approach incorporating deep metric learning was demonstrated to perform better than previous methods in sharpening spatial information and preserving spectral information (Xing et al., 2018).

Remote sensing is used to monitor changes in land use and vegetation in various situations, for example, changes caused by temporary events, such as human-induced disturbances and natural disasters, to relatively long-term events, such as climate change. The recent developments in observation technology have provided various possibilities of wide-area monitoring by aircraft and satellites. However, there are also many challenges, such as the processing and analysis of increasingly complex data and the spatiotemporal imbalance in the amount of data. An approach that incorporates deep metric learning to address these issues may enable more stable and accurate global monitoring. On the contrary, since it is difficult to predict the ecological impacts of environmental changes caused by recent climate change and anthropogenic loading, it has become paramount to develop frameworks that can handle unknown categories in the future.

Possible application of deep metric learning to ecological field observation data

In ecosystem monitoring, many computer vision techniques were employed in surveys conducted at the ground level. In recent years, the miniaturization of observation equipment, improved battery performance, and the development of recording media have made it possible to obtain large-scale field data. Against such backdrops, the handling, analysis, classification, and clustering of large amounts of data, are challenged by limited training data and extreme data-volume imbalances among classes. Deep metric learning has provided many effective solutions to such problems in remote sensing. However, their application in ground-level ecosystem monitoring is still in its initial stages. Miele et al. (2021) proposed revisiting animal re-identification using image similarity networks and metric learning with CNNs to re-identify individual giraffes based on their body surface patterns. In this study, deep metric learning techniques were utilized to overcome an unknown class, which is difficult to solve using current deep learning models. Despite the limited number of images per individual in the training dataset, the CNN re-identification performance reached a top-1 accuracy of approximately 90%, whereas it performed slightly worse for unknown individuals.

Recently, its application to DNA sequence data obtained from environmental DNA metabarcoding has been reported. For dimensionality reduction and clustering of the huge amount of complex and high-dimensional sequence data generated, Lamperti et al. (2023) proposed a deep learning model that incorporates deep metric learning and combines multiple neural networks. They used this method to visualize ecological characteristics from environmental DNA datasets in two-dimensional space and demonstrated that features could be extracted more effectively than compared with previous methods.

While these methods demonstrate the potential of deep metric learning for other flora and fauna and various datasets, they also suggest challenges to overcome, such as the need for more valid training datasets, diverse data collection, training time proportional to the data volume, and the identification of unknown classes. Recently, a zero-shot deep metric learning approach using only a few samples (or even one sample) was proposed to identify diseases and pests in plant leaves (Zabihzadeh & Masoudifar, 2023). The proposed method uses General Discriminative Feature Learning (Al-Kaabi et al., 2023) as the deep feature extractor and uses a proxy-based loss that effectively captures the overall structure of the embedding space with fast convergence. This approach is effective in rare cases or when it is difficult to collect large datasets. If such few-shot or zero-shot learning techniques are applied to surveys of rare species of wild plants and animals, there will be an increasing number of opportunities to provide important information for the maintenance and conservation of biodiversity.

Like individual identification, wildlife behavior recognition provides important information for ecosystem monitoring. Information obtained from wildlife behavior recognition is very important for maintaining and conserving ecosystems, such as understanding the ecology of animals and their distribution and movement based on this information. Behavior recognition using deep metric learning has already been applied to humans and has shown promising results as a solution to the problems of intra-class variation due to individual differences and the processing of untrained behaviors (Gutoski et al., 2021; Li et al., 2024). These challenges could also apply to wildlife behavior detection, which is difficult to encounter, and extreme imbalances in data can be expected. Therefore, applying human-applied technologies to wildlife will accelerate the development of ecosystem monitoring technologies. However, this may be a barrier to accelerated technological development, because the amount and variety of wildlife behavior data are far less balanced than those of humans. Advancements in observation equipment and methods, as well as the development of new technologies, such as few-shot learning or zero-shot learning, may provide solutions to this challenge.

Future remarks

Deep metric learning is very effective at learning distances and similarities between data and can be a powerful tool for extracting similarities and changes in complex data often observed in ecosystems. In past ecosystem monitoring, some data, especially ground-level data, were often spatiotemporally heterogeneous and difficult to handle. However, for themes such as climate change, where long-term data validation is important, incorporating heterogeneous data with the latest large-scale data may lead to new hypotheses. Developing methods for monitoring biodiversity and population dynamics using deep metric learning will provide opportunities to solve ecosystem monitoring challenges, such as manual image analysis in the big data era, and open new avenues for biodiversity research and conservation. However, future challenges such as data dependency, model overfitting, and interpretability remain. Collaboration among researchers from various disciplines is essential for maximizing the potential of deep metric learning to protect ecosystems and mitigate the impacts of human activities and climate change (Carey et al., 2019).

Author contributions

Yurika Oba: Conceptualization, Writing - Original Draft. Hideyuki Doi: Conceptualization, writing-review, and editing.

Funding

This study did not receive any specific grants from funding agencies in the public, commercial, or non-profit sectors.

Figure captions

Fig. 1

Visual representation of framework of deep learning and deep metric learning.

Fig. 2

Distance Relationship for a Siamese Network (A) Desired handwritten data discrimination for 3 and 8 digits (B) after Siamese network applied to MNIST data for 3 and 8 digits. The figures and captions are taken from Kaya and Bilge (2019). Note: The number of epochs indicates how many times to iterate over the entire training dataset.

Figures

Fig. 1

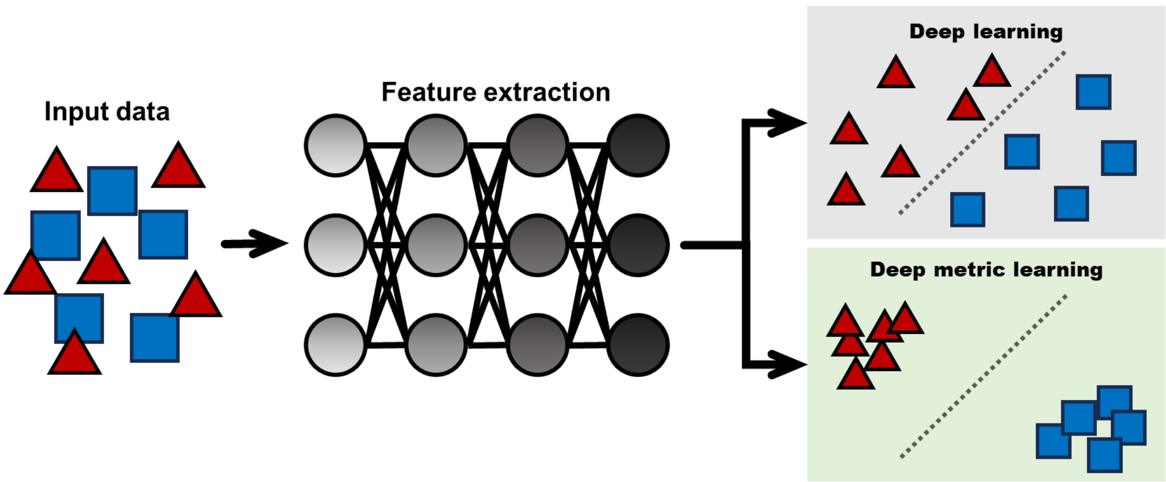
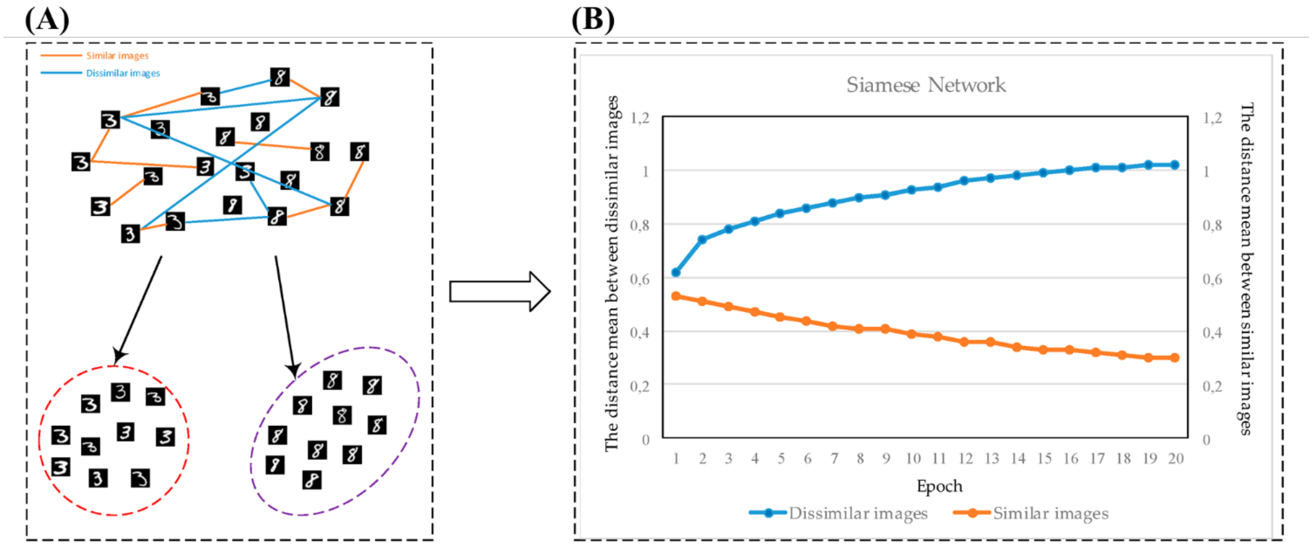


Fig. 2



References

Al-Kaabi, K., Monsefi, R., & Zabihzadeh, D. (2023). A framework to enhance generalization of deep metric learning methods using general discriminative feature learning and class adversarial neural networks. *Applied Intelligence*, *53*(8), 8693–8711. doi:10.1007/s10489-022-03959-6.

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., & Le Toan, T. (2018). Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. *Remote Sensing*, *10*(8), 1250. doi:10.3390/rs10081250.

Cao, Y., Wang, Y. B., Peng, J. H., Zhang, L. Q., Xu, L. L., Yan, K., & Li, L. (2020). DML-GANR: Deep metric learning with generative adversarial network regularization for high spatial resolution remote sensing image retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(12), 8888–8904. doi:10.1109/TGRS.2020.2991545.

Carey, C. C., Ward, N. K., Farrell, K. J., Lofton, M. E., Krinos, A. I., McClure, R. P., & Arzberger, P. (2019). Enhancing collaboration between ecologists and computer scientists: Lessons learned and recommendations forward. *Ecosphere*, *10*(5), e02753. doi:10.1002/ecs2.2753.

Cheng, Q. M., Gan, D. Q., Fu, P., Huang, H. Y., & Zhou, Y. Z. (2021). A novel ensemble architecture of residual attention-based deep metric learning for remote sensing image retrieval. *Remote Sensing*, *13*(17), 3445. doi:10.3390/rs13173445.

Deng, B., Jia, S., & Shi, D. M. (2019). Deep metric learning-based feature embedding for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(2), 1422–1435. doi:10.1109/TGRS.2019.2946318.

Dong, Y. N., Yang, C., & Zhang, Y. X. (2021). Deep metric learning with online hard mining for hyperspectral classification. *Remote Sens.*, *13*(7), 1368. doi:10.3390/rs13071368.

Franceschini, S., Meier, A. C., Suan, A., Stokes, K., Roy, S., & Madin, E. M. (2023). A deep learning model for measuring coral reef halos globally from multispectral satellite imagery. *Remote Sensing of Environment*, *292*, 113584. doi:10.1016/j.rse.2023.113584.

Golwalkar, R., & Mehendale, N. (2022). Masked-face recognition using deep metric learning and FaceMaskNet-21. *Applied Intelligence*, *52*(11), 13268–13279. doi:10.1007/s10489-021-03150-3.

Gong, Z. Q., Zhong, P., Yu, Y., Hu, W. D., & Li, S. T. (2019). A CNN with multiscale convolution and diversified metric for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(6), 3599–3618. doi:10.1109/TGRS.2018.2886022.

Gutoski, M., Lazzaretti, A. E., & Lopes, H. S. (2021). Deep metric learning for open-set human action recognition in videos. *Neural Computing & Applications*, *33*(4), 1207–1220. doi:10.1007/s00521-020-05009-z.

Hu, J., Lu, J., & Tan, Y. P. (2014). Discriminative deep metric learning for face verification in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1875–1882). doi:10.1109/CVPR.2014.242.

Jamali, A., Mahdianpari, M., Brisco, B., Mao, D., Salehi, B., & Mohammadimanesh, F. (2022). 3DUNetGSFormer: A deep learning pipeline for complex wetland mapping using generative adversarial networks and Swin transformer. *Ecological Informatics*, *72*, 101904. doi:10.1016/j.ecoinf.2022.101904.

Kang, J., Fernandez-Beltran, R., Duan, P., Kang, X., & Plaza, A. J. (2021). Robust normalized softmax loss for deep metric learning-based characterization of remote sensing images with label noise. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(10), 8798–8811. doi:10.1109/TGRS.2020.3042607.

Kang, J., Fernández-Beltrán, R., Ye, Z., Tong, X. H., Ghamisi, P., & Plaza, A. (2020). High-rankness regularized semi-supervised deep metric learning for remote sensing imagery. *Remote Sensing*, *12*(16), 2603. doi:10.3390/rs12162603.

Kaya, M., & Bilge, H. Ş. (2019). Deep metric learning: A survey. *Symmetry*, *11*(9), 1066. doi:10.3390/sym11091066.

Kosuge, A., Yu, L. X., Hamada, M., Matsuo, K., & Kuroda, T. (2023). A deep metric learning-based anomaly detection system for transparent objects using polarized-image fusion. *IEEE Open Journal of the Industrial Electronics Society*, *4*, 205–213. doi:10.1109/OJIES.2023.3284014.

Lamperti, L., Sanchez, T., Si Moussi, S. S., Mouillot, D., Albouy, C., Flück, B., . . . Manel, S. (2023). New deep learning-based methods for visualizing ecosystem properties using environmental DNA metabarcoding data. *Molecular Ecology Resources*, *23*(8), 1946–1958. doi:10.1111/1755-0998.13861.

Li, W. Z., Zhou, J. W., Li, X., Cao, Y., & Jin, G. (2023). Few-shot object detection on aerial imagery via deep metric learning and knowledge inheritance. *International Journal of Applied Earth Observation and Geoinformation*, *122*, 103397. doi:10.1016/j.jag.2023.103397.

Li, X. F., Lu, J., Chen, X. G., & Zhang, X. D. (2024). Spatial-temporal adaptive metric learning network for one-shot skeleton-based action recognition. *IEEE Signal Processing Letters*, *31*, 321–325. doi:10.1109/LSP.2024.3351070.

Li, X. X., Yang, X. C., Ma, Z. Y., & Xue, J. H. (2023). Deep metric learning for few-shot image classification: A review of recent developments. *Pattern Recognition.*, *138*, 109381.

doi:10.1016/j.patcog.2023.109381.

Liu, H. Y., Luo, R. Y., Shang, F. H., Meng, X. C., Gou, S. P., & Hou, B. (2020). Semi-supervised deep metric learning networks for classification of polarimetric SAR data. *Remote Sensing*, *12*(10), 1593. doi:10.3390/rs12101593.

Lu, J., Hu, J., & Zhou, J. (2017a). Deep metric learning for visual understanding: An overview of recent advances. *IEEE Signal Process. Mag*, *34*(6), 76–84. doi:10.1109/MSP.2017.2732900.

Lu, J. W., Hu, J. L., & Tan, Y. P. (2017b). Discriminative deep metric learning for face and kinship verification. *IEEE Transactions on Image Processing*, *26*(9), 4269–4282. doi:10.1109/TIP.2017.2717505.

Marini, S., Fanelli, E., Sbragaglia, V., Azzurro, E., Del Rio Fernandez, J., & Aguzzi, J. (2018). Tracking fish abundance by underwater image recognition. *Scientific Reports*, *8*(1), 13748. doi:10.1038/s41598-018-32089-8.

Miele, V., Dussert, G., Spataro, B., Chamaillé-Jammes, S., Allainé, D., & Bonenfant, C. (2021). Revisiting animal photo-identification using deep metric learning and network analysis. *Methods in Ecology and Evolution*, *12*(5), 863–873. doi:10.1111/2041-210X.13577.

Pasher, J., Smith, P. A., Forbes, M. R., & Duffe, J. (2014). Terrestrial ecosystem monitoring in Canada and the greater role for integrated earth observation. *Environmental Reviews*, *22*(2), 179–187. doi:10.1139/er-2013-0017.

Vance, T. C., Huang, T., & Butler, K. A. (2024). Big data in earth science: Emerging practice and promise. *Science*, *383*(6688), eadh9607. doi:10.1126/science.adh9607.

Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G. A.& Wald, L. (2015). A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(5), 2565–2586. doi:10.1109/TGRS.2014.2361734.

Wang, J., Chen, H., Ma, L., Chen, L., Gong, X. D., & Liu, W. C. (2021). Sphere loss: Learning discriminative features for scene classification in a hyperspherical feature space. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–19. doi:10.1109/TGRS.2021.3059101.

Xing, Y. H., Wang, M., Yang, S. Y., & Jiao, L. C. (2018). Pan-sharpening via deep metric learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 165–183. doi:10.1016/j.isprsjprs.2018.01.016.

Zabihzadeh, D., & Masoudifar, M. (2023). ZS-DML: Zero-shot deep metric learning approach for plant leaf disease classification. *Multimedia Tools and Applications*, 83(18), 54147–54164. doi:10.1007/s11042-023-17136-5.

Zhang, X., Han, L., Han, L., & Zhu, L. (2020). How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery? *Remote Sensing*, *12*(3), 417. doi:10.3390/rs12030417.

Zhang, Y. H., Wu, H. A., Li, M. J., Kang, Y. H., & Lu, Z. (2021). Investigating ground subsidence and the causes over the whole Jiangsu Province, China using Sentinel-1 SAR data. *Remote Sensing*, *13*(2), 179. doi:10.3390/rs13020179.

Zhao, H. W., Yuan, L., & Zhao, H. Y. (2020). Similarity retention loss (SRL) based on deep metric learning for remote sensing image retrieval. *ISPRS International Journal of Geo-Information*, *9*(2), 61. doi:10.3390/ijgi9020061.