Spatiotemporal Data Augmentation of MODIS-LANDSAT Water Bodies Using Generative Adversarial Networks

Ashit Neema
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Data Science Commons

Recommended Citation
https://digitalcommons.usu.edu/etd/8580
SPATIOTEMPORAL DATA AUGMENTATION OF MODIS-LANDSAT WATER BODIES USING GENERATIVE ADVERSARIAL NETWORKS

by

Ashit Neema

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Data Science

Approved:

Soukaina Filali Boubrahimi, Ph.D.  
Major Professor

Xiaojun Qi, Ph.D.  
Committee Member

Hamid Karimi, Ph.D.  
Committee Member

D. Richard Cutler, Ph.D.  
Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2022
ABSTRACT

Spatiotemporal Data Augmentation of MODIS-LANDSAT Water Bodies using Generative Adversarial Networks

by

Ashit Neema, Master of Science
Utah State University, 2022

Major Professor: Soukaina Filali Boubrahimi, Ph.D.
Department: Computer Science

With the rising need for water resource management tasks, such as lake coastal zone management, rising seas border shift detection, and erosion monitoring, there is a rising necessity in human effort to provide precise water body metadata. 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 efficient monitoring of various environmental resources, several Earth science applications require data at both high spatial and temporal resolutions. To address this need, many data fusion approaches have been described in the literature, that rely on combining data snapshots from multiple sources. Most of the fusion methods learn a mapping of data at different resolutions utilizing the local correlation between the pixel values. The main limitations of data fusion methods are their reliability on data correlations that are sensitive to atmospheric disturbances and other climatic factors that result in noise, outliers, and missing data. To address this challenge, we propose Hydrological Generative Adversarial Network (Hydro-GAN), a novel machine-learning-based method that maps the available data at low resolution to a high-resolution data counterpart. Our proposed model integrates generative adversarial networks that we modified to better the shape of
the water body boundaries. We limited our research scope to mapping water bodies images acquired from Moderate Resolution Imaging Spectroradiometer (MODIS) at low resolution, and Land Remote-Sensing Satellite (LANDSAT) at high resolution.
PUBLIC ABSTRACT

Spatiotemporal Data Augmentation of MODIS-LANDSAT Water Bodies using Generative Adversarial Networks

Ashit Neema

The monitoring of the shape and area of a water body is an essential component for many Earth science and Hydrological applications. For this purpose, these applications require remote sensing data which provides accurate analysis of the water bodies. In this thesis the same is being attempted, first, a model is created that can map the information from one kind of satellite that captures the data from a distance of 500m to another data that is captured by a different satellite at a distance of 30m. To achieve this, we first collected the data from both of the satellites and translated the data from one satellite to another using our proposed Hydro-GAN model. This translation gives us the accurate shape, boundary, and area of the water body. We evaluated the method by using several different similarity metrics for the area and the shape of the water body. The second part of this thesis involves augmenting the data that we obtained from the Hydro-GAN model with the original data and using this enriched data to predict the area of a water body in the future. We used the case study of Great Salt lake for this purpose.

The results indicated that our proposed model was creating accurate area and shape of the water bodies. When we used our proposed model to generate data at a resolution of 30m it gave us better areal and shape accuracy. If we get more data at this resolution, we can use that data to better predict coastal lines, boundaries, as well as erosion monitoring.
To my family, for everything they have done to get me to this day.
ACKNOWLEDGMENTS

I would like to thank my supervisor Professor Soukaina Filali Boubrahimi for the exceptional support and guidance throughout the research. I would also like to thank Ayman Naseer for the encouragement and the engaging collaborations: You have taught me a lot! Special thanks to Utah state University, EarthData and Nasa.gov for providing the infrastructure and resources. Last but not least, I want to thank my family and friends for their overwhelming support.

Ashit Neema
CONTENTS

ABSTRACT .......................................................... iii
PUBLIC ABSTRACT ....................................................... v
ACKNOWLEDGMENTS ...................................................... vii
LIST OF TABLES ........................................................ x
LIST OF FIGURES ........................................................ xi
ACRONYMS ............................................................. xiii
1 INTRODUCTION .......................................................... 1
  1.1 Organization of this work: ...................................... 4
2 BACKGROUND AND RELATED WORK ................................. 5
  2.1 Background on LSR-HSR Image Mapping: ...................... 5
  2.2 Background on Planform Change Use Case: .................... 5
  2.3 Background on Data Pre-processing ............................ 6
  2.4 Background on Generative Networks ............................ 6
3 METHODOLOGY AND DATA PROCESSING ............................. 8
  3.1 Water Bodies Data Sources ..................................... 8
  3.2 Data Pre-processing .............................................. 9
  3.3 Hydrological Generative Adversarial Network (Hydro-GAN) ........ 12
  3.4 Hydrological Generator (Hydro-GEN) Model .................... 14
  3.5 Hydrological Discriminator (Hydro-DIS) Model ................. 15
  3.6 Model Training Challenges .................................... 16
  3.7 Loss Functions used ............................................ 16
  3.8 Experimental Methodology ..................................... 18
    3.8.1 Areal accuracy measures ................................ 18
    3.8.2 Jaccard Similarity Index ................................ 19
    3.8.3 Cosine Similarity Index ................................ 19
    3.8.4 Shape and Distance measure ............................ 20
    3.8.5 Dynamic Time Warping metric .......................... 21
  3.9 Water Body Areal Forecasting .................................. 22
    3.9.1 Data Augmentation Pipeline ............................. 23
    3.9.2 Areal Forecasting Evaluation Metrics .................... 23
  3.10 Dataset Preparation .......................................... 24
4 EXPERIMENTAL RESULTS

4.1 Hydro-GAN Hyper-parameter Search

4.2 Hydro-GAN training result

4.2.1 Areal accuracy result

4.2.2 Shape and distance accuracy result

5 CASE STUDY: AREA EVALUATION AND FORECASTING

5.1 Lake Tharthar: Area Evaluation

5.2 Great Salt Lake: Area Forecast

6 CONCLUDING REMARKS

6.1 Future Work:

REFERENCES
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary of areal and shape evaluation metrics</td>
<td>23</td>
</tr>
<tr>
<td>4.1</td>
<td>Evaluation of LSR to HSR mapping</td>
<td>31</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary of the Great Salt Lake Areal forecasting.</td>
<td>35</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Spatial and temporal resolutions of earth’s different satellites.</td>
<td>2</td>
</tr>
<tr>
<td>3.1 Example of images from MODIS and LANDSAT satellites of multiple reservoirs</td>
<td>9</td>
</tr>
<tr>
<td>3.2 Data pre-processing pipeline</td>
<td>10</td>
</tr>
<tr>
<td>3.3 Structuring element example</td>
<td>11</td>
</tr>
<tr>
<td>3.4 Image pre-processing performed on 2 MODIS examples (Kariba Reservoir and Lake Argyle)</td>
<td>12</td>
</tr>
<tr>
<td>3.5 Architecture of Hydro-GEN and Hydro-DIS sub-models of the Hydro-GAN applied on Lake Mead (part of the Colorado River)</td>
<td>15</td>
</tr>
<tr>
<td>3.6 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</td>
<td>19</td>
</tr>
<tr>
<td>3.7 Overview of the proposed data mapping and forecasting models applied on Great Salt Lake</td>
<td>23</td>
</tr>
<tr>
<td>4.1 Fréchet Inception distance metric of Hydro-GAN generated polygons with $\beta$ values ranging from 0 to 100. The lower the FID value, the better (i.e., the generated and target polygons are similar).</td>
<td>25</td>
</tr>
<tr>
<td>4.2 (a) Example input MODIS image from Qapshaghay Bogeni Reservoir (Kazakhstan) fed to Hydro-GAN and transformed into (b) an HSR after 30 training epochs (2) and (c) 100 training epochs compared to the (d) ground truth LANDSAT image.</td>
<td>26</td>
</tr>
<tr>
<td>4.3 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</td>
<td>28</td>
</tr>
<tr>
<td>4.4 Dynamic time warping measure applied by (a) aligning two polygons shape signatures and (b) computing their warping distance by finding the optimal path along the warping matrix diagonal</td>
<td>29</td>
</tr>
<tr>
<td>4.5 The normalized alignment cost for DTW matrix in test set images</td>
<td>30</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>4.6</td>
<td>Correlation matrix of Jaccard, Cosine and DTW similarity measures used for area, shape and distance accuracy evaluation, when measured between generated and original HSR test set images.</td>
</tr>
<tr>
<td>5.1</td>
<td>Comparison of area variation of Lake Tharthar between generated and actual polygons across 7 years from 2015-2021.</td>
</tr>
<tr>
<td>5.2</td>
<td>Comparing Mean Absolute error for forecasting models with interpolated and non-interpolated data</td>
</tr>
<tr>
<td>5.3</td>
<td>Comparing Mean Squared error for forecasting models with interpolated and non-interpolated data</td>
</tr>
</tbody>
</table>
ACRONYMS

GAN Generative Adversarial Network
Hydro-GAN Hydrological Generative Adversarial Network
MODIS Moderate Resolution Imaging Spectroradiomete
LANDSAT Land Remote-Sensing Satellit
OLI Operational Land Imager
TIRS Thermal Infrared Sensors
GIS Geographic Information Systems
LSR Low Spatial Resolution
HSR High Spatial Resolution
FAIR Findability, Accessibility, Interoperability, and Reusability
ORBIT Ordering Based Information Transfer Across Space and time
SDP Spatially Distributed Probabilistic method
RGD Reg Green Blue
HSV Hue Saturation and Value
Hydro-GEN Hydrological Generator
Hydro-DIS Hydrological Discriminator
DTW Dynamic Time Warping
ERF Effective Receptive Field
FID Fréchet Inception Distance
MLP Multilayer Perceptron
GSL Great Salt Lake
CHAPTER 1
INTRODUCTION

Water bodies monitoring is essential to guide evidence-based decision making which is necessary for hydrological and ecological sustainability [1]. 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 [2]. Previous studies have suggested that lakes and reservoirs are good sentinels of global climate change because they are sensitive to environmental changes [3].

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 [4]. Following the 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 [5]. 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 [6]. In the fifty years since the first satellite was launched, remote sensing satellites have advanced from small-scale production of low-resolution images to daily acquisitions of over 10 terabytes of information [7]. This progress has been largely shepherded by the need for Earth observation sciences. More than 150 Earth observation satellites are currently in orbit, carrying sensors that acquire data at different spatial and temporal resolutions [8]. Due to the differences in satellite missions that directly influence their sensor designs, there is often a trade-off in
different spatiotemporal data resolutions across the remote sensing spectrum [9]. Figure 1.1 illustrates the spatiotemporal 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. For example, MODIS sensors onboard TERRA and AQUA satellites capture Earth’s surface every 8 days at a coarse spatial resolution (500m). On the other hand, OLI and TIRS sensors onboard the LANDSAT satellite captures the earth’s surface every 16 days but at a high spatial resolution of 30m (refer Figure. 1.1 ). Given the limitation of individual sensors at either delivering high spatial resolution or high temporal cadence, there is a clear need to develop interpolation methods that can transfer information across spatiotemporal scales [9]. 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.

Fig. 1.1: Spatial and temporal resolutions of earth’s different satellites.
One among many studies that can benefit from high-quality interpolated data is the analysis of channel planform change and sediment-supply changes of a water body [10]. The precise extraction of the shape and area of water bodies at high spatiotemporal scales from remote sensing satellite imagery is of great significance for planform research [11]. Similarly interpolated data will directly enrich satellite data, making it possible to train data-hungry forecasting models such as deep learning methods [12]. In addition, 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 [13]. Furthermore, interpolated data will facilitate the integration of remote sensing data with Geographic Information Systems (GIS) for automatic or semiautomatic water body extraction and mapping [2].

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 high spatial resolution (HSR) image instances of water bodies and reservoirs collected across a period of seven years from MODIS and LANDSAT satellites. Such a mapping can then be used to generate HSR data for time steps 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 used our mapping to generate interpolated data and and used it on the case study of Great Salt lake.

- We made our source code and data open-source in a project website\(^1\) that meets the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR) wilkinson2016fair.

---

\(^1\)https://sites.google.com/view/hydro-ml/
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 [14]. Timely monitoring of surface water and delivering data on the dynamics of surface water are also essential for policy and decision-making processes [2]. 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 [15].

1.1 Organization of this work:

This work is organized as follows:

- In Chapter 2, We review the background information and related works.

- Chapter 3, elaborates on the methods that were used to build our porposed Hydro-GAN model and how the dataset was collected and preprocessed. It also explains the techniques that were used to obtain and evaluate the results.

- Chapter 4, includes the results that were achieved as part of this thesis.

- Chapter 5, includes the case study of area evaluation and area forecast after using the enriched data obtained from the proposed model.

- Chapter 6, includes the conclusions made by analyzing the results of the studies performed as a part of this work.
CHAPTER 2
BACKGROUND AND RELATED WORK

2.1 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 [9]. Due to factors like noise, outliers, and vast amounts of missing data (due to clouds and sensor failures), the accuracy of the classification methods has been relatively low. To overcome the latter limitations, a new approach has emerged called the Ordering Based Information Transfer Across Space and time (ORBIT) [9]. The main idea behind ORBIT is to use the inherent ordering among instances due to the elevation structure and temporal context. Specifically, if a region is filled with water, it is understood that due to the gravity all the regions in the basin that have lower elevation than the given location should also be filled with water [16]. Another key assumption that was made by this approach is that a water body grows and shrinks smoothly (except for sudden events such as floods). In other terms, the water surface extents of nearby dates are likely to be very similar [17]. Hence, ORBIT can map data from LSR to HSR data only at time steps when noiseless LSR data is available [9].

2.2 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 [18]. 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 [18]. 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 [18]. 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 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.

2.3 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 a-priori known polygons. Image processing techniques have also been used recently for automatically extracting the water bodies’ outline from satellite data without any a-priori map knowledge [19]. 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 [19]. 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 [19].

2.4 Background on Generative Networks

GAN models have been used in a variety of applications, including image synthesis, semantic image editing, style transfer, and classification [20]. 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 [21]. Regular GANs hypothesize the discriminator as a classifier with the sigmoid cross-entropy loss function [22]. 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 involv-
ing a pair of competing networks. Our GAN architecture utilizes a generator model for
outputting new plausible synthetic images, and a discriminator model that classifies images
as real (from the dataset) or fake (generated). As such, the two competing 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 images [20].
CHAPTER 3
METHODOLOGY AND DATA PROCESSING

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 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 [23]. In this section, we will discuss the data preprocessing pipeline and the proposed model.

3.1 Water Bodies Data Sources

Our dataset is collected from MODIS and LANDSAT Earth Observation satellites. Figure 3.1 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 [24]. Terra MODIS and Aqua MODIS are viewing the entire Earth’s surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths [25]. We used the bands 1-2-1 and 7-2-1 to obtain the images as it separates the water bodies and land surfaces (refer Figure 3.1-(a)). The data collected from the MODIS sensors are merged at a temporal resolution of eight days and a spatial resolution of 500 meters. LANDSAT 8 consists of two science instruments called the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The two sensors collect data at a temporal resolution of 16 days and a spatial resolution of 30 meters [26]. We used bands 7-4-3 and 1-5-7 to obtain the images of water bodies (refer Figure 3.1-(b)). 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 \cite{27}. Our dataset 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 satellites. 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 datasets publicly accessible on the project website.

3.2 Data Pre-processing

Prior to the HSR-LSR mapping process, we curated our image datasets 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 3.2 and described in Algorithm 1. The steps are (1) convert the input images to HSV format (lines 3-4), (2) binarize the image into a black-and-white format (lines 5-8), (3) denoise the image using morphological operations (lines 9-11), and (4) apply an image mask to extract the polygon shape.

The first step involved converting the colored image in Red Green Blue (RGB) into a Hue Saturation and Value (HSV) color space image. This step is crucial as the HSV
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 $extracted$.

1. $extracted \leftarrow []$
2. for $currentImage \in [D]$ do
   3. $hsv = convertToHSV(currentImage)$
   4. $v = extract(hsv)$
   5. $binaryIm \leftarrow binarize(v, getHistogramPartition(v))$
   6. $binaryIm \leftarrow (1 - binaryIm) * 255$
   7. $newImage \leftarrow erosion(binaryIm, \theta_{erosion})$
   8. $newImage \leftarrow dialtion(newImage, \theta_{dilation})$
   9. $polygonMask = ones(size(currentImage))$
10. $polygonMask(:, ;, 1) \leftarrow 255 - newImage$
11. $polygonMask(:, ;, 2) \leftarrow 255 - newImage$
12. $extracted \leftarrow append(polygonMask)$
3. end for
4. 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 color space is the most efficient image format for color-based image segmentation [28]. 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 [29]. 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.
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).

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. 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 $\theta_{\text{erosion}}$ 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 $\theta_{\text{dilation}}$ parameter in Algorithm 1.

![Fig. 3.3: Structuring element example](image-url)
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 3.3 shows an example of a morphological opening that removes the noisy patches from the original image. 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 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 3.4 illustrates the entire data pipeline steps using two MODIS LSR images (previously shown in Figure 3.1-(a)).

![Image Pre-processing Steps](image)

Fig. 3.4: Image pre-processing performed on 2 MODIS examples (Kariba Reservoir and Lake Argyle)

### 3.3 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 datasets at times when only LSR data are available.

We used generative adversarial network (GAN) frameworks as the backbone of our proposed model. Our motive for utilizing GAN is their ability to generate crispy sharp
images [30]. 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 [31].

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 dataset. Hydro-GAN is composed of two sub-models Hydro-GEN and Hydro-DIS. The first sub-model is the generator model that we train to generate new synthetic images. The second sub-model is the discriminator model that is trained for classifying whether the generated water bodies are real (from the HSR domain) or synthetic (i.e., generated by the Hydro-GEN 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: $G : x \rightarrow y$ [20]. Hydro-GEN utilizes a conditional generative adversarial network (cGAN), where the output $y$ is conditioned on some input $z$, resulting in a mapping: $G : (x, z) \rightarrow y$ [20]. The generator is trained via adversarial loss, and an additional soft Dynamic Time Warping (DTW) loss term, measured between the generated image and the expected ground truth water body images. This additional loss encourages the generator model to create accurate boundaries shape that aligns well with typical water bodies' shape signatures. Equation 3.1 defines the objective function that a traditional GAN network is optimizing for.

$$
\min_{D, G} \max \left( \mathbb{E}[\log(D(z, y))] + \mathbb{E}[\log(1 - D(z, G(x, z)))] \right)
$$

(3.1)
In equation 3.1 \( D(z, y) \) refers to the probability that a sample image \( y \) pertains to the real dataset 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 fake instance is real. 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 \( \log(1 - D(G(z)) \).

The discriminator \( D \) is trained to maximize the function, while the generator \( G \) is trained to minimize it. 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 \( \log(1 - D(G(z)) \).

### 3.4 Hydrological Generator (Hydro-GEN) Model

Our proposed Hydro-GEN sub-model is an encoder-decoder model that uses a U-Net architecture [32]. The model takes a source water body pre-processed LSR image and generates a target HSR image. To achieve this purpose, it first down-samples the input image down to a bottleneck latent layer, then up-sampling the bottleneck representation back to the size of the output image [32]. 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 encoding layers and the corresponding decoding layers, forming a U-shape. The skip connections are used to pass the low-level data of the HSR image through the bottleneck layer. The encoder and decoder of the generator are comprised of standardized blocks of convolutional, batch normalization, dropout, and activation layers. Figure 3.5 shows the skip connections used by the U-Net network [20].

The Hydro-GEN generator model is trained via the Hydro-DIS discriminator model. The weights of Hydro-GEN are updated to reflect the discriminator loss when predicting synthetic images as either real or fake 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 weights are also updated to minimize the mean absolute error between the generated images and the target ground truth image.
known as the L1 loss. The final objective function is made of a weighted sum of both the
adversarial and the L1 losses. The weights range between 1 to 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.

Fig. 3.5: Architecture of Hydro-GEN and Hydro-DIS sub-models of the Hydro-GAN
applied on Lake Mead (part of the Colorado River)

3.5 Hydrological Discriminator (Hydro-DIS) Model

The Hydro-DIS discriminator is a deep convolutional neural network model that per-
forms conditional image classification [33]. It takes both the LSR source image and the
target HSR image pairs as input and predicts the likelihood of the target HSR image being
real. The discriminator design is based on the Effective Receptive Field (ERF) of the model
which are the regions that contain input pixels with a non-negligible impact [34]. ERFs
provide a one-to-many mapping between the pixel in the LSR input image and the pixels in
the HSR target image. Each value of the model activation map defines the likelihood that a
patch in the input image is real. These values are averaged to give an overall classification
score for the images. The architecture of our Hydro-GAN model is shown in (Figure 3.5).
3.6 Model 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 [35]. 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 [36]. The second challenge is the gradient vanishing problem that might occur if the Hydro-DIS discriminator performs significantly better than the generator [35]. In other terms, if the Hydro-GEN is highly accurate, the hydro-DIS discriminator will result in a random guess (50% accuracy), which poses a threat to the convergence of the Hydro-GAN as a whole.

3.7 Loss Functions used

To overcome the two aforementioned challenges, we propose to explore a new loss function. First, we modify the traditional GAN model defined in Equation 3.1 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 3.1, the generator maximizes the log of the discriminator probabilities i.e., –log(D(G(z))). This change is inspired by framing the problem from a different perspective, where the generator seeks to maximize the probability of images being real, instead of minimizing the probability of an image being fake. This avoids generator saturation through a more stable weight update mechanism [35].

The second loss adjustment that we propose is to use a least-squares loss, where the discriminator seeks to minimize the sum of the squared difference between predicted and expected values for real and synthetic images as defined in Equation 3.2. 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 3.3. The benefit of the least-squares loss is that it gives more penalty to larger errors, in turn resulting in a large correction rather than a vanishing gradient and no model update [22].
\[
\text{minimize}(D(z, y) - 1)^2 + (D(z, G(x, z)))^2
\]  
(3.2)

\[
\text{minimize}(D(z, G(x, z)) - 1)^2
\]  
(3.3)

Along with using non-saturating and least-squares losses, we also propose a new Dynamic Time Warping (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 [37]. Our motivation for adding a DTW loss is that it emphasizes the polygons shape accuracy which equips the generator to better fool the discriminator. Our new generator loss, defined in Equation 3.4, is now a weighted sum of the cross-entropy adversarial loss \( \mathcal{L}_{\text{adversarial}} \), the L1 loss \( \mathcal{L}_{L1} \), and the DTW loss \( \mathcal{L}_{DTW} \).

\[
G(x, z) = \mathcal{L}_{\text{adversarial}} + \mathcal{L}_{L1} + \mathcal{L}_{DTW}
\]  
(3.4)

where \( \mathcal{L}_{\text{adversarial}} \) is the adversarial loss is the probabilistic loss which is also used in the discriminator. \( \mathcal{L}_{L1} \) loss is the mean absolute error between the generated and the expected 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 \( \beta \) parameter than can be used to balance the terms.

\[
G(x, z) = (100 - \beta) \times (\mathcal{L}_{\text{adversarial}} + \mathcal{L}_{L1}) + \beta \times \mathcal{L}_{DTW}
\]  
(3.5)

We evaluated the synthetic polygon accuracy using the Fréchet Inception Distance
(FID), which is a commonly used metric for evaluating generative models [38]. The FID metric is defined as the squared Wasserstein metric between two multidimensional Gaussian distributions of the real and synthetic data. FID summarizes how similar the two groups are in terms of statistics on computer vision features of the raw images calculated using the Inceptionv3 model used for image classification [39]. The FID is defined in Equation 3.6.

\[
\text{FID} = \|\mu - \mu_w\|^2_2 + \text{tr} \left( \Sigma + \Sigma_w - 2 \left( \Sigma^{1/2} \Sigma_w \Sigma^{1/2} \right)^{1/2} \right)
\]  

(3.6)

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. A low FID score indicates that the two groups of images are similar, while a higher score suggests that the images have dissimilar characteristics.

3.8 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. For areal accuracy, Jaccard and Cosine measures were selected. The DTW similarity metric was used for shape and distance accuracy. In the following subsections, we will detail each measure and explain its usefulness and limitations.

3.8.1 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 3.6-(b). The area similarity is computed using the Jaccard and Cosine similarity indices based on the extracted polygon boundaries.

![Generated Image](a) Generated Image ![Polygon Boundary](b) Polygon Boundary ![Distance from centroid](c) Distance from centroid

**Fig. 3.6:** 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

### 3.8.2 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):

\[
J(P_{\text{real}}, P'_{\text{predicted}}) = \frac{\text{Area}(P_{\text{real}} \cap P'_{\text{predicted}})}{\text{Area}(P_{\text{real}} \cup P'_{\text{predicted}})} * 100 \tag{3.7}
\]

Since the Jaccard similarity index considers the size of the shared area relative to the size of the combined union area of two polygons it is a scale-invariant measure [40].

### 3.8.3 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 [40]:

\[
C(P_{\text{real}}, P'_{\text{predicted}}) = \frac{\text{Area}(P_{\text{real}} \cap P'_{\text{predicted}})}{\sqrt{\text{Area}(P_{\text{real}} \cdot P'_{\text{predicted}})}} * 100 \tag{3.8}
\]
Both indices range between zero and one hundred and the higher the metric values the better the areal match of the real \(P_{\text{real}}\) and generated \(P'_{\text{predicted}}\) polygons.

### 3.8.4 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.

**Algorithm 2 Shape Signature Derivation**

**Input:** The polygon image \(\text{image}\), and an image threshold value \(\text{thres}\)

**Output:** Distance array \(\text{dist}\).

1. \(\text{dist} = []\)
   \(\triangleright\) Convert images to binary format
2. \(\text{binaryImage} \leftarrow \text{getBinary(image, thres)}\)
   \(\triangleright\) Find Polygon contours
3. \(\text{contourImage} \leftarrow \text{getContour(binaryImage)}\)
   \(\triangleright\) Calculate centroid points coordinates
4. \(M1 = \text{getCentroidCoordinates(contourImage)}\)
5. \(cx = \text{getXCoordinate(M1)}\)
6. \(cy = \text{getYCoordinate(M1)}\)
   \(\triangleright\) Calculating distance from centroid
7. \(\text{points} \leftarrow \text{getEdgeCoordinates(contourImage)}\)
8. **for** \(\text{point} \in \text{points} \text{ do} \)
9.   \(\text{distance} \leftarrow \text{getDistance([cx, cy], [point[0], point[1]])}\)
10. \(\text{dist} \leftarrow \text{append(distance)}\)
11. **end for**
12. **return** \(\text{dist}\)

In this work, we consider the centroid of the polygon as the reference point. We extracted the centroids of the water body polygons and used the euclidean distance between the centroid of the polygon and the contour point coordinates (edge coordinates) in a clockwise direction to create a sequence of distances. The latter was used to create the polygon shape signatures (i.e., time series). Algorithm 2 describes the shape signature derivation algorithm. Figure 3.6-(c) 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.

### 3.8.5 Dynamic Time Warping metric

DTW method is defined in Algorithm 3 [41]. DTW arranges the two input centroid shape signatures on the sides of an $n \times m$ grid ($n$ and $m$ being the length of the two centroid shape signatures). Both of the sequences start on the bottom left of the grid. Within each cell of the grid, a distance value is given based on the corresponding elements of the two sequences. The best alignment between the two centroid shape signatures is acquired by looking for the path from the bottom left corner to the top right corner that minimizes the total incremental distance. The total distance is called the warping distance, and it represents the minimum of the sum of the distances between the individual elements on the path. If the shapes match perfectly, which is the best-case scenario, there is a one-to-one pairwise matching between the two sequences, which results in a zero warping distance. The canonical form of DTW is shown in Equation 3.9. $M$ and $N$ represent 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 cost function $g$ in Equation 3.9 is usually chosen to be the square of the differences between $x_i$ and $y_j$.

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

(3.9)

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

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. Along with the warping path, DTW generates a warping distance between two centroid shape signatures that represents the dissimilarity between these two shapes. The larger the warping distance value, the more discrepancy exists between the two polygon shapes.
Algorithm 3 Algorithm for finding minimum-cost path through a DTW matrix.

**Input:** DTW distance matrix \( \text{dist} \text{mat} \)

**Output:** minimum cost path \( \text{path} \), cost matrix \( \text{cost} \text{mat} \).

1: \( N, M = \text{getShape}(\text{dist} \text{mat}) \) \( \triangleright \) Initialize cost matrix
2: \( \text{cost} \text{mat} = [N, M] \) \( \triangleright \) Fill the cost matrix
3: \( \text{traceback} \text{mat} = \text{np.zeros}(N, M) \)
4: **for** \( i \in \text{range}(N) \) **do**
5:     **for** \( j \in \text{range}(M) \) **do**
6:         \( \text{penalty} = \text{getPenalty}(\text{cost} \text{mat}) \)
7:         \( i\text{penalty} = \text{getIndexPenalty}(\text{penalty}) \)
8:         \( \text{cost} \text{mat} = \text{updateMat}(\text{dist} \text{mat}, \text{penalty}) \)
9:         \( \text{traceback} \text{mat} = \text{updateTraceback}(i\text{penalty}) \)
10:     **end for**
11: **end for** \( \triangleright \) Traceback from bottom right
12: \( i = N - 1, j = M - 1, \text{path} = [(i, j)] \)
13: **while** \( i > 0 \ OR \ j > 0 \) **do**
14:     \( \text{traceback} \text{type} = \text{traceback} \text{mat}[i, j] \)
15:     IF \( (\text{traceback} \text{type} == 0), \text{Then} \rightarrow i = i - 1, j = j - 1 \)
16:     IF \( (\text{traceback} \text{type} == 1), \text{Then} \rightarrow i = i - 1 \)
17:     IF \( (\text{traceback} \text{type} == 2), \text{Then} \rightarrow j = j - 1 \)
18:     \( \text{path} \leftarrow \text{append}((i, j)) \)
19: **end while**
20: \( \text{cost} \text{mat} = \text{updateCostMatrix}(\text{cost} \text{mat}) \)
21: **return** \( \text{path}, \text{cost} \text{mat} \)

Table 3.1 shows the summary of the areal and shape evaluation metrics that were used in this paper, along with their advantages and drawbacks.

### 3.9 Water Body Areal Forecasting

After training our GAN model on the LSR and HSR mapping task, we interpolated the ground truth HSR data by generating new synthetic HSR water body images. We extracted the shape signatures and the area of the water bodies across 7 years using the new interpolated dataset. The new dataset is then used to train a model for areal forecasting of water bodies. We developed a Multilayer Perceptron (MLP) model to perform the areal regression of the Great Salt Lake water body. We used the areal time series of the 5 historical years \( TS = \{ Ar_{n-4}, Ar_{n-3}, \ldots, Ar_n \} \) to predict the area of the the next year \( Ar_{n+1} \).
### Evaluation measure

<table>
<thead>
<tr>
<th>Evaluation measure</th>
<th>Advantages</th>
<th>Drawbacks</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaccard Similarity</td>
<td>Useful when assessing polygons with similar characteristics</td>
<td>Operates poorly with polygons that are very large or small</td>
<td>[0,100]</td>
</tr>
<tr>
<td>Cosine Similarity</td>
<td>Better for assessing areal similarity between imbalanced polygons</td>
<td>Does not penalize shapes that have highly different areas</td>
<td>[0,100]</td>
</tr>
<tr>
<td>DTW</td>
<td>Good for shape similarity when starting point is well selected</td>
<td>Needs shape alignment and is vulnerable to changes in starting points</td>
<td>[0,∞]</td>
</tr>
</tbody>
</table>

Table 3.1: Summary of areal and shape evaluation metrics

---

**Fig. 3.7:** Overview of the proposed data mapping and forecasting models applied on Great Salt Lake

### 3.9.1 Data Augmentation Pipeline

Figure 3.7 shows the processing pipeline of our proposed data augmentation solution. Our method comprises two main steps: (1) data augmentation through adversarial image generation and (2) areal water body forecasting.

### 3.9.2 Areal Forecasting Evaluation Metrics

To evaluate the results of our forecasting model, we applied the three metrics: Mean
Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

**Mean Squared Error (MSE).** calculates as the average of the squared differences between predicted and expected target values in a dataset [42] (area in our case), as defined in Equation eq1.

\[
\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2
\]  

(3.10)

where \(Y_i\) is the \(i_{th}\) expected value in the dataset and \(\hat{Y}_i\) is the \(i_{th}\) predicted value.

**Root Mean Squared Error (RMSE).** is the square root of average value of squared error in a set of predicted values, without considering direction. The lower the value the better the model [43], as defined in Equation 3.11.

\[
\sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}
\]

(3.11)

**Mean Absolute Error (MAE).** is the average value of error in a set of predicted values, without considering direction, , as defined in Equation 3.12

\[
\frac{1}{n} \sum_{i=1}^{n} |(Y_i - \hat{Y}_i)|
\]

(3.12)

3.10 **Dataset 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 dataset 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 6720 MODIS satellite images and 336 out of 3360 LANDSAT-8 images were included in the test set.
CHAPTER 4
EXPERIMENTAL RESULTS

4.1 Hydro-GAN Hyper-parameter Search

To identify the optimal weights that produce the highest polygon accuracy based on Fréchet Inception Distance (FID), we performed a grid search of the weight parameters, defined in Equation 3.5. We trained our proposed Hydro-GAN model on different weight ratios. Figure 4.1 shows a heatmap that summarizes the FID scores obtained by each model when trained on different $\beta$ weights, ranging from 1 to 100 for balancing the $L_{\text{adversarial}}$, $L_{L1}$ and $L_{DTW}$ losses.

![Frechet Inception Distance (FID) between generated and target Images](image)

**Fig. 4.1:** Fréchet Inception distance metric of Hydro-GAN generated polygons with $\beta$ values ranging from 0 to 100. The lower the FID value, the better (i.e., the generated and target polygons are similar).
The result in Figure 4.1 reveals that when more weight is assigned to the DTW loss, the Hydro-GAN model is generating 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, 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 $\beta$ weight is small, the FID scores are high which signify that the Hydro-GAN loss is overpowered by the adversarial and L1 losses (refer to the bottom left corner of Figure 4.1). 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 $\beta$ value of 100.

Fig. 4.2: (a) Example input MODIS image from Qapshagay Bogeni Reservoir (Kazakhstan) fed to Hydro-GAN and transformed into (b) an HSR after 30 training epochs (2) and (c) 100 training epochs compared to the (d) ground truth LANDSAT image.

4.2 Hydro-GAN training result

After identifying the optimal weight ratios from Figure 4.1, 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 4.2 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 4.2-(a) shows an input LSR MODIS image fed to the Hydro-GAN model and
a temporary results that were learned after 30 epochs (Figure 4.2-(b)) and after 100 epochs (Figure 4.2-c). 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 4.2-(d). 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 5.

4.2.1 Areal accuracy result

To measure the areal accuracy between the generated polygons (from the Hydro-GAN model using the test dataset) and the ground truth HSR LANDSAT polygons, we used the Jaccard and Cosine similarity percentages as illustrated in Figure 4.3. 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 3.1, one reason that can explain this correlation is that Jaccard and Cosine have the same numerator as shown in Equations (3.7) and (3.8).

While Jaccard measures the predicted and ground-truth geometries area of intersection with respect to their union area, Cosine measures the area of intersection with respect to the geometric mean of the two geometry areas. Figure 4.3-(a) shows that the Jaccard similarity measure is relatively pessimistic compared to the Cosine measure. This is due to Jaccard’s usage of the geometric mean of the polygons’ areas which is smaller than the areal union of the polygons’ areas. There are distinct advantages of using Jaccard and Cosine. When the
Fig. 4.3: 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

similarity is measured between a large geometry and a smaller geometry, Jaccard will be more pessimistic than the Cosine measure. This shows that the Cosine measure is assessing the areal similarity between imbalanced shapes more accurately than Jaccard. On the other hand, when assessing the similarity between geometries with similar areal characteristics, using a pessimistic similarity measure is a better choice [40].

4.2.2 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 4.4-(a). As shown in the Figure 4.4-(a), the result indicate that when there is a dip in the shape signature of 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 1000. 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 ground truth 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 4.4-(b) which shows the DTW alignment and warping matrix that were produced when comparing the two polygons. The graph in Figure 4.4-(b) shows that, the warping path (denoted by the blue line starting from the bottom left corner 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.

![Dynamic time warping measure](image)

Fig. 4.4: Dynamic time warping measure applied by (a) aligning two polygons shape signatures and (b) computing their warping distance by finding the optimal path along the warping matrix diagonal

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 4.5, which shows that the distribution of the normalized 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.

![Normalized alignment cost for DTW matrix in test set images](image)

Fig. 4.5: The normalized alignment cost for DTW matrix in test set images

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 4.6. 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 4.6 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 which 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 need
to study the water volume change, then introducing a new areal loss $L_{\text{area}}$ could be useful.

![Fig. 4.6: Correlation matrix of Jaccard, Cosine and DTW similarity measures used for area, shape and distance accuracy evaluation, when measured between generated and original HSR test set images.](image)

Table 4.1 shows the result of the image mapping and interpolation step. The results suggest that our GAN-based model achieves high accuracy in both shape and area of a water body across all the three aforementioned metrics.

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>Average Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaccard similarity (Areal)</td>
<td>91%</td>
</tr>
<tr>
<td>Cosine similarity (Areal)</td>
<td>93%</td>
</tr>
<tr>
<td>Dynamic Time Warping (Shape)</td>
<td>1.31</td>
</tr>
</tbody>
</table>
5.1 Lake Tharthar: Area Evaluation

In this case study, we evaluated our proposed Hydro-GAN model on lake Tharthar’s shrinking and expansion behaviors. Our study area consists of lake Tharthar’s polygon area across a period of seven 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 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 5.1. The blue bar plot in Figure 5.1 indicates that, Lake Tharthar (one of the reservoirs from the dataset) first shrinks in surface area from 2015 to 2020 with a little increase in the year 2018, then shrinks again 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 [44].

The increase in the area of the water body in 2021 can be seen in Figure 5.1-(g) 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 area of a water body across the seven years. 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 5.1 is that the area calculated from the LSR MODIS images, shown in the red bar plot, does not 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 5.1, are more realistic area estimations that produce significantly lower errors than LSR MODIS. This result shows that the LSR to HSR conversion improved the prediction of the area of the lake Tharthar water body. In addition, we note that our model produces generally a slight under-estimation.
of the area. Following our suggestion in Section 4.2.2, using an areal loss can correct the under-estimation.

5.2 Great Salt Lake: Area Forecast

One of the saltiest inland lakes in the world is the Great Salt Lake (GSL), which is situated in northern Utah. The water flows into the lake from Weber, Bear, and Jordan rivers with no outlet. Since the GSL is closed basin, the water level rises in wet season due to water supply and falls during droughts [45]. With increased water use for urban, industrial, and agricultural purposes together with a decline in the lake’s natural water supply through time, the lake’s elevation has declined. As the lake’s elevation drops, its volume decreases, its salt concentration increases, and evaporation decreases as a result [45].

5.2.1 Areal Forecast

We used the enhanced dataset to predict the water body area after creating the HSR water body polygons. We trained our forecasting model using both the ground truth data and the enriched data produced by our GAN model in order to test the effectiveness of our enriched dataset [45]. Figure 5.2 shows the prediction results generated by our forecasting model. The results show that training the model with interpolated data improves its mean absolute error performance levels (i.e., more data is used for learning short-term changes in the water body area). When we compared the MSE of the two forecasting models in Figure 5.3, we found similar results.

The results of the three metrics are summarized in Table 5.1, which demonstrates that training on the enriched dataset improves forecasting across all the evaluation metrics. The interpolated data reduces error when used to train the MLP model, suggesting that mapping from LSR to HSR to produce more data is helpful in predicting a water body’s surface area. This result also supports our findings when evaluating our adversarial approach. Since the data generated in the adversarial step is only used for training the forecasting model, the forecasting improvement suggest that the synthetic HSR data is of high quality (i.e., within the ground truth data distribution).
Fig. 5.2: Comparing Mean Absolute error for forecasting models with interpolated and non-interpolated data

(a) Forecasting results obtained using interpolated data  
(b) Forecasting results obtained without using interpolated data

Fig. 5.3: Comparing Mean Squared error for forecasting models with interpolated and non-interpolated data

(a) Forecasting results obtained using interpolated data  
(b) Forecasting results obtained without using interpolated data

Table 5.1: Summary of the Great Salt Lake Areal forecasting.

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>MLP + Ground Truth Data</th>
<th>MLP + Enriched Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>49.46</td>
<td>21.97</td>
</tr>
<tr>
<td>RMSE</td>
<td>7.03</td>
<td>4.68</td>
</tr>
<tr>
<td>MAE</td>
<td>6.10</td>
<td>3.58</td>
</tr>
</tbody>
</table>
In this paper, we have proposed Hydro-GAN which is a deep learning-based generative method that uses a conditional generative adversarial network for mapping the remote sensing information available at low resolution (MODIS satellites) to high resolution (LANDSAT satellites). 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 datasets 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.

To test the effectiveness of our proposed model we used the case study of the Great Salt Lake. We used our mapping to generate interpolated data and used it to forecast the surface area of Great Salt lake. Our interpolated data increased the accuracy of the forecasting model, which indicates that the model achieves better predictions when more data was used for training for the task of short term areal prediction.

6.1 Future Work:

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 plan to make improvements to our mapping algorithm and our Hydro-GAN model so that it can generate mapping of those water bodies on which the model was never trained on. 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.
REFERENCES


[38] Y. Yu, W. Zhang, and Y. Deng, “Frechet inception distance (fid) for evaluating gans.”


