Water, as one of the most essential resources on the Earth, not only sustains human life and ecological cycles, but also plays a vital role in economic development and mineral exploration (Vorosmarty et al., 2000). Therefore, constructing large-scale to over all water conditions, combining multi-modal and multi-temporal RS data and domain knowledge to improve the interpretation generalization in other areas and long-term perception ability would be the future research directions of water body classification. In this context, water identification aims at classifying if a pixel from RS images is water or not, which has gradually evolved into a hot research topic in RS, and scholars have carried out many studies to extract water bodies from various RS images (Hollstein et al., 2016; Dang and Li, 2021). Focusing on multiple-sensor images processing, Li et al. (2021c) presented an encoder-decoder-based dense-localfeature-compression (DLFC) network to extract valuable spatial and spectral details. How to enhance the interpretation effect with limited samples has always been a hot topic for scholars in RS. Li et al. (2021b) used a region of interest (ROI) to build water labels and then proposed a pixel-based CNN to synchronously combine spectral and texture information for water segmentation. To fully use microwave images, Xue et al. (2021) introduced a densecoordinate-feature concatenate network (DCFCN) for merging water body features from dual-polarimetric SAR images and thus address the ground interference in single-polarization SAR images. A portion of this is surface water and mainly involves rivers, lakes, canals, and ponds; the oceans are always excluded from this category due to their large extent and salty characteristic (Huang et al., 2018). Later, many improved water indexes were developed, such as the modified NDWI (MNDWI) (Xu, 2006), EWI (Yan et al., 2007), NWI (Feng, 2012), HRWI (Yao et al., 2015) and two-step TSUWI (Wu et al., 2018). Next, a restricted receptive field DeconvNet (RRF DeconvNet) was presented to compresses redundant layers and then uses an edge weighting loss to extract accurate water edges (Miao et al., 2018). With the growth of a severe global water shortage and increasing flood disasters, the efficient and accurate assessment of available surface water has become an essential part of ecological protection, urban planning, and industrial production. Over the past decades, Earth observation techniques have advanced significantly and many kinds of high-resolution RS images are accessible for various real-world applications, including efficiently identifying available water bodies. Therefore, data-driven ML- and DL based methods can adaptively leverage spectral and spatial information to build discriminative features for efficient and accurate water classification. Due to the simplicity of DT, Hollstein et al. (2016) developed several DTbased classification methods for simultaneously extracting water, snow, clouds, and so on from Sentinel 2 images. Differently, (Morsy et al., 2018) designed a unsupervised land/water classification method to automatically monitor the changes of coastal areas from multi-spectral airborne LiDAR data. Isikdogan et al. (2017) demonstrated an early application of a data-driven FCN for surface water mapping of Landsat 7 images, which was different from the threshold methods in adopting different regions and imaging conditions.