

Water Resources Research®

RESEARCH ARTICLE

10.1029/2023WR036342

Key Points:

- Remote sensing data augmentation for improving the accuracy of environmental assessments can be achieved using adversarial networks
- High spatiotemporal resolution of water bodies data enhances the precision of their areal forecasting
- Shape and areal accuracies play an important role for efficient spatiotemporal data interpolation

Correspondence to:

S. Filali Boubrahimi,
soukaina.boubrahimi@usu.edu

Citation:

Filali Boubrahimi, S., Neema, A., Nassar, A., Hosseinzadeh, P., & Hamdi, S. M. (2024). Spatiotemporal data augmentation of MODIS-landsat water bodies using adversarial networks. *Water Resources Research*, 60, e2023WR036342. <https://doi.org/10.1029/2023WR036342>

Received 16 OCT 2023

Accepted 24 FEB 2024

Author Contribution:

Writing – review & editing:
Pouya Hosseinzadeh

Spatiotemporal Data Augmentation of MODIS-Landsat Water Bodies Using Adversarial Networks

Soukaina Filali Boubrahimi¹ , Ashit Neema¹, Ayman Nassar^{2,3} , Pouya Hosseinzadeh¹ , and Shah Muhammad Hamdi¹ 

¹Department of Computer Science, Utah State University, Logan, UT, USA, ²Utah Water Research Laboratory, Department of Civil and Environmental Engineering, Utah State University, Logan, UT, USA, ³Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT, USA

Abstract With increasing demands for precise water resource management, there is a growing need for advanced techniques in mapping water bodies. The currently deployed satellites provide complementary data that are either of high spatial or high temporal resolutions. As a result, there is a clear trade-off between space and time when considering a single data source. For the efficient monitoring of multiple environmental resources, various Earth science applications need data at high spatial and temporal resolutions. To address this need, many data fusion methods have been described in the literature, that rely on combining data snapshots from multiple sources. Traditional methods face limitations due to sensitivity to atmospheric disturbances and other environmental factors, resulting in noise, outliers, and missing data. This paper introduces Hydrological Generative Adversarial Network (Hydro-GAN), a novel machine learning-based method that utilizes modified GANs to enhance boundary accuracy when mapping low-resolution MODIS data to high-resolution Landsat-8 images. We propose a new non-saturating loss function for the Hydro-GAN generator, which maximizes the log of discriminator probabilities to promote stable updates and aid convergence. By focusing on reducing squared differences between real and synthetic images, our approach enhances training stability and overall performance. We specifically focus on mapping water bodies using MODIS and Landsat-8 imagery due to their relevance in water resource management tasks. Our experimental results demonstrate the effectiveness of Hydro-GAN in generating high-resolution water body maps, outperforming traditional methods in terms of boundary accuracy and overall quality.

Plain Language Summary This study addresses the imperative challenges of water resource management, including coastal zone oversight, detecting sea border shifts due to rising waters, and erosion tracking. Satellite data currently offers a choice between high spatial detail with infrequent updates or lower spatial detail with more frequent updates, presenting a trade-off between data precision and frequency. To efficiently monitor environmental resources like water bodies, we require data with both high spatial detail and frequent updates. To meet this need, we introduce the Hydrological Generative Adversarial Network, a novel machine learning tool that enhances data clarity, particularly in outlining water bodies. In testing, we employed images from the Moderate Resolution Imaging Spectroradiometer satellite, providing less detailed images, and the Land Remote-Sensing Satellite, offering highly detailed imagery. In essence, this study enhances water resource management by effectively combining data from multiple sources, even in adverse conditions, potentially advancing environmental protection and management efforts.

1. Introduction

Water bodies monitoring is essential to guide evidence-based decision making which is necessary for hydrological and ecological sustainability (Njue et al., 2019). Surface water is an irreplaceable resource for ecological systems, human uses, industrial uses, hydro-power generation, social development, and recreation. Reliable information about the dynamic changes of open surface water (e.g., lakes, reservoirs, and rivers) is critically important for various scientific disciplines, such as flood prediction, coastal zone management, coast erosion, agricultural sustainability, watershed analysis, climate models, and the assessment of present and future water resources (Sarp & Ozelik, 2017). Previous studies have suggested that lakes and reservoirs are good sentinels of global climate change because they are sensitive to environmental changes (Adrian et al., 2009).

Recent advancement in satellite based-remote sensing information with different spatial, spectral, radiometric, and temporal resolutions has given new dimensions to the water bodies studies (Ismail et al., 2021). Following the

© 2024. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

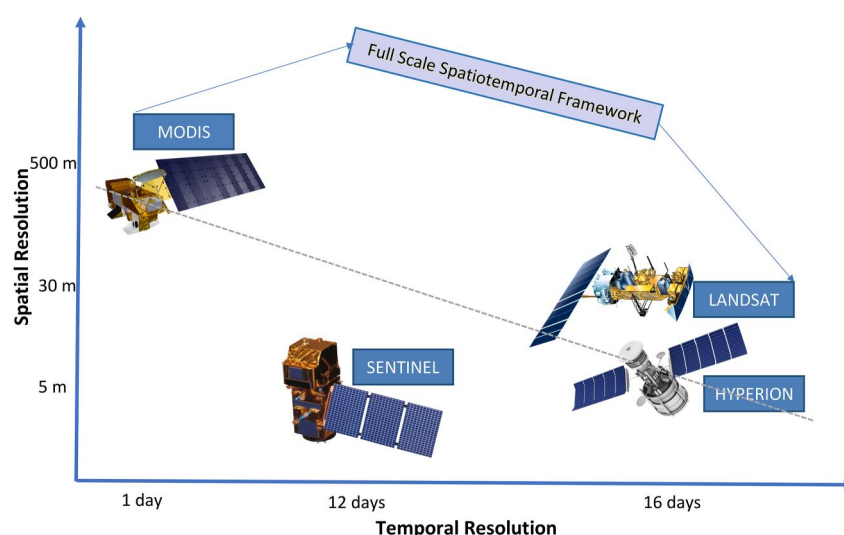


Figure 1. Spatial and temporal resolutions of earth's different satellites.

increase in the availability of satellite images, and image processing techniques, numerous research studies have attempted to extract and delineate water bodies from these images (Ouma & Tateishi, 2006). These technological and methodological advancements, shift the analysis of surface water bodies from regional-scale to global scale for a better understanding of the Earth's natural processes.

Earth observation data is acquired through a large number of satellites that have unique spatiotemporal resolutions making them complementary data sources (Amato et al., 2020). In the 50 years since the first satellite launch, remote sensing satellites have evolved from producing low-resolution images to daily data acquisition exceeding 10 terabytes (Campbell & Wynne, 2011). This transformation is driven by Earth observation science, with over 150 observation satellites in orbit, equipped with sensors operating at various spatial and temporal resolutions (Fu et al., 2020).

Due to the differences in sensor designs, there is often a trade-off in different spatiotemporal data resolutions across the remote sensing spectrum (Khandelwal, Karpatne, & Kumar, 2017). Figure 1 illustrates the spatio-temporal resolution of four active satellites which shows the aforementioned space and time trade-off. It can be noted that the highest spatial resolution satellite is compromising time resolution and vice versa. MODIS satellite data, obtained from both TERRA and AQUA instruments, observes the entire Earth's surface every 1–2 days. While the temporal resolution of individual MODIS observations is 8 days, due to the combined coverage from both TERRA and AQUA satellites, data from the same location on Earth is acquired approximately every 8 days. On the other hand, Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) sensors onboard the Landsat satellite capture the earth's surface every 16 days but at a high spatial resolution (HSR) of 30 m (refer Figure 1). In other words, We utilize MODIS data, which is available in various spatial resolutions, including 250 m, 500 m, and 1 km. For our study, we specifically employ MODIS data at a spatial resolution of 500 m. Additionally, Landsat data, captured by OLI with a spatial resolution of 30 m, is used in our research. TIRS on Landsat has a spatial resolution of 100 m. Given the limitation of individual sensors at either delivering HSR or high temporal cadence, it is essential to create interpolation techniques that can transmit information across spatiotemporal scales (Khandelwal, Karpatne, & Kumar, 2017). The knowledge transfer has the ability to equip the science communities with synthetic data sets that approximate real data on the go. For the case of water bodies interpolation, one of the desirable properties of the method is to achieve interpolated water bodies that are similar to the ground truth data in both shape and area.

High spatiotemporal resolution data will allow timely monitoring of the surface water and dynamics which are crucial elements for policy and decision-makers in hydrology and geomorphology (Maiti & Bhattacharya, 2009). Furthermore, interpolated data will facilitate the integration of remote sensing data with Geographic Information Systems for automatic or semiautomatic water body extraction and mapping (Sarp & Ozcelik, 2017). In the light of the aforementioned need for a flexible spatiotemporal resolution domain, in this paper, we propose a method

that aims to learn a mapping between low spatial resolution (LSR) image instances and HSR image instances of water bodies and reservoirs collected across a period of 7 years from MODIS and Landsat satellites. The proposed mapping can be utilized to generate HSR data at times when only LSR data is available. Our contributions are as follows:

- We designed an image processing pipeline that uses computer vision tools to extract target water bodies' polygon boundaries from satellite imagery.
- We developed a new Generative Adversarial Network (GAN) for mapping the data between any LSR-HSR satellite pair.
- We made our source code and data open-source in a project website (<https://sites.google.com/view/hydro-ml/>) that meets the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR) (Wilkinson et al., 2016).

Our contribution can help in extracting the precise shape and area of a water body which can further be used to measure the expansion or shrinking of a water body over a period of time. This can be crucial as water body extraction is an important task in different disciplines, such as lake coastal zone management, coastline change, and erosion monitoring, flood prediction, climate and environmental change, and evaluation of water resources (Haghnazar et al., 2024; Hosseinzadeh et al., 2023, 2024; Jiang et al., 2018). Timely monitoring of surface water and delivering data on the dynamics of surface water are also essential for policy and decision-making processes (Sarp & Ozcelik, 2017). It can also help in detecting the changes in urban water bodies that make a huge difference to human lives and may cause disasters, such as surface subsidence, urban inland inundation, and health problems (Y. Chen et al., 2018).

The rest of the paper is organized as follows: in Section 1 we review the related works; in Section 2 we define the architecture of our proposed GAN-based method and its different variants; in Section 3 we discuss the results and evaluation of our experiments. We introduce a case study on Lake Thathar in Section 4. Finally, Section 5 concludes the paper and shows potential directions for future works.

1.1. Background on LSR-HSR Image Mapping

Early efforts to learn surface water mapping between LSR and HSR images include supervised learning methods of remote sensing images (Khandelwal, Karpatne, & Kumar, 2017). Factors such as noise, outliers, and enormous volumes of missing data (due to clouds and sensor failures) can impact the accuracy of the aforementioned classification systems. To overcome the latter limitations, a new approach has emerged called the Ordering Based Information Transfer Across Space and time (ORBIT) (Khandelwal, Karpatne, & Kumar, 2017). The fundamental idea behind ORBIT is to take advantage of the natural ordering of instances that results from the elevation structure and temporal context. Specifically, if an area is filled with water, it is interpreted that due to the gravity all the regions in the basin with lower elevation than the current area are also occupied by water (Pekel et al., 2016). One key assumption in this approach is that a water body always expands and contracts smoothly (with the exception of floods and other natural phenomena) which suggests that the surface area of the water body at close time steps are highly similar (Khandelwal, Karpatne, Marlier, et al., 2017). To put it another way, the water surface extents of dates close together are likely to be highly similar.

Hence, ORBIT can map data from LSR to HSR data only at time steps when noiseless LSR data is available (Khandelwal, Karpatne, & Kumar, 2017).

1.2. Background on Planform Change Use Case

For the use case of channel planform change analysis in river bodies, the change of a water body channel is due to natural or human-made fluctuations in the streamflow or sediment supply (Leonard et al., 2020). It has been found that although recurrent images from remote sensing satellites have been widely used to measure the channel change, these measurements are only significant if the measure of the change is more than the uncertainty threshold (Leonard et al., 2020). To address this challenge, a generalized method was introduced by Christina M. et al. for quantifying the uncertainty associated with measurements of channel change from remote sensing images based on spatially varying estimates of uncertainty called the spatially distributed probabilistic (SDP) method (Leonard et al., 2020). The SDP approach leverages image co-registration error, interpretation uncertainty, and digitization uncertainty for quantifying uncertainty. It has been established that SDP can be used to

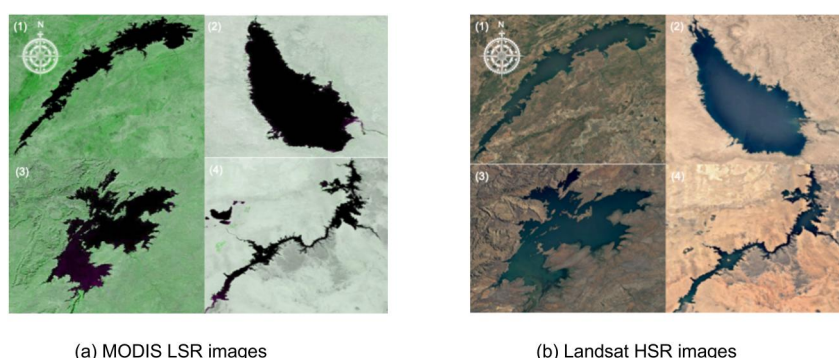


Figure 2. Images from MODIS and Landsat satellites of (1) Kariba reservoir (Zambia), (2) Lake Thathar (Iraq), (3) Lake Argyle (Australia), and (4) Lake Nasser (Egypt).

calculate uncertainty at specific locations of linear channel adjustment or polygons of erosion and deposition, while also estimating the central tendency of the net planform change.

1.3. Background on Data Pre-Processing

Prior to building the LSR to HSR data mapping, it is necessary to build a data pipeline that prepares the LSR-HSR corresponding data pairs. For this purpose, if the scope of the study is limited to a few water bodies, it is common to use data maps that extract the apriori known polygons (Frazier & Page, 2000; Muster et al., 2013). Image processing techniques have also been used recently for automatically extracting the water bodies' outlines from satellite data without any apriori map knowledge (Rokni et al., 2014). Single-band methods are automated polygon extraction approaches that utilize a selected threshold value to extract water bodies' boundaries. Similarly, multi-band methods combine different reflective bands for improved surface water extraction (Rokni et al., 2014). The weakness of using a pixel threshold is that it is prone to errors caused by the mixing of water pixels with those of different cover types. A more sophisticated approach for automated polygon extraction is to employ image segmentation. The latter technique is relatively more accurate compared with single-band methods (Rokni et al., 2014).

1.4. Background on Generative Networks

GAN models have been used in a variety of applications, including image synthesis, semantic image editing, style transfer, and classification (Goodfellow et al., 2014). These networks not only learn the mapping from an input image to an output image but also learn a loss function to train this mapping. This makes it possible to apply the same generic approach to problems that traditionally would require very different loss formulations (Isola et al., 2016). Regular GANs hypothesize the discriminator as a classifier with the sigmoid cross-entropy loss function (Mao et al., 2016). In this paper, we develop a machine learning-empowered synthetic satellite that is capable of spatiotemporal interpolation across a pair of real satellites. Our approach is built on the GAN model that we used to accurately represent water bodies' shapes and generate realistic synthetic HSR image instances at times when no measurements are available. Our proposed model pertains to the unsupervised learning paradigm that is capable of learning deep representations without extensively labeled training data. The novelty of our approach consists of deriving back-propagation signals through a competitive training process involving a pair of competing networks. Our GAN architecture utilizes a generator model and a discriminator model, where the first model is used to generate new synthetic images, and the second model is used to classify the data as either real or synthetic. The two competing models are trained concurrently in an adversarial process, where the role of the generator is to mislead the discriminator, on the other hand, the discriminator tries to detect the generated images (Goodfellow et al., 2014).

2. Methodology and Data Sources

An ideal LSR-HSR image mapping method produces images that accurately describe the shape of the water bodies' boundaries. For this purpose, we develop a machine-learning model (i.e., a GAN model) that is specifically equipped to focus on the shape and areal accuracy of the water bodies' interpolated polygons. To better

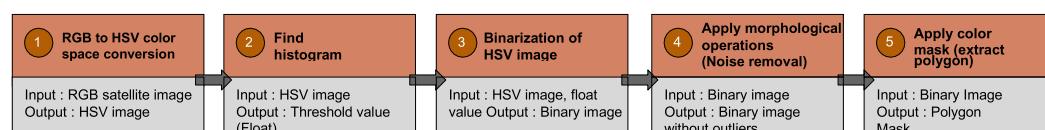


Figure 3. Data pre-processing pipeline.

represent the polygon ground truth shapes, we use an optimization algorithm that minimizes the loss with respect to the polygon shapes. Our model is trained on historical image data captured from both Landsat and MODIS satellites that are HSR and LSR respectively (L. Yu et al., 2018). In this section, we will discuss the data pre-processing pipeline and the proposed model.

2.1. Water Bodies Data Sources

Our data set is collected from MODIS and Landsat Earth Observation satellites. Figure 2 shows a sample of LSR water bodies captured onboard MODIS and their corresponding HSR captured by the Landsat satellite. MODIS is a key instrument aboard the Terra and Aqua satellites. While Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, Aqua passes south to north over the equator in the afternoon (Justice et al., 2002). Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1–2 days, acquiring data in 36 spectral bands, or groups of wavelengths (Barnes et al., 2003). We used Terra MODIS with bands 1-2-1 and 7-2-1 to obtain 256 by 256 pixel size images, as the bands allow the separation between the water bodies and land surfaces (refer Figure 2a). The data collected from the MODIS sensors are merged at a temporal resolution of 8 days and a spatial resolution of 500 m. It is important to note that although the temporal resolution of MODIS is 1 day, the satellite acquires data from the same location on Earth approximately every 8 days. This is known as the 8-day repeating cycle (i.e., MODIS 8-day composite) (Y. Chen et al., 2013). Landsat 8 consists of two data collection instruments called OLI and TRIS. The two sensors collect data at a temporal resolution of 16 days. OLI provides a spatial resolution of 30 m, capturing visible and near-infrared bands, while TIRS has a spatial resolution of 100 m, focusing on thermal infrared bands (Verpoorter et al., 2012). We used bands 7-4-3 and 1-5-7 to obtain the images of water bodies (refer Figure 2b). Likewise, we used 256 by 256 pixel size images for Landsat. The multi-band method takes advantage of reflective differences of each involved band and extracts water based on the analysis of signature differences between water and others (Qiao et al., 2012). Our data set contains 20 reservoirs, across 7 years from 2015 to 2021. We have a total of 6,720 images of MODIS sensors and 3,360 images of Landsat 8 satellite. In our study, we utilized satellite data products from both MODIS and Landsat. For MODIS, we employed Level 1B products, which include radiometrically calibrated and geo-located data. These Level 1B products undergo geometric correction, making them suitable for our mapping purposes. Regarding Landsat, we utilized Level 1T products, which are terrain-corrected and georeferenced. We have outlined more details on the water bodies used in this study in the Appendix section. Following the FAIR guiding principles, we have made our data sets publicly accessible on the project website.

2.2. Data Pre-Processing

Prior to the HSR-LSR mapping process, we curated our image data sets to obtain cleaned machine-learning ready data. We applied computer vision methods to extract the shapes of the water bodies' polygons. The image processing step involves extracting useful metadata from the image. In this case, the metadata is the shape of the water bodies' polygons. Our data curation is a four-step process that is shown in Figure 3 and described in Algorithm 1. The steps are (a) convert the input images to Hue Saturation and Value (HSV) format (lines 3–4), (b) binarize the image into a black-and-white format (lines 5–8), (c) denoise the image using morphological operations (lines 9–11), and (d) apply an image mask to extract the water body polygons.

While existing methods on segmentation-based methods for water bodies extraction are powerful techniques for water bodies image analysis tasks, but have a limited capacity to capture fine-grained details and nuances in water bodies, especially in cases where the polygon boundaries are not well-defined (Verma et al., 2021; Xia et al., 2020). Therefore, our proposed data pre-processing pipeline (i.e., binarization and morphological operations) offers a better approach to accurate boundary extraction. Binarizing the image and applying morphological operators are essential steps in image processing for accurate water body detection and shape extraction. On one hand, the binarization process simplifies the image by converting it into a binary format, where pixels are

categorized as either foreground (object of interest) or background. On the other hand, morphological operations, such as erosion, dilation, opening, and closing, are valuable for enhancing and refining the edges of objects in the binary image. These operations can fill gaps, connect broken lines, and smoothen object boundaries.

The first step involved converting the colored image in Red Green Blue (RGB) into HSV color space image. This step is crucial as the HSV color space is the most efficient image format for color-based image segmentation (T.-W. Chen et al., 2008). The HSV color conversion also enables the normalization of colors across images captured from satellites that have water bodies of different colors relative to their surroundings. The HSV color space consists of three matrices, hue, saturation, and value whose ranges are (0–179), (0–255), and (0–255) respectively (Loesdau et al., 2014). While the hue represents the color, saturation measures the amount to which the given color (hue) is mixed with white. The value matrix represents the amount to which the given color (hue) is mixed with black.

Algorithm 1. Data Pre-Processing Algorithm to Extract Polygon Out of a Satellite Image

Input: Dataset D containing satellite images, $\theta_{dilation}$ dilation structuring element, and $\theta_{erosion}$ the erosion structuring element.

Output: Set of extracted polygon images $extracted$.

```

 $extracted \leftarrow \emptyset$ 
for  $current\_image \in [D]$  do
     $hsv = convertToHSV(currentImage)$ ;           ▷ Convert RGB image to HSV format
     $v = extract(hsv)$ ;                             ▷ Extract the gray-scale threshold
     $binaryIm \leftarrow binarize(v, getHistogramPartition(v))$ ;
     $binaryIm \leftarrow (1 - binaryIm) * 255$ ;
     $newImage \leftarrow erosion(binaryIm, \theta_{erosion})$ ;           ▷ Remove outliers
     $newImage \leftarrow dialtion(newImage, \theta_{dilation})$ ;
     $polygonMask = ones(size(currentImage))$            ▷ Create the polygon boundary
     $polygonMask(:, :, 1) \leftarrow 255 - newImage$ 
     $polygonMask(:, :, 2) \leftarrow 255 - newImage$ 
     $extracted \leftarrow append(polygonMask)$ 
end for
return  $extracted$ 

```

The second step of the data pipeline involves binarizing the converted grayscale input images by converting them to black and white pixels. The latter step reduces the domain of colors contained in the image from 256 shades of gray to a binary set (black or white). To achieve the binarization, we first perform a hyperparameter search for a water body-color pixel threshold value based on the distribution of grayscale pixels in the image. We then establish that the pixel is converted to white (value of zero) if the grayscale value of the pixels is greater than the threshold. Similarly, if the grayscale value of the pixel is lower than the threshold, then it is converted into black (value of one). Both MODIS and Landsat images follow an identical pre-processing, which involves converting their original RGB format into grayscale and subsequently into a binarized image.

The third data processing step consists of removing the outliers and noise patches around the water body by applying morphological operations to the binary images. The morphological operations rely on the relative ordering of the pixel values with respect to their values in order to infer outliers (Comer & Delp, 1999). In this operation, we probe an image with a small shape or template called a *structuring element*. The *structuring element* is positioned at all of the possible locations in the image and it is compared with the corresponding pixel neighborhood. We employed two groups of morphological operations. On one hand, the first operation group tests whether the *structuring element* touches or intersects the neighborhood, which is governed by the $\theta_{erosion}$ parameter in Algorithm 1. On the other hand, the second operation group test whether the *structuring element* fits well within the neighborhood, which is governed by the $\theta_{dilation}$ parameter in Algorithm 1.

This is known as morphological opening, and it removes small objects (noise) from an image while preserving the shape and size of larger objects in the image. Figure 4 shows an example of a morphological opening that removes the noisy patches from the original image (Said et al., 2016). The resulting binary image contains a non-zero value only if the *structuring element* morphological tests are successful at a location in the input image. Finally, the last

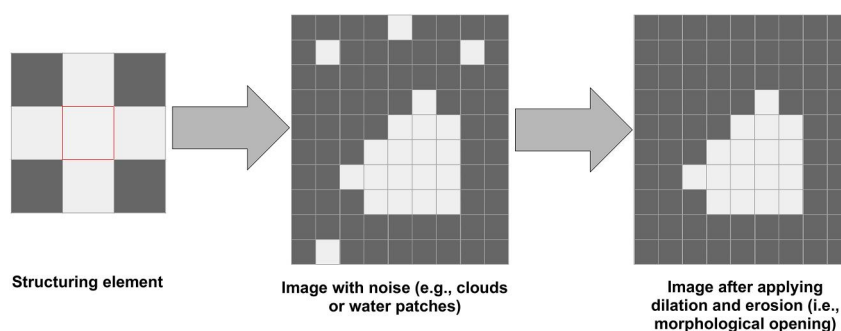


Figure 4. Structuring element example applied on a noisy image.

step consists of applying a color mask to the binary image that extracts the shape of the water body in the form of a polygon. Figure 5 illustrates the entire data pipeline steps using two MODIS LSR images (previously shown in Figure 2a).

2.3. Hydrological Generative Adversarial Network (Hydro-GAN)

Our proposed Hydrological Generative Adversarial Network (Hydro-GAN) is an image-to-image translation model that transforms imagery from the LSR domain to the HSR domain by learning the non-linear mapping between the two. The data product will directly enrich the current HSR data sets at times when only LSR data are available.

We used GAN frameworks as the backbone of our proposed model. Our motive for utilizing GAN is their ability to generate crispy sharp images (C. Li et al., 2017). These networks provide a way to learn deep representations without requiring any extensively annotated training data. GAN models are able to learn the representations by deriving back-propagation signals through a competitive process involving a pair of networks (Martinez & Heiner, 2020).

Hydro-GAN is an unsupervised learning model that automatically learns the regularities and patterns in the input data and tries to mimic the same patterns when generating synthetic samples. The success of Hydro-GAN is determined by the output plausibility in comparison with the original data set. Hydro-GAN is composed of two

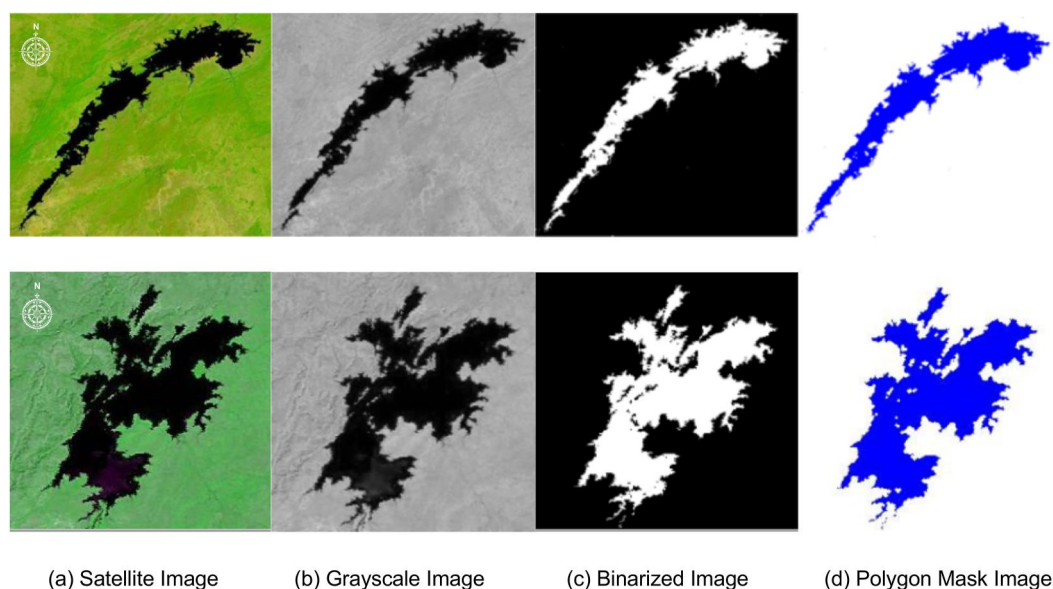


Figure 5. Pre-processing pipeline of (a) Kariba reservoir (Zambia) and Lake Argyle (Australia) MODIS images that are transformed into (b) grayscale images, (c) binarized images, before the (d) water bodies polygons extraction.

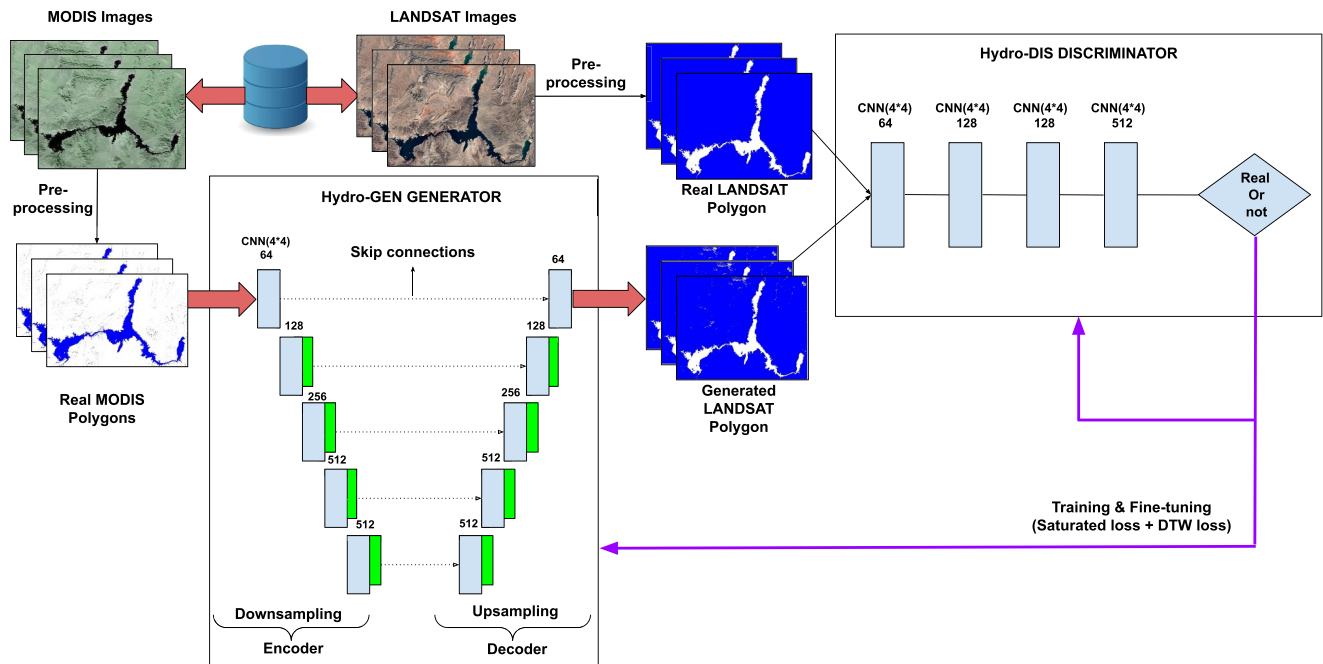


Figure 6. Architecture of Hydro-GEN and Hydro-DIS sub-models of the Hydro-GAN applied on Lake Mead (part of the Colorado River).

sub-models Hydro-GEN and Hydro-DIS. The first sub-model is the generator model which is trained to generate new synthetic images. The second sub-model is the discriminator model that tries to classify the generated water body image as either are real (from the HSR domain) or synthetic (i.e., created by the Hydro-GEN sub-model). While Hydro-DIS learns to optimize its loss function, Hydro-GEN learns to fool the Hydro-DIS discriminator model. As such, the two Hydro-GAN sub-models are trained simultaneously in an adversarial process where the generator seeks to better fool the discriminator and the discriminator seeks to better identify the synthetic generated water bodies images.

Hydro-GEN learns a mapping from a random input x in latent space to an output y that matches the data distribution of the HSR Landsat images: $G : x \rightarrow y$ (Goodfellow et al., 2014). Hydro-GEN utilizes a conditional generative adversarial network (cGAN), where the output y is conditioned on some input z , resulting in a mapping: $G : (x, z) \rightarrow y$ (Goodfellow et al., 2014). The generator is trained via adversarial loss, and an additional soft Dynamic Time Warping (DTW) loss term, calculated between the synthetic image and the ground truth water body images. This added loss motivates the generator to develop accurate water body shapes that match the shape signatures of the original water bodies. In this context, randomness refers to the incorporation of unpredictable and stochastic input values, drawn from probability distributions, to introduce variability and produce diverse and data-matching outputs. Equation 1 defines the objective function that a traditional GAN network is optimizing for.

$$\min \max L(D, G) = \mathbb{E}[\log(D(z, y))] + \mathbb{E}[\log(1 - D(z, G(x, z)))] \quad (1)$$

In Equation 1 $D(z, y)$ refers to the probability that a sample image y pertains to the real data set given the condition z . E is the expected value over all the real image data instances. $G(x, z)$ refers to the synthetic image samples conditioned on z . $D(z, G(x, z))$ is the discriminator's estimate of the probability that a synthetic instance is real.

The discriminator D attempts to maximize the function, on the other hand, the generator G attempts to do the opposite. The generator can't directly affect the $\log(D(x))$ term in the function, so, for the generator, minimizing the loss is equivalent to minimizing $\log(1 - D(G(z)))$.

2.3.1. Hydrological Generator (Hydro-GEN) Model

Our proposed Hydro-GEN sub-model is an encoder-decoder model that uses a U-Net design (Ronneberger et al., 2015). The model takes a source water body pre-processed LSR image and generates a target HSR image.

The model accomplishes this by first down-sampling the input image to the final layer, then up-sampling this layer's data back to the dimensions of the output image (Ronneberger et al., 2015). The most prominent patterns from the input LSR image are learned to be retained and encoded in the bottleneck layer prior to constructing the output HSR image representation. The U-Net architecture contains skip-connections that connect the corresponding sampling layers. The skip connections are used to pass the low-level data of the HSR image through the bottleneck layer.

The generator model consists of standardized blocks of convolution layers, with batch normalization, dropout, and activation applied to it. Figure 6 shows the skip connections used by the U-Net network (Goodfellow et al., 2014). While both U-Net and Hydro-GEN models share the same foundational structure involving an encoder-decoder design, Hydro-GEN focuses on generating high-resolution water body images from low-resolution inputs, making it particularly suited for remote sensing applications. Hydro-GEN incorporates unique loss functions, including adversarial, L1, and DTW losses, to address the challenges specific to water body image enhancement.

The Hydro-GEN model is trained via the Hydro-DIS model. The weighting of the Hydro-GEN model is updated to reflect the discriminator loss when predicting synthetic images as either real or generated (known as the adversarial loss). Hydro-GEN is penalized when generating synthetic samples that are easily distinguishable by Hydro-DIS from the real training data distribution. The Hydro-GEN weighting is also updated to minimize the L1 between the synthetic images and the original ground truth image, which represents the mean absolute error between these images. The final objective function is made of a weighted sum of the combination of these losses. The weights range between 1 and 100 in favor of the adversarial loss and vice versa. We used weighting between the two losses to find an optimal balance between adversarial and L1 losses that can maximize the synthetic image plausibility. A detailed summary of our Hydro-GEN model's architecture in Appendix B1.

2.3.2. Hydrological Discriminator (Hydro-DIS) Model

The Hydro-DIS model is a deep convolutional neural network model that performs image classification based on some conditions (Tran et al., 2017). Hydro-DIS takes both the LSR source image and the target HSR image pairs as input and classifies the target image as real or synthetic. The Hydro-DIS model is designed using the Effective Receptive Field, which are the regions that contain input pixels with a non-negligible impact (Luo et al., 2017). ERFs provide a one-to-many mapping between the pixels in the LSR input image and the pixels in the HSR target image. The chance that a region in the input image is real is determined by each value of the model activation map. The average of these values is used to generate an overall classification score for the input images. The architecture of our Hydro-DIS model is shown in Figure 6. A detailed summary of our Hydro-DIS model's architecture in Appendix B2.

2.3.3. Hydro-GAN Training Challenges

There are several challenges that perturb the training process of the Hydro-GAN model. One of the major issues is that the original minimax loss function can cause the Hydro-GAN to get stuck in a local minimum at the early stages of training when the Hydro-DIS discriminator's job is relatively easier than Hydro-GEN generator (Dwivedi, 2022). Ideally, Hydro-GAN should learn patterns represented in different water bodies' locations and avoid building expertise only in a subset of the training data distribution (Hui, 2018). The second challenge is the gradient vanishing problem that might occur if the Hydro-DIS discriminator performs significantly better than the generator (Dwivedi, 2022).

2.3.4. Hydro-GAN Loss Functions

To overcome the two aforementioned challenges, we propose to explore a new loss function. First, we modify the traditional GAN model defined in Equation 1 to be non-saturating. To achieve non-saturation, we use a variation of the standard loss function where instead of minimizing the $\log(1 - D(G(z)))$ in Equation 1, the generator maximizes the log of the discriminator probabilities that is, $-\log(D(G(z)))$. This change is inspired by depicting the problem from a different outlook, where the generator is trained to maximize the likelihood of images being real, instead of minimizing the likelihood of an image being synthetic. This avoids generator saturation through a more stable weight update mechanism (Dwivedi, 2022).

The second loss adjustment that we propose is to use a least-squares loss, where the Hydro-DIS is trained to effectively reduce the sum of the squared difference between generated and original values for real and synthetic images as defined in Equation 2. Similarly, the generator seeks to minimize the sum of the squared difference between predicted and expected values as though the generated images were real as defined in Equation 3. One advantage of using least-squares loss is that it produces a significant adjustment to the model in case of large errors, which helps in preventing the vanishing gradient (Mao et al., 2016).

$$\text{minimize}(D(z,y) - 1)^2 + (D(z,G(x,z)))^2 \quad (2)$$

$$\text{minimize}(D(z,G(x,z)) - 1)^2 \quad (3)$$

Along with using non-saturating and least-squares losses, we also propose a new DTW loss term that assesses the generator's accuracy in producing accurate boundary shapes. DTW algorithm is an elastic distance measure that has demonstrated good performance with sequence-based data, and in particular, time-series data (Muller, 2007). Our motivation for adding a DTW loss is that it emphasizes the polygon shape accuracy which equips the generator to better fool the discriminator. Our new generator loss, defined in Equation 4, is now a weighted sum of the cross-entropy adversarial loss ($\mathcal{L}_{adversarial}$), the L1 loss ($\mathcal{L}_{L1}$), and the DTW loss ($\mathcal{L}_{DTW}$).

$$G(x,z) = \mathcal{L}_{adversarial} + \mathcal{L}_{L1} + \mathcal{L}_{DTW} \quad (4)$$

where $\mathcal{L}_{adversarial}$ is the adversarial loss is the probabilistic loss which is also used in the discriminator. $\mathcal{L}_{L1}$ loss is the mean absolute error in the synthetic and the original images. $\mathcal{L}_{DTW}$ loss is our proposed loss which minimizes the difference between the Euclidean distance of generated and expected polygons from their respective centroids to the boundary coordinates.

We hypothesize that the traditional adversarial and L1 losses contributions are different than the DTW loss given that they optimize the Hydro-GEN and Hydro-DIS losses and the shape accuracy loss respectively. Therefore we weighted the loss terms using a β parameter than can be used to balance the terms.

$$G(x,z) = (100 - \beta) * (\mathcal{L}_{adversarial} + \mathcal{L}_{L1}) + \beta * \mathcal{L}_{DTW} \quad (5)$$

We evaluated the synthetic polygon accuracy using the Fréchet Inception Distance (FID), which is a measure for evaluating generative models (Y. Yu et al., 2021). The FID metric is defined as the squared Wasserstein metric between two multidimensional Gaussian distributions of the real and synthetic data. FID uses the Inceptionv3 model to show the similarity between two groups of images by using the computer vision features of the images (Szegedy et al., 2015). The FID is defined in Equation 6.

$$\text{FID} = \|\mu - \mu_w\|_2^2 + \text{tr}(\Sigma + \Sigma_w - 2(\Sigma^{1/2}\Sigma_w\Sigma^{1/2})^{1/2}) \quad (6)$$

where $\mathcal{N}(\mu, \Sigma)$ is the distribution of the neural network features of the images generated by the GAN model and $\mathcal{N}(\mu_w, \Sigma_w)$ is the distribution of the same neural network features of the real images used to train the GAN model. $\text{tr}()$ represents the trace of a matrix. The trace of a square matrix is the sum of its diagonal elements. μ is the mean vector of the neural network features obtained from the images generated by the GAN model and σ is the covariance matrix of the neural network features obtained from the images generated by the GAN model. A low FID score indicates that the two groups of images are similar, while a higher score suggests that the images have dissimilar characteristics.

2.4. Experimental Methodology

An ideal LSR to HSR mapping should produce a water body polygon that is similar in area, shape, and distance with respect to the original HSR polygon. To achieve this purpose, we evaluate the generated polygons against the ground truth polygons by using three criteria: areal, shape, and distance accuracy. Jaccard and Cosine measurements were chosen for assessing the areal precision. For evaluating the shape and distance measures, the

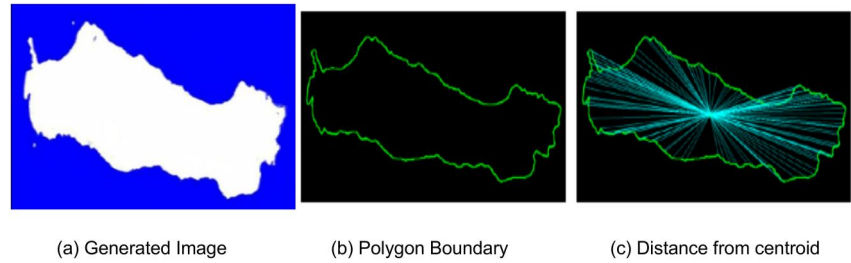


Figure 7. An input (a) water body image (Qapshaghay Bogeni Reservoir) used to extract the (b) water body polygon and (c) the shape signature based on the centroid distances.

DTW similarity metric was utilized. Each metric will be described in detail in the subsections that follow, along with its utility and drawbacks.

2.4.1. Areal Accuracy Measures

Areal accuracy assesses the correctness of the area-generated water bodies when compared to the area of the HSR ground truth. Prior to generating the area of the water bodies' polygons, we first used the contour detection tool in OpenCV that detects the boundary of the polygons. The polygon contours are considered as the curve joining all the continuous points (along the boundary), having the same color or intensity. The process of creating the polygon contour/boundary based on an input image is shown in the Figure 7b. The area similarity is computed using the Jaccard and Cosine similarity indices based on the extracted polygon boundaries.

2.4.2. Jaccard Similarity Index

The Jaccard similarity index is a metric that compares the area of two polygons in Euclidean space by quantifying the ratio of the shared area (intersection) between the two polygons with respect to their combined areas (Union):

$$J(P_{real}, P'_{predicted}) = \frac{Area(P_{real} \cap P'_{predicted})}{Area(P_{real} \cup P'_{predicted})} * 100 \quad (7)$$

This similarity metric follows the property of scale invariance, as it evaluates the size of the common area in relation to the combined area of the two polygons (Boubrahimi, Aydin, Kempton, & Angryk, 2016; Boubrahimi et al., 2018).

2.4.3. Cosine Similarity Index

The Cosine similarity index is a metric that compares the area of two polygons in Euclidean space by quantifying the ratio of the shared area (intersection) between the two polygons with respect to the square root of the product of the two areas (Boubrahimi, Aydin, Kempton, Mahajan, & Angryk, 2016; Boubrahimi et al., 2018).

$$C(P_{real}, P'_{predicted}) = \frac{Area(P_{real} \cap P'_{predicted})}{\sqrt{Area(P_{real} * P'_{predicted})}} * 100 \quad (8)$$

The areal similarity metrics range between zero and one hundred and the higher the metric values the better the areal match of the real (P_{real}) and generated ($P'_{predicted}$) polygons.

Algorithm 2. Shape Signature Derivation

Input: The polygon image *image*, and an image threshold value *thres*

Output: Distance array *dist*.

dist = []

binaryImage ← *getBinary*(*image*, *thres*)

▷ Convert images to binary format

contourImage ← *getContour*(*binaryImage*)

▷ Find Polygon contours

```

M1 = getCentroidCoordinates(contourImage) ▷ Calculate centroid points
coordinates
cx = getXCoordinate(M1)
cy = getYCoordinate(M1)
points ← getEdgeCoordinates(contourImage)
for point ∈ points do
    distance ← getDistance([cx, cy], [point[0], point[1]]) ▷ Calculating
distance from centroid
    dist ← append(distance)
end for
return (dist)

```

2.4.4. Shape and Distance Measure

To measure the shape and distance accuracy of the ground truth and generated polygons, we created a shape signature of the polygons. A shape signature is a sequence of time-ordered distance values between the polygon coordinates and a reference point in the plane. In this work, we consider the centroid of the polygon as the reference point. We retrieved the centroids of the water body polygons and created a time series of distances by using the Euclidean distance measure between the extracted centroid and the contour point coordinates (edge coordinates) in a clockwise manner. This time series was used to create the polygon shape signatures. Algorithm 2 describes the shape signature derivation algorithm. Figure 7c shows an example of the distance between the polygon centroid and its point coordinates. Finally, DTW was applied to compare the shape signatures of the generated and the real polygons.

2.4.5. Dynamic Time Warping Metric

DTW method is defined in Algorithm 3 (Giorgino, 2009). The two input shape signatures are organized in the form of a $n \times m$ matrix by using DTW algorithm where n and m are the lengths of the inputs. The lower left corner of the matrix is treated as the starting point for the matrix. The value of every cell in the matrix is calculated by measuring the distance between relative elements of the two input time series. As we move from the lower-left corner to the top-right corner of the matrix, we calculate the accumulated distance. The path that gives the minimum accumulated distance is considered as the ideal alignment between the two input time series. This path is referred to as the warping path and the total accumulated distance on this path is known as the warping distance. In an ideal situation, the input time series are identical and the warping path comes out to be a straight line from the lower-left corner to the top-right corner of the matrix, which results in a zero warping distance.

Algorithm 3. Algorithm for Finding Minimum-Cost Path Through a DTW Matrix.

```

Input: DTW distance matrix (dist_mat)
Output: minimum cost path (path), cost matrix (cost_mat).
    N, M = getShape(dist_mat) ▷ Initialize the cost
matrix
    cost_mat = [N, M]
    traceback_mat = np.zeros((N, M))
    for i ∈ range(N) do ▷ Fill the cost matrix
        for j ∈ range(M) do
            penalty = getPenalty(cost_mat)
            i_penalty = getIndexPenalty(penalty)
            cost_mat = updateMat(dist_mat, penalty)
            traceback_mat = updateTraceback(i_penalty)
        end for
    end for
    i = N - 1, j = M - 1, path = [(i, j)]
    while i > 0 OR j > 0 do ▷ Traceback from bottom right

```

```

    traceback_type = traceback_mat[i, j]
    IF(traceback_type == 0), Then -> i = i - 1, j = j - 1
    IF(traceback_type == 1), Then -> i = i - 1
    IF(traceback_type == 2), Then -> j = j - 1
    path ← append((i, j))
  end while
  cost_mat = updateCostMatrix(cost_mat)
  return path, cost_mat

```

The mathematical form of DTW algorithm as accepted in the literature is shown in Equation 9. N and M are the lengths of the input time series x and y . Initially the D matrix is initialized to $D_{0,0} = 0$ and $D_{i,j} = \infty$. The square of the differences between y_j and x_i is generally used as the cost function g in Equation 9.

$$D_{i,j} = g(x_i, y_j) + \min\{D_{i,j-1}, D_{i-1,j}, D_{i-1,j-1}\} \quad (9)$$

s.t : $i \in (1, M), j \in (1, N)$

In addition to using DTW as a loss term, we used DTW as a dissimilarity measure by calculating the optimal matching between the two shape signatures. We used DTW to match the points of the polygon by aligning their centroid shape signatures. The DTW algorithm generates a warping distance between the two input time series, which indicates the discrepancy between the two shapes. The greater the warping distance, the greater the difference between the two polygon shape signatures.

Table 1 shows the summary of the areal and shape evaluation metrics that were used in this paper, along with their advantages and drawbacks.

2.5. Data Set Preparation

Prior to training Hydro-GAN, we created image data pairs that map each curated MODIS image and its corresponding curated Landsat image that was captured during the same day. The data set was split into a training set and a testing set. We reserved 90% of the image data pairs from each reservoir for the training, and the remaining 10% image data pairs were used for testing. Namely, 672 out of 6,720 MODIS satellite images and 336 out of 3,360 Landsat-8 images were included in the test set.

3. Results and Discussion

3.1. Hydro-GAN Hyper-Parameter Search

To identify the optimal weights that produce the highest polygon accuracy based on FID, we performed a grid search of the weight parameters, defined in Equation 5. We trained our proposed hydro-GAN model on different weight ratios. Figure 8 shows a heatmap that summarizes the FID scores obtained by each model when trained on different β weights, ranging from 1 to 100 for balancing the $\mathcal{L}_{adversarial}$, $\mathcal{L}_{L1}$ and $\mathcal{L}_{DTW}$ losses.

The result in Figure 8 reveals that when more weight is assigned to the DTW loss, the Hydro-GAN model generates images that are more similar to the target images. This is reflected by the low FID scores that are approximating a value of zero. Specifically, the lowest FID scores (of 0.2) were obtained in the top right corner,

Table 1
Summary of Areal and Shape Evaluation Metrics

Evaluation measure	Advantages	Drawbacks	Interval (%)
Jaccard similarity	Effective when evaluating polygons with comparable properties	Operates poorly with polygons that are very large or small	[0, 100]
Cosine similarity	Effective for evaluating the areal similarity between inconsistent polygons	Does not penalize shapes that have highly different areas	[0, 100]
DTW	Effective for shape similarity when starting point is carefully chosen	Highly dependent on starting point selection and requires shape alignment	[0, ∞]

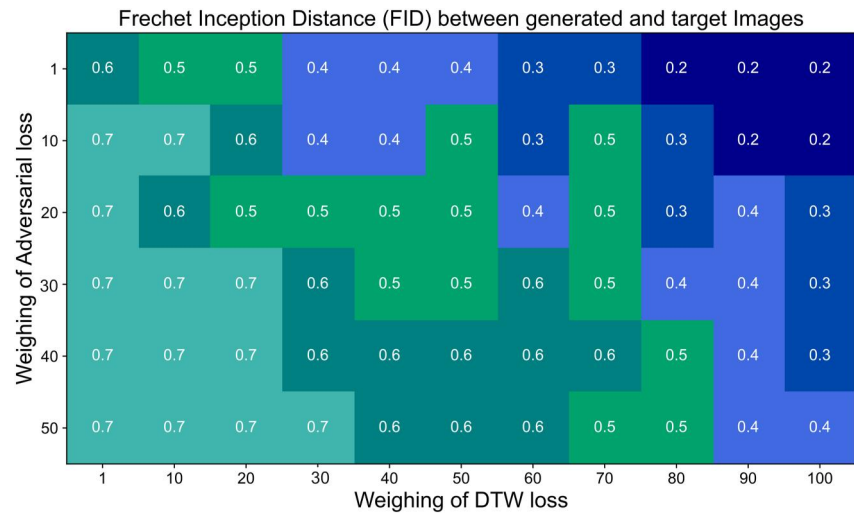


Figure 8. Fréchet Inception Distance (FID) metric of Hydro-GAN generated polygons with β values ranging from 0 to 100. The lower the FID value, the better (i.e., the generated and target polygons are similar).

where $\beta \in [90 - 100]$, which indicates that the generated images are significantly similar to the ground truth. Following the same line of thought, when the β weight is small, the FID scores are high which signifies that the Hydro-GAN loss is overpowered by the adversarial and L1 losses (refer to the bottom left corner of Figure 8). Selecting an equal weight ($\beta = 50$) for the adversarial and distance losses does not achieve optimality. From the aforementioned observations, we trained our Hydro-GAN model using the β value of 100.

3.2. Hydro-GAN Training Result

After identifying the optimal weight ratios from Figure 8, we evaluate our proposed Hydro-GAN model based on the generated HSR images from the LSR MODIS image inputs. We compare the HSR-generated output with the original HSR Landsat ground truth images. Figure 9 shows the evolution of the input LSR image after being fed to the Hydro-GAN model in comparison with the ground truth HSR image captured by Landsat. Figure 9a shows an input LSR MODIS image fed to the Hydro-GAN model and a temporary results that were learned after 30 epochs (Figure 9b) and after 100 epochs (Figure 9c). The results indicate that the model learning curve is improving when comparing the shape and area of the generated polygons after 30 and 100 epochs. After 100 iterations, the generated water body polygon morphology is similar to the polygon contained in the ground truth HSR Landsat image as shown in Figure 9d. After training our Hydro-GAN model for 100 epochs, we measured the accuracy of the model by conducting a quantitative analysis of the test set images based on three evaluation criteria: areal, shape, and distance accuracy as discussed in Section 2.4.

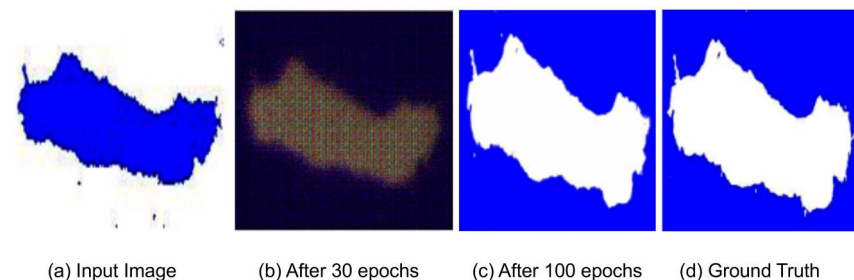


Figure 9. (a) Example input MODIS image from Qapshaghay Bogeni Reservoir (Kazakhstan) fed to Hydro-GAN and transformed into (b) an high spatial resolution after 30 training epochs (2) and (c) 100 training epochs compared to the (d) ground truth Landsat image.

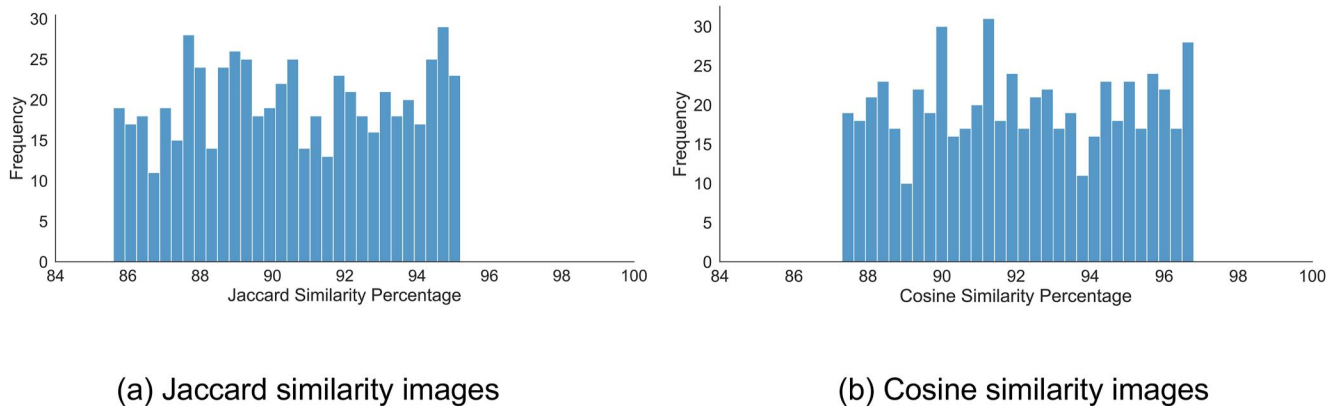


Figure 10. Distribution of area similarity measured on the test set images by plotting histogram of Jaccard and Cosine similarity percentage between real and generated polygons in the test set images.

3.2.1. Areal Accuracy Result

To measure the areal accuracy between the generated polygons (from the Hydro-GAN model using the test data set) and the ground truth HSR Landsat polygons, we used the Jaccard and Cosine similarity percentages as illustrated in Figure 10. The figure shows the histogram plot of Jaccard and Cosine similarity metrics when assessed on the entire testing data of LSR-HSR image pairs. The results of the Jaccard distribution indicate a good agreement between the generated and the original Landsat polygons with a high similarity percentage, ranging approximately between 86% and 95%. A similar result was also obtained from the Cosine histogram distribution where the values span approximately from 87% to 97%. This indicates that there is a strong correlation between Jaccard and Cosine metrics and implies that the Hydro-GAN model is performing well, resulting in highly accurate synthetic HSR polygons. Although the two metrics are different, as outlined in Table 1, one explanation for this correlation is that the numerators of both the areal metrics are the same (i.e., area of intersection of the two polygons), as demonstrated in Equations 7 and 8.

3.2.2. Shape and Distance Accuracy Result

To measure the shape accuracy between the actual and the generated polygons from the test set, the polygons were converted into their respective shape signatures by calculating the Euclidean distance from the centroid to the coordinates of the points on the edges of the polygon and converting these distances into a time series. The result of which was used to plot a similarity graph as shown in Figure 11a. As shown in the Figure 11a, the result indicates that when there is a dip in the shape signature of the ground truth polygon around 300 and 700 indices, there is a similar dip in the shape signature of the generated polygon as well. The same observation can also be made about the rise in the shape signature of the polygons being compared around indices 0, 400, and 1,000. It can also be seen that the shape signatures of the generated and the ground truth polygon are nearly overlapping. This indicates that our Hydro-GAN model is producing HSR images that are highly accurate in shape signature with respect to the reference Landsat images.

To measure the distance accuracy, we used DTW to measure the extent of alignment between the actual and the generated polygon shape signatures. This process was repeated for all the images in the test set. To illustrate the result we chose a random LSR image from the test set to generate an HSR Landsat image. The generated image was then compared with the ground truth HSR image and the DTW was applied to their respective polygons. The result can be seen in Figure 11b which shows the DTW alignment and warping matrix that were produced when comparing the two polygons. The graph in Figure 11b shows that the warping path (denoted by the blue line from the lower-left to the top-right corner) between the generated and actual polygons is approximating a diagonal line (which is the ideal case). This implies that the HSR images produced from our Hydro-GAN model are highly accurate in the distance metric as well.

We evaluate the distance metric of the entire test set by calculating the normalized alignment cost of the DTW matrix between all the generated and actual polygons. A normalized alignment cost of 1 is considered the ideal case. The result of this evaluation is reported in Figure 12, which shows that the distribution of the normalized

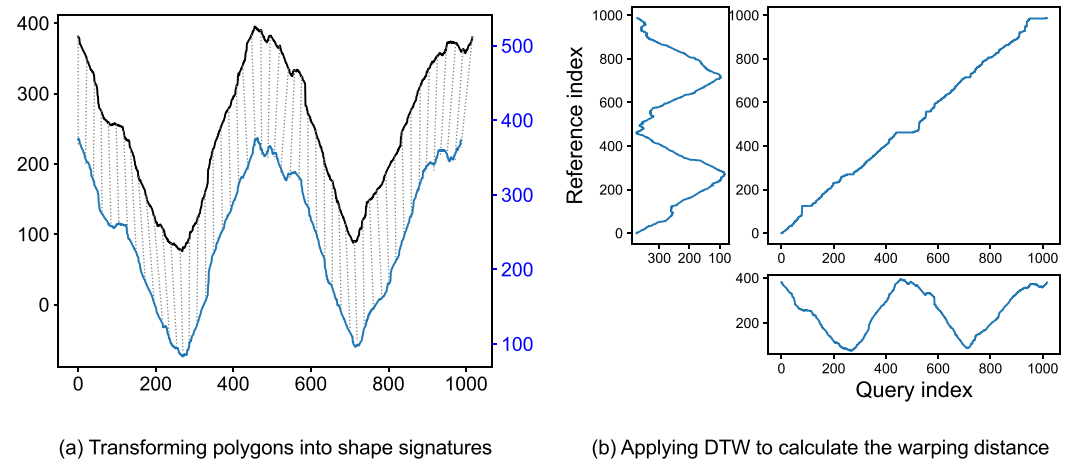


Figure 11. Dynamic time warping measure applied by (a) aligning two polygons shape signatures (Ground truth in blue and generated in black) and (b) computing their warping distance by finding the optimal path along the warping matrix diagonal. Query Index refers to the index of the shape signature in the generated polygon, while Reference Index refers to the index of the shape signature in the ground truth (reference) polygon.

alignment cost for the test set images ranges between 1.13 and 1.51, and the average alignment cost is 1.31 which are all close to the ideal value of 1. This indicates that our Hydro-GAN model is producing HSR images that have lower DTW distances when compared with the actual ground truth.

To better understand how the three areal and distance measures are correlated with each other, we present their correlation matrix which is shown in Figure 13. We note that the Jaccard and Cosine metrics have a strong positive correlation (0.96) which reveals that when there is an increase in the values of Jaccard similarity, the Cosine similarity also increases. This is due to the common numerator shared between Jaccard and Cosine. It can also be inferred from Figure 13 that the DTW alignment cost generates a moderate strength negative correlation with the Jaccard and Cosine similarities (around -0.50). This result indicates that in the cases when the Hydro-GAN model produces highly accurate areas, the exact shapes of the polygons are not highly accurate. Similarly, when Hydro-GAN produces highly underestimated (or overestimated) areas, the water body shape is relatively more accurate. The finding of this experiment is important for fine-tuning the model depending on the hydrology research needs. For example, if the purpose of generating HSR images is to study the sedimentation phenomena that require a precise polygon boundary, then Hydro-GAN can be tuned for optimizing $\mathcal{L}_{L1}$ and $\mathcal{L}_{DTW}$ losses. In case the research needs to study the water volume change, then introducing a new areal loss $\mathcal{L}_{area}$ could be useful.

In recent years, various state-of-the-art image-to-image translation models have been developed to address the challenges of transforming images from one domain to another while preserving their essential features (X. Li et al., 2021). Here, we briefly introduce three prominent models. SPA-GAN, or Spatial Attention GAN, is a model

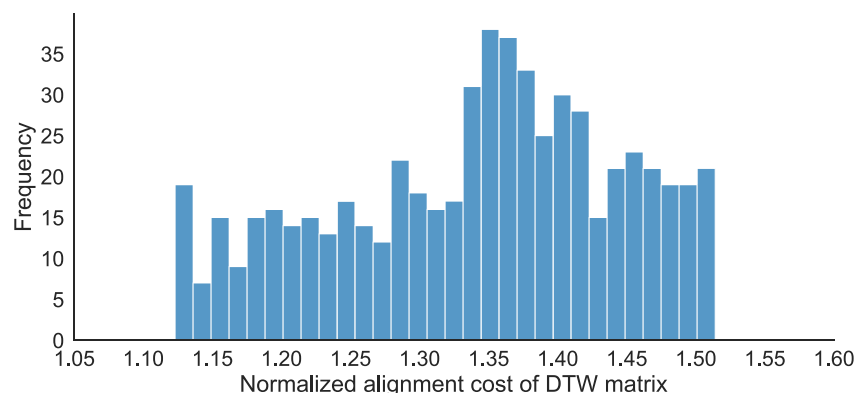


Figure 12. The normalized alignment cost for dynamic time warping matrix in test set images.

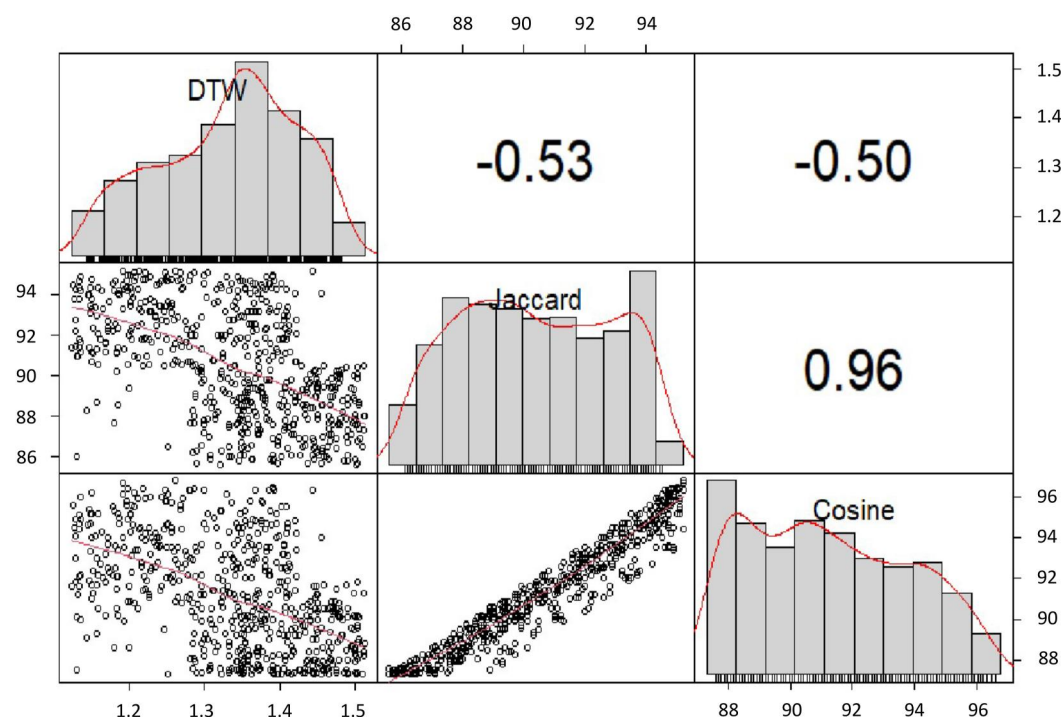


Figure 13. Correlation matrix of Jaccard, Cosine and dynamic time warping similarity measures used for area, shape and distance accuracy evaluation, when measured between generated and original high spatial resolution test set images.

designed to enhance image-to-image translation by incorporating spatial attention mechanisms (Emami et al., 2020). It focuses on selectively attending to specific regions of the input image during the translation process, allowing for more precise and context-aware transformations.

Cycle-GAN, on the other hand, is known for its ability to perform unpaired image translation. It introduces a cycle consistency loss to ensure that the translation from one domain to another and back results in the original input, enforcing a cycle-consistent mapping between domains without the need for paired data during training (Zhu et al., 2017). Pix2Pix is a conditional GAN that excels in paired image-to-image translation tasks (Isola et al., 2017). It learns a mapping from input to output images by training on aligned pairs, making it particularly effective for tasks where there is a clear correspondence between the input and output images.

In our experimental comparison, we evaluated these models alongside your proposed Hydro-GAN using the Jaccard Index. The Jaccard Index measures the similarity between two sets by dividing the size of their intersection by the size of their union. As shown in Figure 14, the results indicate that Hydro-GAN outperformed the other models with a Jaccard Index of 0.98, showcasing its effectiveness in image-to-image translation. This comparison provides a comprehensive assessment of Hydro-GAN's performance relative to existing state-of-the-art models, contributing to a thorough validation of the proposed approach.

4. Case Study: Lake Tharthar's

In this case study, we evaluated our proposed Hydro-GAN model on Lake Tharthar's shrinking and expansion behaviors. Lake Tharthar's is a focal point for hydrological research and water resource management in Iraq (Kornijów et al., 2001). Covering an area of approximately 2,100 square kilometers (810 square miles) at its maximum extent, it is one of the largest lakes in the country. Its seasonal dynamics, ecological significance, and role in sustaining water supply make it a critical area of study for hydrologists and environmental scientists aiming to address water-related challenges in the region.

Our study area consists of Lake Tharthar's polygon area across a period of 7 years, from 2015 to 2021. We compared the area of the lake water body from three different image sources. First, the original area of the water body was calculated from the original HSR Landsat images. Second, the area of the water body was extracted

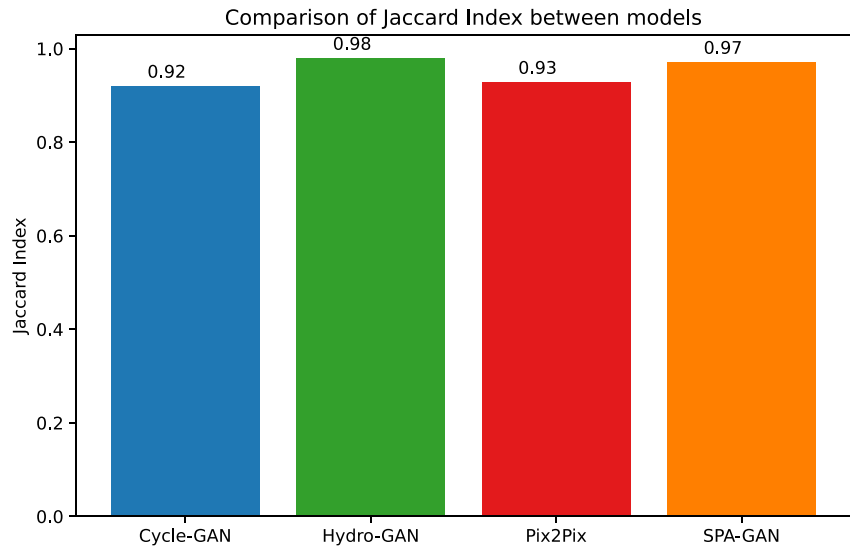
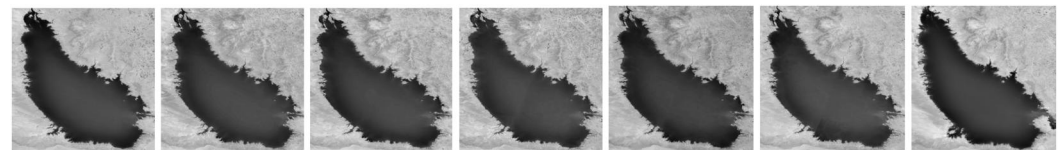


Figure 14. Performance of the state-of-the-art models.

from the generated Landsat images (using our Hydro-GAN model). Third, the area of the water body was calculated using the LSR MODIS satellite images. The three polygon areas were compared with each other as plotted in Figure 15. The blue bar plot in Figure 15 indicates that Lake Tharthar's (one of the reservoirs from the data set) first shrinks in surface area from 2015 to 2020 with a little increase in the year 2018, then shrinks again



(a) 2015 (b) 2016 (c) 2017 (d) 2018 (e) 2019 (f) 2020 (g) 2021

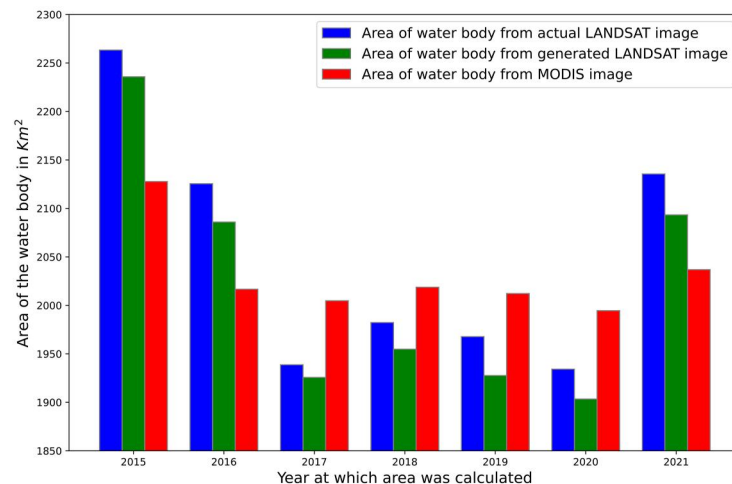


Figure 15. (a–g) Snapshot of Lake Tharthar's water body for each year from 2015 to 2021, providing a visual record of changes over time. The bar plots compare area variation of Lake Tharthar's between generated and actual polygons across 7 years from 2015 to 2021.

until 2020. The area of the water body again increases in the year 2021. The shrinking of the water body can be due to numerous hydrological as well as human factors, such as sedimentation as well as an increase in freshwater consumption (Mahajan & Martinez, 2021).

The increase in the area of the water body in 2021 can be seen in Figure 15g where a small portion of the surface boundary has increased on the top left corner and bottom compared to 2020. It can also be noticed that the images produced by our Hydro-GAN model also follow the same pattern of change in the real ground-truth area of a water body across the 7 years (shown in the green bar plot). This result indicates that our Hydro-GAN model is predicting the shrinking or expansion of the surface area of a water body accurately. Another important observation that can be noted from Figure 15 is that the area calculated from the LSR MODIS images (i.e., our competing baseline), shown in the red bar plot, fails to follow the same pattern of change in the water body accurately. We note that the LSR MODIS extracted area sometimes overestimates the actual area as reported in the years 2017, 2018, 2019, and 2020. Similarly, the LSR MODIS extracted area shows an under-estimation for the years 2015, 2016, and 2021. Our converted LSR to HSR extracted areas, shown in the blue bar plot of Figure 15, are more realistic area estimations that produce significantly lower errors than LSR MODIS. This result shows that our proposed LSR to HSR conversion using Hydro-GAN improved the prediction of the area of the Lake Tharthar's water body. In addition, we note that our model produces generally a slight under-estimation of the area. Following our suggestion in Section 3.2.2, using an areal loss can correct the under-estimation.

5. Conclusion and Future Work

In this paper, we propose Hydro-GAN a deep learning-based generative method that uses a cGAN for mapping the remote sensing information available at low resolution (MODIS satellite) to high resolution (Landsat satellite). In particular, we have used the case study of water bodies and reservoirs. Our proposed method uses a weighted DTW loss function along with the adversarial and L1 loss, to generate the HSR-LSR mapping. The results were evaluated using areal, shape, and distance similarity measures. The evaluation shows that our weighted Hydro-GAN model improved the accuracy of the generated water bodies polygons compared to state-of-the-art GAN models. Since the availability of accurate data on water bodies' boundaries at high spatial and temporal resolution is important for assessing the role it plays in multiple hydrological research tasks, our work can provide complementary data sets for hydrological studies. Hydro-GAN can generate high-resolution data at historical time steps when such data is unavailable which can be used in areas where a large amount of historical data is required for forecasting purposes. As future work, we aim to further improve our mapping by making enhancements to our computer vision algorithm (which extracts the polygon boundary), so that it can extract the polygon even when the images obtained from the remote sensing satellites contain noise like clouds or other distortions. We also aim to extend the domain of our method beyond water bodies, like forests and vegetation. Finally, we plan to make our model generic, so that it can be used to map the images from any remote sensing satellite to another, which can include Sentinel and Hyperion satellites as well.

Appendix A: Data

A1. Description of the Water Bodies

The remote sensing data set from the MODIS and Landsat satellites is available courtesy of Earthdata (<https://earthdata.nasa.gov/>) which provides open access to the data for analysis and research purposes. The data set used in this research consists of 20 reservoirs with a time span of 7 years from 2015 to 2021. The metadata associated with these reservoirs can be found in the table below.

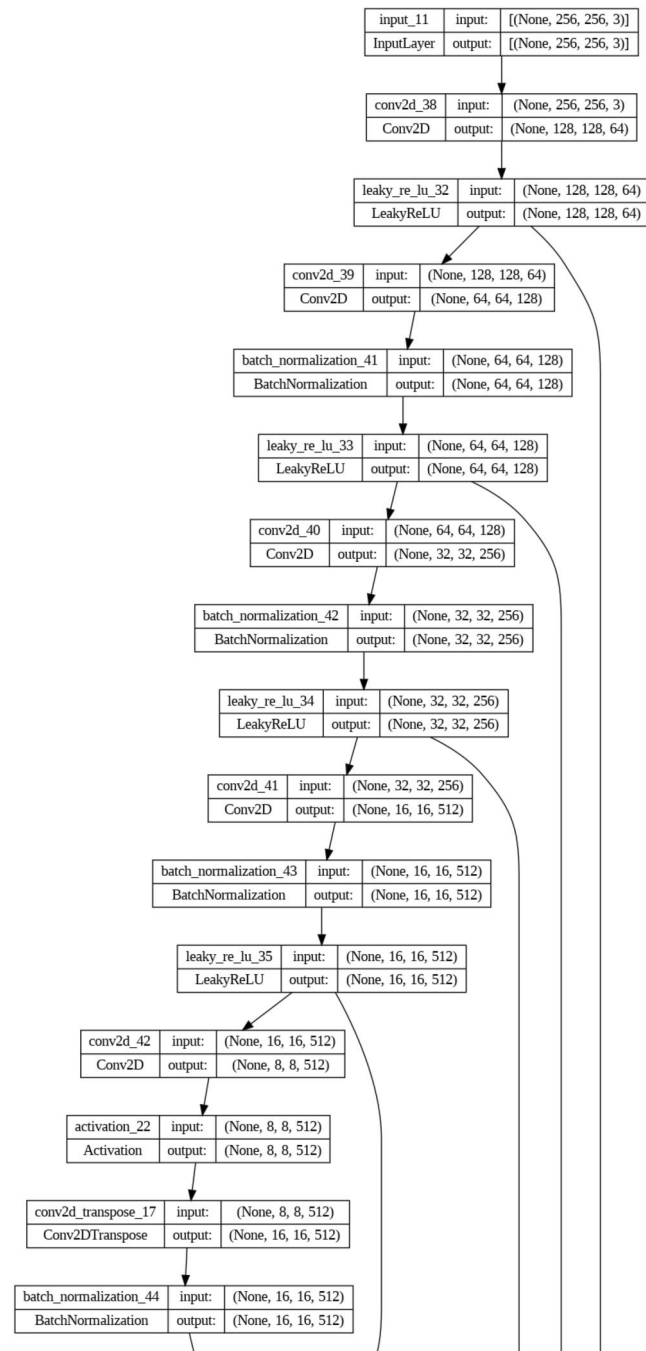
Lake ID	Reservoir name	Longitude, latitude (°)	Country	Continent	Area capacity (KM2)	Storage capacity (KM3)
1	Lake Nasser	32.89, 23.97	Egypt	Africa	6,500	162
2	Kariba Reservoir	28.76, −16.52	Zambia	Africa	5,400	180
3	Lake Zaysan	83.35, 49.66	Kazakhstan	Asia	5,490	49.8
4	Cahora Bassa Reservoir	32.7, −15.58	Mozambique	Africa	2,739	55.8
5	Qapshaghay Bogeni Reservoir	77.1, 43.92	Kazakhstan	Asia	1,850	28.1

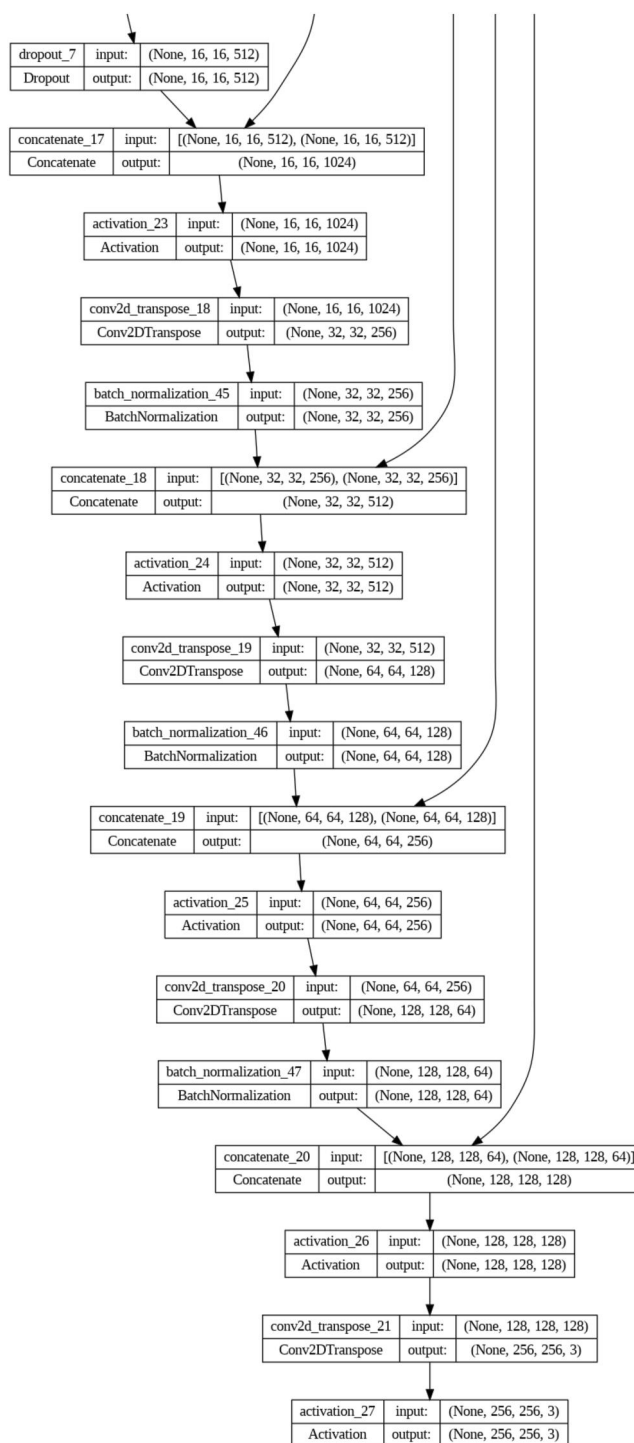
Continued

6	Fort Berthold Reservoir	−101.43, 47.51	United States of America	North America	1,477.4	29.38
7	Okeechobee reservoir	−81.1, 26.94	United States of America	North America	1,536.8	5.2
8	Lake Volta	0.06, 6.3	Ghana	Africa	8,502	148
9	Guri Reservoir	−63, 7.77	Venezuela	South America	4,250	135
10	Lake Mead (Colorado river)	36.15, −114.41	United States of America	North America	659.3	32.236
11	Lake Erie	42.13, −81.28	United States of America	North America	25,667	480
12	Fort Peck lake	47.82, −106.58	United States of America	North America	991.47	23.05
13	Lake Tanganyika	−6.42, 29.63	Tanzania	Africa	32,900	189
14	Lake Argyle	128.74, −16.12	Australia	Oceania	981.21	10.76
15	La Angostura	16.07, −92.48	Mexico	North America	566.3	145
16	Lake Oahe (Missouri river)	45.29, −100.30	United States of America	North America	1,092.6	123
17	Lake Tharthar's	43.46, 33.69	Iraq	Asia	2,135.54	85.59
18	Chardarinskoye	67.96, 41.25	Kazakhstan	Asia	800.66	5.7
19	Furnas	−46.31, −20.67	Brazil	South America	1,127.07	22.59
20	Flathead Lake	−114.23, 47.68	United States of America	North America	5,102	3.2

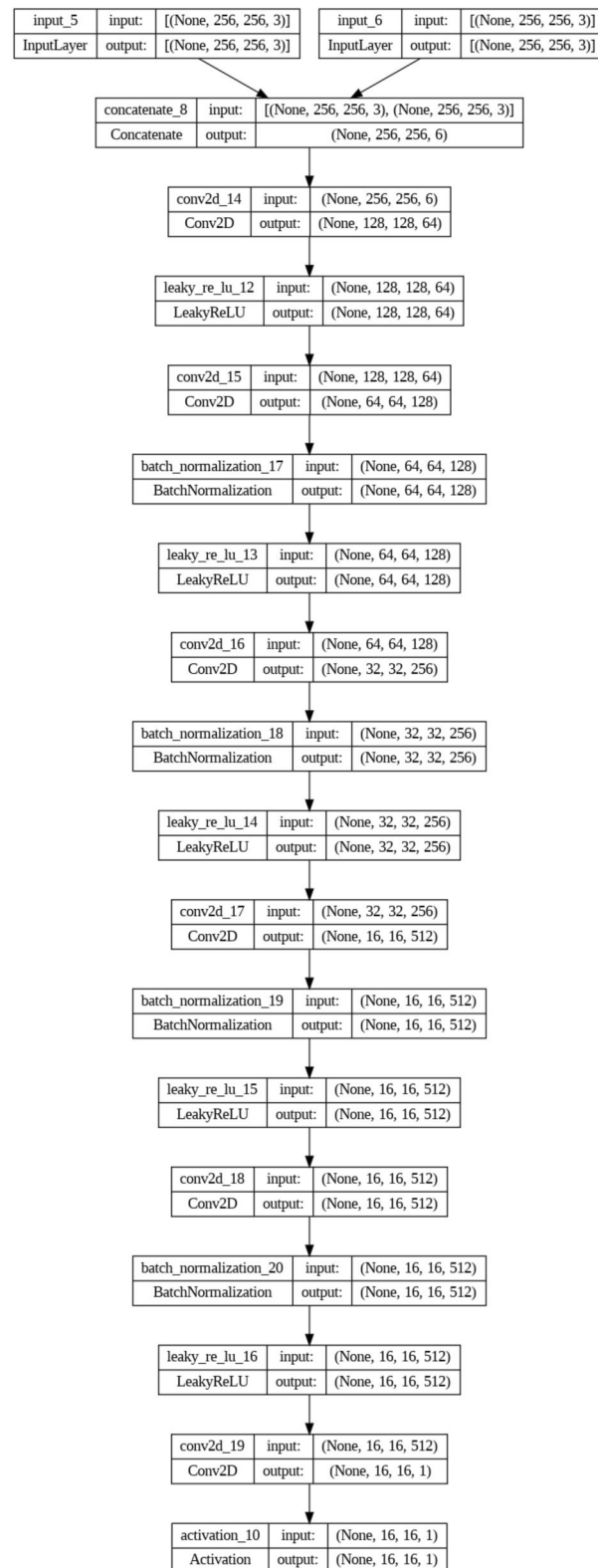
Appendix B: Model Summaries

B1. Detailed Summary of the Hydro-GEN





B2. Detailed Summary of Hydro-DIS



Data Availability Statement

The Hydro-GAN model source code and experimental evaluations are available via (<https://zenodo.org/uploads/10005635>). The remote sensing data from MODIS and Landsat satellites is available via (<https://earthdata.nasa.gov/>) and (<https://zenodo.org/uploads/10005635>).

Acknowledgments

This project has been supported in part by funding from the Division of Atmospheric and Geospace Sciences within the Directorate for Geosciences, under NSF awards #2204363, #2240022, and #2301397 and by funding from the Office of Advanced Cyberinfrastructure within the Directorate for Computer and Information Science and Engineering, under NSF awards #2305781 and #2118329.

References

- Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., et al. (2009). Lakes as sentinels of climate change. *Limnology & Oceanography*, 54(6part2), 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283
- Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020). A novel framework for spatio-temporal prediction of environmental data using deep learning. *Scientific Reports*, 10(1), 22243. <https://doi.org/10.1038/s41598-020-79148-7>
- Barnes, W., Xiong, X., & Salomonson, V. (2003). Status of terra modis and aqua modis. *Advances in Space Research*, 32(11), 2099–2106. [https://doi.org/10.1016/s0273-1177\(03\)90529-1](https://doi.org/10.1016/s0273-1177(03)90529-1)
- Boubrahimi, S. F., Aydin, B., Kempton, D., & Angryk, R. (2016). Spatio-temporal interpolation methods for solar events metadata. In *2016 IEEE international conference on big data (Big Data)*. <https://doi.org/10.1109/bigdata.2016.7840970>
- Boubrahimi, S. F., Aydin, B., Kempton, D., Mahajan, S. S., & Angryk, R. (2016). Filling the gaps in solar big data: Interpolation of solar filament event instances. In *2016 IEEE international conferences on big data and cloud computing (BDCLOUD), social computing and networking (SocialCom), sustainable computing and communications (SustainCom) (BDCLOUD-SocialCom-SustainCom)*. <https://doi.org/10.1109/bdcloud-socialcom-sustaincom.2016.25>
- Boubrahimi, S. F., Aydin, B., Schuh, M. A., Kempton, D., Angryk, R. A., & Ma, R. (2018). Spatiotemporal interpolation methods for solar event trajectories. *The Astrophysical Journal Supplement Series*, 236(1), 23. <https://doi.org/10.3847/1538-4365/aab763>
- Campbell, J., & Wynne, R. (2011). *Introduction to remote sensing*. Guilford Press.
- Chen, T.-W., Chen, Y.-L., & Chien, S.-Y. (2008). Fast image segmentation based on k-means clustering with histograms in HSV color space. In *2008 IEEE 10th workshop on multimedia signal processing* (pp. 322–325). <https://doi.org/10.1109/MMSP.2008.4665097>
- Chen, Y., Fan, R., Yang, X., Wang, J., & Latif, A. (2018). Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water*, 10(5), 585. <https://doi.org/10.3390/w10050585>
- Chen, Y., Huang, C., Ticehurst, C., Merrin, L., & Thew, P. (2013). An evaluation of modis daily and 8-day composite products for floodplain and wetland inundation mapping. *Wetlands*, 33(5), 823–835. <https://doi.org/10.1007/s13157-013-0439-4>
- Comer, M. L., & Delp, E. J., III. (1999). Morphological operations for color image processing. *Journal of Electronic Imaging*, 8(3), 279–289. <https://doi.org/10.1117/1.482677>
- Dwivedi, H. (2022). Understanding GAN loss functions. Retrieved from <https://neptune.ai/blog/gan-loss-functions>
- Emami, H., Aliabadi, M. M., Dong, M., & Chinnam, R. B. (2020). Spa-gan: Spatial attention gan for image-to-image translation. *IEEE Transactions on Multimedia*, 23, 391–401. <https://doi.org/10.1109/tmm.2020.2975961>
- Frazier, P. S., & Page, K. J. (2000). Water body detection and delineation with landsat tm data. *Photogrammetric Engineering & Remote Sensing*, 66(12), 1461–1468.
- Fu, W., Ma, J., Chen, P., & Chen, F. (2020). Remote sensing satellites for digital earth. In *Manual of digital earth* (pp. 55–123). Springer.
- Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of Statistical Software*, 31(7), 1–24. <https://doi.org/10.18637/jss.v031.i07>
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial networks. *arXiv*. <https://doi.org/10.48550/ARXIV.1406.2661>
- Haghnazar, H., Abbasi, Y., Morovati, R., Johannesson, K. H., Somma, R., Pourakbar, M., & Aghayani, E. (2024). Polycyclic aromatic hydrocarbons (PAHs) in the surficial sediments of the Abadan freshwater resources-northwest of the Persian gulf. *Journal of Geochemical Exploration*, 258, 107390. <https://doi.org/10.1016/j.gexplo.2024.107390>
- Hosseinzadeh, P., Boubrahimi, S. F., & Hamdi, S. M. (2024). Improving solar energetic particle event prediction through multivariate time series data augmentation. *The Astrophysical Journal Supplement Series*, 270(2), 31. <https://doi.org/10.3847/1538-4365/ad1de0>
- Hosseinzadeh, P., Nassar, A., Boubrahimi, S. F., & Hamdi, S. M. (2023). ML-based streamflow prediction in the upper Colorado river basin using climate variables time series data. *Hydrology*, 10(2), 29. <https://doi.org/10.3390/hydrology10020029>
- Hui, J. (2018). Generative-adversarial-networks-and-its-application. Retrieved from <https://jonathan-hui.medium.com/gan-whats-generative-adversarial-networks-and-its-application-f39ed278ef09>
- Ismail, M. A., Waqas, M., Ali, A., Muzzamil, M. M., Abid, U., & Zia, T. (2021). Enhanced index for water body delineation and area calculation using Google earth engine: A case study of the Manchar Lake. *Journal of Water and Climate Change*, 13(2), 557–573. <https://doi.org/10.2166/wcc.2021.282>
- Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2016). Image-to-image translation with conditional adversarial networks. In *CoRR*, abs/1611.07004. Retrieved from <http://arxiv.org/abs/1611.07004>
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125–1134).
- Jiang, W., He, G., Long, T., Ni, Y., Liu, H., Peng, Y., et al. (2018). Multilayer perceptron neural network for surface water extraction in landsat 8 oii satellite images. *Remote Sensing*, 10(5), 755. <https://doi.org/10.3390/rs10050755>
- Justice, C., Townshend, J., Vermote, E., Masuoka, E., Wolfe, R., Saleous, N., et al. (2002). An overview of modis land data processing and product status. *Remote Sensing of Environment*, 83(1), 3–15. [https://doi.org/10.1016/S0034-4257\(02\)00084-6](https://doi.org/10.1016/S0034-4257(02)00084-6)
- Khandelwal, A., Karpatne, A., & Kumar, V. (2017). ORBIT: Ordering based information transfer across space and time for global surface water monitoring. In *CoRR*, abs/1711.05799. Retrieved from <http://arxiv.org/abs/1711.05799>
- Khandelwal, A., Karpatne, A., Marlier, M. E., Kim, J., Lettenmaier, D. P., & Kumar, V. (2017). An approach for global monitoring of surface water extent variations in reservoirs using modis data. *Remote Sensing of Environment*, 202, 113–128. (Big Remotely Sensed Data: tools, applications and experiences). <https://doi.org/10.1016/j.rse.2017.05.039>
- Kornijow, R., Szczerbowski, J. A., Krzywosz, T., & Bartel, R. (2001). The macrozoobenthos of the iraqi lakes tharthar, habbaniya and razzazah. *Archives of Polish Fisheries*, 9(10), 27–145.

- Leonard, C., Legleiter, C., Lea, D., & Schmidt, J. (2020). Measuring channel planform change from image time series: A generalizable, spatially distributed, probabilistic method for quantifying uncertainty. *Earth Surface Processes and Landforms*, 45(11), 2727–2744. <https://doi.org/10.1002/esp.4926>
- Li, C., Chang, W., Cheng, Y., Yang, Y., & Póczos, B. (2017). MMD GAN: Towards deeper understanding of moment matching network. In *CoRR*, *abs/1705.08584*. Retrieved from <http://arxiv.org/abs/1705.08584>
- Li, X., Du, Z., Huang, Y., & Tan, Z. (2021). A deep translation (GAN) based change detection network for optical and SAR remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179, 14–34. <https://doi.org/10.1016/j.isprsjprs.2021.07.007>
- Loesdau, M., Chabrier, S., & Gabillon, A. (2014). Hue and saturation in the RGB color space. In A. Elmoataz, O. Lezoray, F. Nouboud, & D. Mammass (Eds.), *Image and signal processing* (pp. 203–212). Springer International Publishing.
- Luo, W., Li, Y., Urtasun, R., & Zemel, R. S. (2017). Understanding the effective receptive field in deep convolutional neural networks. In *CoRR*, *abs/1701.04128*. Retrieved from <http://arxiv.org/abs/1701.04128>
- Mahajan, S., & Martinez, J. (2021). Water, water, but not everywhere: Analysis of shrinking water bodies using open access satellite data. *The International Journal of Sustainable Development and World Ecology*, 28(4), 326–338. <https://doi.org/10.1080/13504509.2020.1851803>
- Maiti, S., & Bhattacharya, A. K. (2009). Shoreline change analysis and its application to prediction: A remote sensing and statistics based approach. *Marine Geology*, 257(1), 11–23. <https://doi.org/10.1016/j.margeo.2008.10.006>
- Mao, X., Li, Q., Xie, H., Lau, R. Y. K., & Wang, Z. (2016). Multi-class generative adversarial networks with the L2 loss function. In *CoRR*, *abs/1611.04076*. Retrieved from <http://arxiv.org/abs/1611.04076>
- Martinez, M. T., & Heiner, O. N. (2020). Conditional generative adversarial networks for solving heat transfer problems. <https://doi.org/10.2172/1673172>
- Muller, M. (2007). Dynamic time warping. *Information Retrieval for Music and Motion*, 2, 69–84. https://doi.org/10.1007/978-3-540-74048-3_4
- Muster, S., Heim, B., Abnizova, A., & Boike, J. (2013). Water body distributions across scales: A remote sensing based comparison of three arctic tundra wetlands. *Remote Sensing*, 5(4), 1498–1523. <https://doi.org/10.3390/rs5041498>
- Njue, N., Stenfort Kroese, J., Gräff, J., Jacobs, S., Weeser, B., Breuer, L., & Rufino, M. (2019). Citizen science in hydrological monitoring and ecosystem services management: State of the art and future prospects. *Science of the Total Environment*, 693, 133531. <https://doi.org/10.1016/j.scitotenv.2019.07.337>
- Ouma, Y. O., & Tateishi, R. (2006). A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: An empirical analysis using landsat TM and ETM+ data. *International Journal of Remote Sensing*, 27(15), 3153–3181. <https://doi.org/10.1080/01431160500309934>
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Qiao, C., Luo, J., Sheng, Y., Shen, Z., Zhu, Z., & Ming, D. (2012). An adaptive water extraction method from remote sensing image based on NDWI. *Journal of the Indian Society of Remote Sensing*, 40(3), 421–433. <https://doi.org/10.1007/s12524-011-0162-7>
- Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sensing*, 6(5), 4173–4189. <https://doi.org/10.3390/rs6054173>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *CoRR*, *abs/1505.04597*. Retrieved from <http://arxiv.org/abs/1505.04597>
- Said, K. A. M., Jambek, A. B., & Sulaiman, N. (2016). A study of image processing using morphological opening and closing processes. *International Journal of Control Theory and Applications*, 9(31), 15–21.
- Sarp, G., & Ozelik, M. (2017). Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey. *Journal of Taibah University for Science*, 11(3), 381–391. <https://doi.org/10.1016/j.jtusci.2016.04.005>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015). Rethinking the inception architecture for computer vision. In *CoRR*, *abs/1512.00567*. Retrieved from <http://arxiv.org/abs/1512.00567>
- Tran, D., Ray, J., Shou, Z., Chang, S., & Paluri, M. (2017). Convent architecture search for spatiotemporal feature learning. In *CoRR*, *abs/1708.05038*. Retrieved from <http://arxiv.org/abs/1708.05038>
- Verma, U., Chauhan, A., Manohara Pai, M. M., & Pai, R. (2021). Deepprivwidth: Deep learning based semantic segmentation approach for river identification and width measurement in SAR images of coastal Karnataka. *Computers & Geosciences*, 154, 104805. <https://doi.org/10.1016/j.cageo.2021.104805>
- Verpoorter, C., Kutser, T., & Tranvik, L. (2012). Automated mapping of water bodies using landsat multispectral data. *Limnology and Oceanography: Methods*, 10(12), 1037–1050. <https://doi.org/10.4319/lom.2012.10.1037>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Xia, M., Qian, J., Zhang, X., Liu, J., & Xu, Y. (2020). River segmentation based on separable attention residual network. *Journal of Applied Remote Sensing*, 14(3), 032602. <https://doi.org/10.1117/1.jrs.14.032602>
- Yu, L., Zhang, R., Tian, S., Yang, L., & Lv, Y. (2018). Deep multi-feature learning for water body extraction from landsat imagery. *Automatic Control and Computer Sciences*, 52(6), 517–527. <https://doi.org/10.3103/S0146411618060123>
- Yu, Y., Zhang, W., & Deng, Y. (2021). Frechet inception distance (fid) for evaluating gans.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223–2232).