MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SIENCE AND TECHNOLOGY

HOÀNG THỊ MINH CHÂU

RESEARCH ON THE DEVELOPMENT OF A MULTIMODAL RESERVOIR WATER LEVEL FORECASTING MODEL USING SATELLITE IMAGERY AND RESERVOIR OPERATION DATA

SUMMARY OF DOCTORAL DISSERTATION ON INFORMATICS

Major: Information System

Code: 9 48 01 04

Hanoi - 2026

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy of Science and Technology

Supervisors:

Supervisor 1: Assoc. Prof. Dr. Tran Thi Ngan

Supervisor 2: Assoc. Prof. Dr. Nguyen Long Giang

Referee 1: _____

Referee 2: _____

Referee 3: _____

The dissertation is examined by Examination Board of Graduate University of Science and Technology, Vietnam Academy of Science and Technology at …. AM, February ….., 2026

This dissertation can be found at:

1) Graduate University of Science and Technology Library
2) National Library of Vietnam

INTRODUCTION

1. Necessity of the research

Reservoirs play a critical role in flow regulation, flood mitigation, water supply and power generation, contributing to socio-economic development and ecosystem stability. However, climate change is altering the hydrological cycle, increasing the frequency and magnitude of extreme events and producing nonlinear variability in reservoir water level time series. Many reservoirs worldwide have recorded declining water levels, leading to water imbalance, deterioration of water quality and increased downstream safety risks, while the reservoir's flood-mitigation function becomes ever more essential.

In reservoir operation, the reliability of water level forecasts is a key factor for release and regulation decisions; erroneous forecasts can cause inundation or water shortages. Therefore, developing accurate, flexible forecasting models for reservoir water level that can adapt to climate change is an urgent requirement. Such development will contribute to more effective water resource management, ensure water security and reduce disaster risk.

2. Summary of studies on reservoir water level forecasting

Previous studies on reservoir water level forecasting can be divided into two main groups: traditional approaches and deep-learning-based approaches.

- **Traditional approaches:** include hydraulic/hydrological models [29–31], regression models such as ARIMA [32, 33], SARIMA [34, 35] ... or conventional machine learning models [36–46].
- **Machine learning / deep learning approaches:** employ machine learning or deep learning models such as RNN/LSTM/BiLSTM/... for forecasting reservoir water levels using satellite imagery data.

In recent years, multimodal reservoir water level forecasting methods have been strongly developed [47–53]. The focus of this research direction is the mechanism for fusing heterogeneous data sources (image data and numerical data) to build comprehensive forecasting models that simultaneously reflect spatial features and temporal dynamics of reservoir systems.

Multimodal reservoir water level forecasting has attracted considerable attention, concentrating on exploiting the combination of satellite imagery and reservoir operational data with deep learning models to estimate reservoir water levels. This approach allows simultaneous utilization of the rich spatial information and wide coverage of satellite imagery together with the high-resolution, continuous temporal characteristics of operational hydrological data, thereby improving forecast accuracy and stability.

In recent years, satellite imagery has become a highly valuable data source and is increasingly exploited across many fields. For example, Ul Islam et al. [59] used remote sensing data to analyze spatial distribution changes and infrastructure scale in Peshawar District (Pakistan), while Song et al. [60] applied satellite imagery in the maritime domain, demonstrating the great potential of imagery to describe complex spatio-temporal phenomena.

In hydrology, satellite imagery plays a particularly important role thanks to its wide spatial coverage, temporal continuity and remote-observation capability. Numerous studies [61, 62] have exploited satellite image time series to monitor and forecast reservoir water level variations at large scales and over long-term periods. Notably, satellite imagery can provide water level information in areas that are difficult or impossible to instrument with field monitoring systems, such as frequently flooded zones, remote river basins or large-scale reservoirs [63]. Additionally, with increasingly shorter revisit cycles and higher spatial resolutions, satellite data can detect both small fluctuations and abrupt changes in reservoir water levels in a timely manner, thereby contributing to improved accuracy and reliability of forecasting models [64].

Fundamentally, multimodal deep-learning methods for forecasting generally follow a three-stage information-processing workflow to map heterogeneous data types into a unified predictive space:

- *Representation stage:* Different data types are preprocessed and transformed into suitable formats (vectors, matrices or tensors).
- *Fusion stage:* This is the most critical step, where information from heterogeneous data sources is combined. This process can be performed via operations such as concatenation, addition or more sophisticated mechanisms like attention. The objective is to obtain a joint representation that encapsulates both spatial and temporal information.
- *Inference stage:* From the joint representation, the model uses fully connected layers to produce the forecasted water level output.

The fundamental difference among studies lies in where the "fusion stage" occurs within the network architecture. Based on this criterion, methods are classified into three main strategies: early fusion, mid-level fusion and late fusion. The methods are illustrated in Figure 0.1.

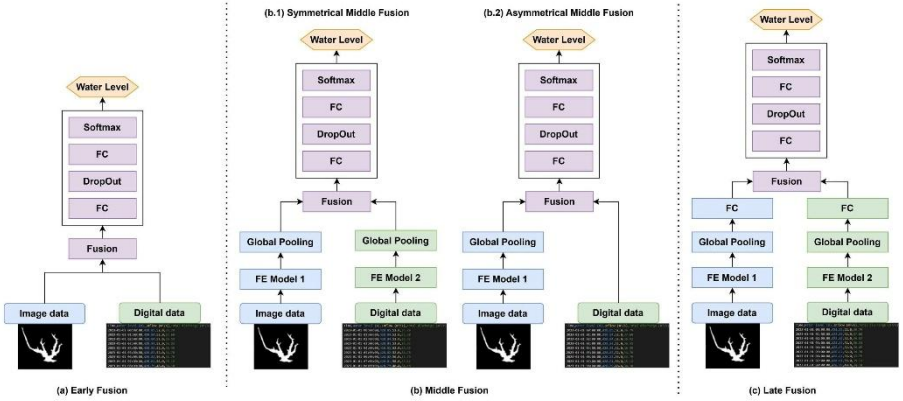

Figure 0.1. Classification of multimodal integration methods

3. Research gaps

a) The exploitation of satellite imagery in reservoir water level forecasting still exhibits significant research gaps, especially in the context of multimodal forecasting models that combine image data and hydrological numerical data.

Most current studies use CNN or ResNet architectures to extract spatial features from satellite imagery, while reservoir operational data and hydrological time-series are often processed independently by simple models such as MLPs or fully-connected layers. The extracted features are primarily normalized and fused using simple fusion strategies, typically vector concatenation at intermediate or output layers of the model.

This approach shows several limitations:

Firstly, the feature fusion mechanism remains largely mechanical and does not adequately reflect intrinsic relationships and dynamic interactions between spatial information from satellite imagery and the temporal dynamics of the hydrological system.

Secondly, long-term temporal dependencies—which are crucial in processes of storage, release and reservoir regulation—are not effectively exploited due to a lack of tight integration with deep architectures specialized for sequential data such as LSTM, GRU or Transformer.

Thirdly, many models are sensitive to noisy data, cloud cover or missing satellite images, resulting in limited robustness to anomalous inputs and reduced forecasting stability.

Finally, the high computational complexity of some existing multimodal models increases training time and hinders deployment in operational forecasting systems.

b) There is a lack of analysis on the multimodal characteristics of reservoir water level data that tightly integrate satellite imagery and reservoir operational data. At the same time, unified multimodal forecasting systems that cohesively integrate satellite imagery and operational reservoir data within a single architecture have not been developed to effectively exploit spatio-temporal dependencies across multiple scales. The objective is to improve model accuracy, stability and generalization capability, while steering toward practical applicability in reservoir management and operation.

4. Research objectives of the thesis

General objective: To research and develop a multimodal reservoir water level forecasting model, using satellite imagery of the reservoir area and reservoir operational data as the model inputs.

Specific objectives:

- Propose a multimodal reservoir water level forecasting model in which spatial features extracted from satellite imagery are fused with numerical features from reservoir operational data into a unified representation space for forecasting.

- Establish experiments to evaluate the validity and forecasting performance of the proposed models with the goal of reducing forecast errors, especially under noisy conditions.
- Concurrently develop a web-based multimodal reservoir water level forecasting system to support integration, storage and analysis of satellite imagery and hydrological data in real time, serving reservoir management and operation.

5. Main contributions of the thesis

To enhance forecasting accuracy of reservoir water levels and thereby improve reservoir operation and management effectiveness, this thesis proposes the following principal scientific contributions:

- First, the thesis proposes the multimodal reservoir water level forecasting model HOG-GRU in [CT1] for daily mean reservoir water level forecasting.
- Second, the thesis proposes the multimodal reservoir water level forecasting model Colubrid-Net in [CT2] for daily mean reservoir water level forecasting.
- Third, the thesis conducts experiments to evaluate the forecasting performance of the proposed models and implements a web-based multimodal reservoir water level forecasting system in works [CT3, CT4, CT5].

6. Thesis structure

- **Introduction**: presents research necessity, literature review, objectives, scope, methodology and contributions of the thesis.
- **Chapter 1:** introduces the reservoir water level forecasting problem and its characteristics. This chapter also provides foundational knowledge necessary for the design and refinement of the forecasting models proposed in subsequent chapters.
- **Chapter 2**: proposes two multimodal reservoir water level forecasting models HOG–GRU and Colubrid-Net for daily mean reservoir water level prediction, including detailed descriptions of architectures, feature extraction and fusion methods from satellite imagery and time-series data, as well as model construction and training procedures.
- **Chapter 3:** presents experiments evaluating the effectiveness of the proposed models, including study area description, Sentinel-2 imagery and reservoir operational data, experimental design, evaluation metrics, results and discussion, and introduces the multimodal reservoir water level forecasting system.
- **Conclusion and future directions:** summarizes achieved results, highlights scientific and applied contributions, points out limitations and proposes directions for future research.
- **List of the author's publications**
- **References**

CHAPER 1. THEORETICAL BASIS

This chapter presents the foundational knowledge related to the problem of reservoir water level forecasting, including the definition of the reservoir water level forecasting problem and its basic characteristics. At the same time, the chapter also introduces knowledge that supports the research directions implemented in Chapter 2 and Chapter 3 of the thesis, such as: Sentinel-2 satellite imagery; regression-based reservoir water level forecasting models; deep-learning-based reservoir water level forecasting models; and image-processing techniques that support the reservoir water level forecasting problem. In addition, model evaluation metrics and the ANOVA analysis method are included.

CHAPER 2. PROPOSAL OF MULTIMODAL RESERVOIR WATER LEVEL FORECASTING MODELS

2.1. Introduction

Based on analysis of the characteristics of the reservoir water level forecasting problem under multi-source data conditions, this chapter focuses on presenting two multimodal reservoir water level forecasting models. The models jointly exploit Sentinel-2 satellite imagery and observed reservoir water-level data to forecast daily mean reservoir water level for reservoirs whose water-surface shape varies with water level. Satellite imagery serves as a supplementary spatial information source, enhancing representation capacity and supporting operational reservoir data during forecasting.

The common novelty of the proposed models lies not only in combining image and numerical data, but also in the methods of image feature extraction and the feature-fusion strategies. Instead of using simple feature concatenation as in many previous studies, the thesis proposes guided fusion mechanisms and improves the image feature-extraction backbone to more effectively exploit spatial information from satellite imagery. The two proposed models are HOG-GRU [CT1] and Colubrid-Net [CT2].

The research results of Chapter 2 are published in works [CT1, CT2] listed in the "Author's Publications" section.

2.2. Proposal of the multimodal forecasting model HOG-GRU

2.2.1. Forecasting model

The proposed model is illustrated in Figure 2.1 and uses both satellite imagery and measured reservoir water-level data as inputs. Satellite images are first converted into reservoir masks, then HOG (Histogram of Oriented Gradients) is applied to extract features in vector form and these vectors are L2-normalized to obtain the final feature representation. This image feature vector is normalized to the same space as the reservoir water-level data, which have been normalized to [0,1], and then combined via the feature-combination function $P(y_t, f_t)$ to create a new exogenous variable. Finally, the exogenous variable and the original water-level time series—both

normalized to the domain [0,1] - are fed into a GRU model to forecast the reservoir water level at the next time step.

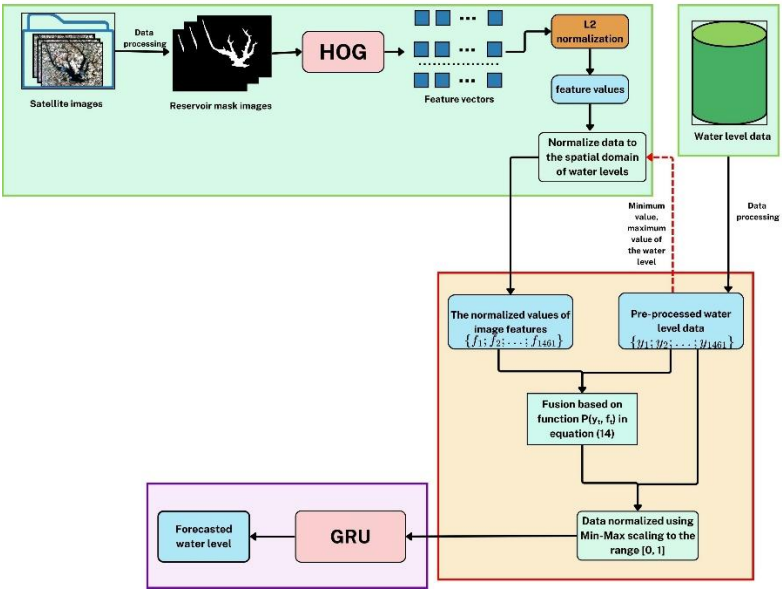

Figure 2.1. Proposed HOG-GRU forecasting model

2.2.2. Data processing

Mask images of the An Khê reservoir area were collected and labeled as described in Section 3.2 (Chapter 3) for the period 01/01/2019–31/12/2022, yielding 18 Sentinel-2 images labeled as boundary masks and stored in .txt format, with pixel value 0 for the object and 255 for the background. The mask images were converted to grayscale to reduce data complexity while preserving necessary information (Figure 2.2), then resized to 320 × 320 × 1. Next, an MMSE estimator was used to reduce noise and improve image quality, combined with multi-point interpolation based on the Pearson correlation coefficient to fill missing months, ensuring one representative image per month. The interpolation added 30 images, bringing the total number of mask images to 48. Finally, the data were smoothed, assigned monthly time indices, and normalized using Min–Max normalization in a Python environment.

The operational data for the An Khê and Ka Nak reservoirs, after preprocessing, consist of 1,461 samples indexed sequentially by day, prioritizing consistency and the ability to integrate into the forecasting model; the input water-level time series were normalized to [0,1].

Figure 2.2. Illustrative example of the An Khê reservoir mask image

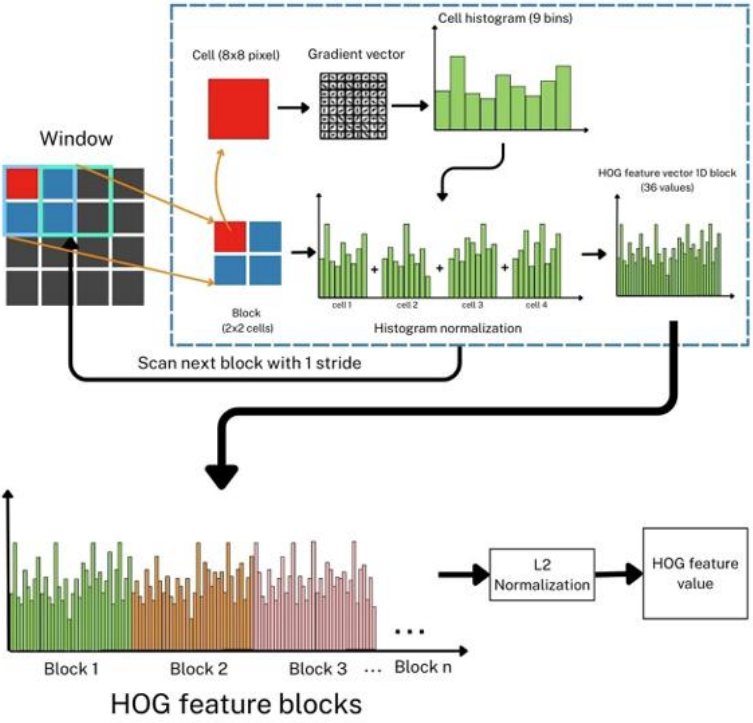

Figure 2.3. HOG algorithm diagram

2.2.3. Image feature extraction and noise handling

Reservoir mask images were intensity-normalized to the range [0,1] and input to the HOG algorithm. Each mask image is represented by a single feature vector. The lengths of these vectors depend on the input image size and are L2-normalized to ensure stability and consistency of the features. After this step, a set of HOG feature values is obtained, as illustrated in Figure 2.3.

L2 normalization is applied according to the following formula:

$$g_t = \sqrt{\sum v_i^2} \quad (2.1)$$

where: v_i is the value of each element in the HOG feature vector, and g_t denotes the overall intensity of the feature, emphasizing salient values and reducing the influence of small noise in the image.

2.2.4. Fusion of water-level data and image features

The fusion function $P(y_t, f_t)$ is expressed in formula (3.1):

$$P(y_t, f_t) = \frac{y_t f_t}{1+f_t}\left[1 - S(y_t - a)\right] + \mu\left[S(y_t - a) - S(y_t - b)\right] + \left(y_t + \frac{\mu}{f_t}\right)S(y_t - b) \quad (2.2)$$

Where:

$$S(x) = \begin{cases} 0; x \le 0 \\ 3x^2 - 2x^3; 0 < x < 1, \\ 1; x \ge 1 \end{cases} \quad (2.3)$$

y_t is the water-level value at time t and $y_t>0$; f_t is the scaling/coefficient associated with each y_t, representing the image-derived feature value that has been normalized to the same space as the reservoir water-level data, with $f_t>0$. $\bar{y}$ is the mean value of the y_t series, and a and b are, respectively, the upper and lower bounds of the confidence interval, determined by:

$$a = \mu - \frac{\sigma}{\sqrt{N}} * \alpha \quad (2.4), \qquad b = \mu + \frac{\sigma}{\sqrt{N}} * \alpha \quad (2.5)$$

The parameter α is not a learning rate but a z-score coefficient used to define the confidence interval for the water-level series. In this study, α = 3.291 corresponding to a 99.9% confidence level. Values within the interval [*a, b*] are considered normal, while values outside this interval are identified as anomalous and highlighted by the fusion function.

2.2.5. Model implementation and GRU training

Reservoir water-level data and the image-derived fused features were Min–Max normalized to [0,1] and split into training, validation and test sets in the ratio 7 : 1.5 : 1.5.

This study builds a GRU model to forecast reservoir water level from a multivariate time series with input shaped as (time steps, features). The architecture consists of two GRU layers (64 and 32 units) interleaved with Dropout layers of 30%, and a final Dense output layer with a sigmoid activation for the normalized forecast value. The model is trained using the Adam optimizer (lr = 1 x 10^{-3}), MSE loss, for 300 epochs, batch size 32, with ModelCheckpoint to save the best-performing model.

2.3. Proposal of the multimodal reservoir water-level forecasting model Colubrid-Net

Reservoir water-level forecasting is a complex spatio-temporal prediction problem that requires capturing both the temporal hydrodynamic behavior and the reservoir's spatial characteristics. Formally, the problem can be described as learning a mapping function:

$$f : X \times S \rightarrow R$$

Where, $X \subseteq R^{T \times d}$ enotes the multivariate time-series dataset consisting of T time steps with d features, $S \subseteq R^{H \times W \times C}$ denotes the satellite image data with height H, width W and C spectral channels, and the output is a scalar value representing the forecasted water level.

2.3.1. Idea of the proposed method

The thesis proposes a novel multimodal architecture that integrates a U-Net-based spatial feature extractor with time-series modeling, as illustrated in Figure 2.4. The core idea is that reservoir water-level dynamics are jointly influenced by short-term temporal patterns in observational hydrological data and by relatively stable, long-term spatial features of the reservoir system.

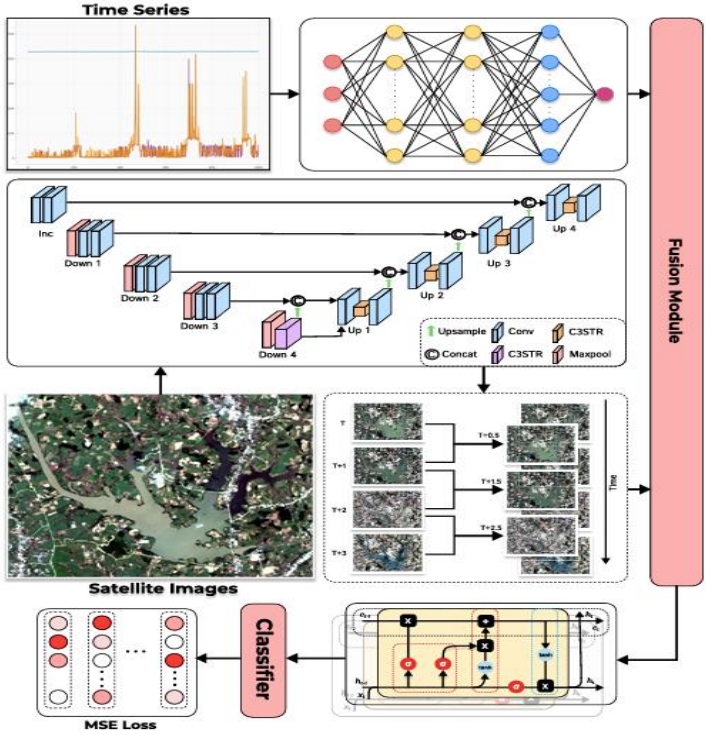

Figure 0.1. Colubrid-Net architecture

The model processing pipeline comprises three main components:

1. U-Net-based multiscale spatial feature extractor (UMSFE): processes reservoir satellite images to extract hierarchical multiscale spatial representations;
2. Multilayer perceptron temporal feature extractor (MTFE): transforms the time-series data into high-density temporal embedding vectors;
3. Temporal fusion Bi-LSTM (TF-BiLSTM): performs joint spatio-temporal inference on the fused multimodal features.

The principal contribution of the proposed method is exploiting multiscale feature extraction from both the encoder and decoder paths of U-Net to capture local patterns and global context simultaneously, thereby improving spatial representation and water-level forecasting accuracy. The model combines spatial features from satellite imagery (UMSFE) and temporal features from time-series data (MTFE), uses temporal interpolation to synchronize resolutions, fuses modalities via concatenation, and forecasts water level using a Bi-LSTM trained with MSE loss.

Basic concepts: Denote $I_t \in R^{H\times W}$ as the satellite image observing the reservoir at time t, and $x_t \in R^d$ as the corresponding multivariate time-series vector, which includes water level, outflow, inflow, and engineered features such as seasonal indicators and lagged variables.

The objective is to forecast the water level y_{t+1} at the next time step based on historical observations $\{I_{t-\tau+1},\ldots, I_t\}$ and $\{x_{t-\tau+1},\ldots, x_t\}$ within a regression window of size τ.

2.3.2. UMSFE architecture

1) C3STR Network (C3 Swin Transformer)

The UMSFE architecture uses a C3 Swin Transformer network to overcome the limitations of CNNs in modeling multi-scale spatial relationships in satellite imagery [96]. The original C3 module proposed in YOLOv5 [97] is effective for local feature extraction but has a limited receptive field and therefore cannot fully capture the global relationships required to represent fine-grained details. In this study, reservoir satellite images contain heterogeneous spatial information at multiple scales — from fine surface-water textures to coarse catchment boundaries — which requires architectures capable of robust and efficient multiscale feature extraction.

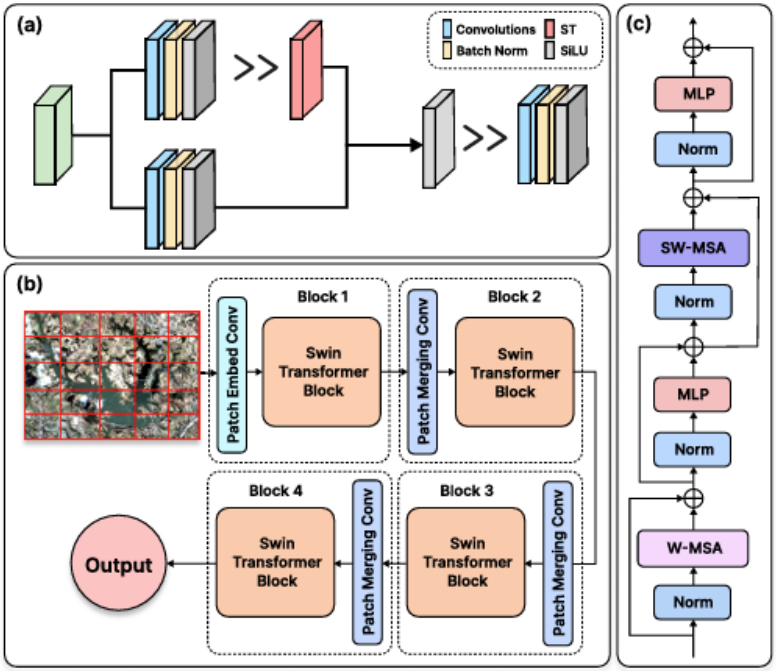

Figure 2.5. Overview of the C3STR module architecture. (a) Two-branch design of C3STR; (b) hierarchical processing structure of the Swin Transformer (ST) module with patch-merging and Transformer blocks; (c) internal structure of an ST block composed of sequential modules.

Swin Transformer (ST) [98] improves on Vision Transformer (ViT) [99] by introducing a hierarchical self-attention mechanism based on shifted windows, reducing computational complexity from quadratic to linear while retaining the ability to model global context. C3STR leverages these advantages via a two-branch architecture, as illustrated in Figure 2.5.(a): one branch processes features using ST to capture long-range dependencies and global context, as shown in Figure 2.5.(b), while the other branch preserves the original features via an identity mapping to ensure stable gradient flow. Each Swin Transformer (ST) block alternates between window-based multi-head self-attention (W-MSA) and shifted-window multi-head self-attention (SW-MSA), as illustrated in Figure 2.5.(c) and expressed by the following formulas:

$$\hat{z}^{l} = \text{W-MSA}\left(\text{LN}(z^{l-1})\right) + z^{l-1} \quad (2.7)$$

$$z^{l} = \text{MLP}\left(\text{LN}(\hat{z}^{l})\right) + \hat{z}^{l} \quad (2.8)$$

$$\hat{z}^{l+1} = \text{SW} - \text{MSA}\left(\text{LN}(z^{l})\right) + z^{l} \quad (2.9)$$

$$z^{l+1} = \mathrm{MLP}\left(\mathrm{LN}(\hat{z}^{l+1})\right) + \hat{z}^{l+1} \quad (2.10)$$

where l is the layer index within the ST block; z^{l-1}, z^{l}, z^{l+1} are the feature representations of the input at successive layers, and $\hat{z}^{l-1}$, $\hat{z}^{l}$, $\hat{z}^{l+1}$ are the intermediate feature representations at layers *l-1*, *l*, *l+1*. These are spatial feature tensors updated sequentially through each Transformer block. LN (Layer Normalization) normalizes across feature channels, stabilizing training and improving convergence.

The window partitioning strategy splits input features into non-overlapping windows of size M×M to perform local self-attention, while shifted windows in subsequent layers enable information exchange across different regions. For reservoir satellite imagery, the C3STR module is capable of capturing scale-aware spatial features — from fine-grained structures such as surface condition patterns to coarser contexts like reservoir geometry and surrounding land-use types.

2) **Dynamic Snake Activation Function**

To enhance nonlinear modeling capability, we use the dynamic Snake activation function [100] for both encoder and decoder branches. The Snake activation provides a learnable periodic activation suitable for modeling complex feature transformations and is defined as:

$$\mathrm{Snake}\ (x) = x + \frac{1}{\alpha}\sin^2(\alpha x) \quad (2.10)$$

where α is a learnable frequency parameter that controls the periodicity and amplitude of the sinusoidal component. Unlike fixed activations such as ReLU or GELU, the learnable α can adapt during training, allowing the model to self-tune its nonlinear characteristics for the specific task, thereby optimizing feature extraction from reservoir satellite imagery.

3) **Proposed architecture**

The spatial feature-extraction module in this study is designed to learn multiscale representations from reservoir satellite images. The encoder follows a hierarchical downsampling strategy described as:

$$e_1 = \mathrm{DoubleConv}(I_t) \quad (2.11)$$

$$e_i = \mathrm{Down}(e_{i-1}),\ i \in \{2, 3, 4\} \quad (2.12)$$

$$e_5 = \mathrm{DownC3STR}(e_4) \quad (2.13)$$

where DoubleConv performs two consecutive convolutions followed by batch normalization and activation; Down executes max-pooling prior to applying the double-convolution block; and DownC3STR integrates the C3STR module to enhance the feature representation at the bottleneck. The decoder then reconstructs spatial information via upsampling and skip connections as follows:

$$d_1 = \mathrm{Up}(e_5, e_4) \quad (2.14)$$

$$d_i = \mathrm{Up}(d_{i-1}, e_{5-i}),\ i \in \{2, 3, 4\} \quad (2.15)$$

We define the encoder feature set E={e_3,e_4,e_5} and the decoder feature set D={d_1,d_2,d_3,d_4}. Each feature map is then processed with Generalized Mean pooling (GeM), defined by:

$$GeM(F) = \left(1/HW \sum_{i=1}^{H} \sum_{j=1}^{W} F_{i,j}^{p}\right)^{1/p} \quad (2.16)$$

Where p is a learnable parameter. The final spatial representation is obtained by concatenating (denoted⊕) all pooled features and applying Layer Normalization:

$$f^{spa} = \text{LayerNorm}\left(\bigoplus \left(\bigcup_{F \in \mathcal{E} \cup \mathcal{D}} \text{GeM}(F)\right)\right) \quad (2.17)$$

which is then projected through a two-layer MLP to produce the final spatial feature vector $z^{spa} \in R^{1024}$.

2.3.3. MTFE architecture

The temporal branch processes multivariate time-series data via MLP layers. First, the input features are standardized as $\tilde{x}_t$ = StandardScale(x_t), and temporal feature extraction proceeds as:

$$h_1 = \text{ReLU}(W_1 \tilde{x}_t + b_1) \quad (2.18)$$

$$h_2 = \text{ReLU}(W_2 h_1 + b_2) \quad (2.19)$$

$$h_3 = \text{ReLU}(W_3 h_2 + b_3) \quad (2.20)$$

$$z^{temp} = \text{ReLU}(W_4 h_3 + b_4) \quad (2.21)$$

Here, W_i and b_i are learnable parameters, and the output vector dimension is $z^{temp} \in R^{1024}$, serving as a compact, information-rich embedding of the hydrological time series before fusion with the spatial feature.

2.3.4. Temporal alignment and image-feature interpolation

To address the temporal incompatibility between monthly satellite imagery and daily time-series data, this thesis develops a temporal interpolation mechanism that expands image features from monthly resolution to daily resolution, while preserving seasonal variation characteristics and introducing realistic temporal dynamics.

Specifically, given a set of monthly image features $\{z_m^{spa}\}_{m=1}^{M}$, where M denotes the number of observed months, we generate a set of daily features $\{z_d^{spa}\}_{d=1}^{D}$ for D days using the following interpolation strategy. For each day d within month m, the base interpolated feature is defined as $z_d^{base} = z_m^{spa}$. o model temporal variations that reflect real changes in reservoir conditions, seasonal and stochastic perturbations are introduced according to the following formulations:

$$s_d = \alpha_s \sin\left(\frac{2\pi \cdot \text{ day\_of\_year } (d)}{365}\right) \quad (2.22)$$

$$\epsilon_d \sim \mathcal{N}(0, \sigma^2 \text{I}) \quad (2.23)$$

$$z_d^{spa} = z_d^{base} + s_d + \epsilon_d \quad (2.24)$$

where α_s=0.05 controls the amplitude of seasonal variation, σ=0.02 determines the magnitude of random perturbations, and I denotes the identity matrix. This interpolation strategy preserves the fundamental spatial features extracted from monthly satellite imagery while introducing controlled temporal variability to reflect the dynamic nature of the reservoir system.

2.3.5. Bi-LSTM for temporal fusion

The multimodal features are fused via concatenation, defined as:

$$z_t = z_t^{spa} \oplus z_t^{temp} \in R^{2048} \quad (2.25)$$

To handle the temporal mismatch between monthly satellite imagery and daily observational time-series data, we employ the temporal interpolation approach to upsample monthly image features to daily resolution, as described in Section 2.3.5. Sequence modeling is performed using a two-layer Bi-LSTM with hidden dimension h=128. The process can be formulated as follows:

$$i_t = \sigma(W_{ii} z_t + b_{ii} + W_{hi} h_{t-1} + b_{hi}) \quad (2.26)$$

$$f_t = \sigma\big(W_{if} z_t + b_{if} + W_{hf} h_{t-1} + b_{hf}\big) \quad (2.27)$$

$$o_t = \sigma(W_{io} z_t + b_{io} + W_{ho} h_{t-1} + b_{ho}) \quad (2.28)$$

$$g_t = \tan \mathrm{h}\big(W_{ig} z_t + b_{ig} + W_{hg} h_{t-1} + b_{hg}\big) \quad (2.29)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (2.30)$$

$$h_t = o_t \odot \tan \mathrm{h}(c_t) \quad (2.31)$$

In the bidirectional setting, the final hidden state at time t is obtained by concatenating the forward and backward hidden states, given by:

$$h_t^{bi} = h_t^{fwd} \oplus h_t^{bwd} \quad (2.32)$$

The final prediction is then produced through a two-layer output head:

$$\hat{y}_{t+1} = W_{out}\, ReLU\big(W_{proj}\, h_t^{bi} + b_{proj}\big) + b_{out} \quad (2.33)$$

to obtain a scalar value representing the forecasted reservoir water level at time step *t+1*.

2.4. Chapter conclusion

This chapter proposes two multimodal reservoir water-level forecasting models, namely HOG-GRU and Colubrid-Net. The results presented in Chapter 2 have been published in works [CT1, CT2], as listed in the *"List of Scientific Publications"*.

CHAPER 3. EXPERIMENTAL RESULTS AND DEVELOPMENT OF THE RESERVOIR WATER LEVEL FORECASTING MODEL

3.1. Study Area

The An Khê – Ka Nak reservoir system (Gia Lai Province) was selected as the representative study area due to its complex natural conditions, steep terrain, and inter-reservoir operational regime, which result in rapid and nonlinear fluctuations in water levels. The two reservoirs are located in a region frequently affected by heavy rainfall, flash floods, and prolonged droughts, with large and difficult-to-predict water level variations. This area presents a high risk for flood regulation and water resource management, thereby requiring accurate and timely water level forecasting models, particularly when integrating multimodal data sources.

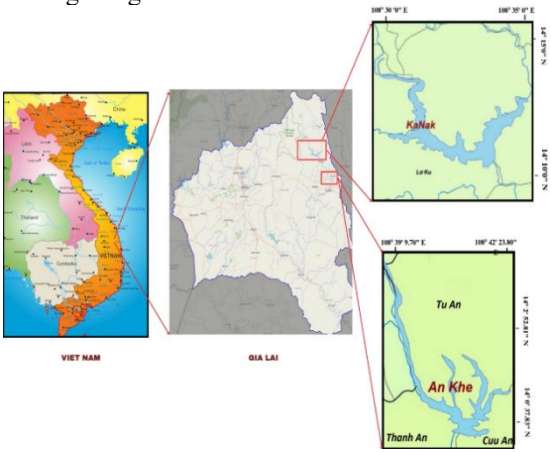

Figure 3.1. Geographic Location of the An Khê and Ka Nak Reservoirs

3.2. Sentinel-2 Satellite Image Collection and Annotation Process

In this section, the thesis describes in detail the process of collecting and annotating Sentinel-2 satellite images. The satellite imagery was obtained from the Copernicus Programme, a key component of the European Union's space initiative. These images were acquired from the Sentinel-2 satellites. Currently, the Copernicus Programme allows global users to freely access the Open Hub platform.:

- Processing level selection: The processing level was set to Level 2A, with a cloud coverage threshold of less than 20%;
- Data collection period: Satellite images for the An Khê reservoir area were collected from January 1, 2018 to December 31, 2023;
- Selected spectral bands: The chosen color bands include Band 4 (Red – R4), Band 3 (Green – G3), and Band 2 (Blue – B2).

With the above selection criteria, the retrieved images exhibit high clarity and closely resemble the natural visual appearance of the reservoir. As a result, 34 satellite images covering the An Khê reservoir area were obtained.

After downloading, the satellite images were processed to extract the An Khê reservoir region using the SNAP software. The image annotation process was conducted according to consistent and appropriate criteria in order to construct a dataset with corresponding labels for each image. Within the SNAP software, the following steps were performed to determine the geographic coordinates of the reservoir area:

Step 1: Select the ZIP file and open it with the color bands R4 (Red), G3 (Green), and B2 (Blue). Zoom in on the displayed image to the location of the An Khê reservoir.

Step 2: Use the rectangle drawing tool to delineate the An Khê reservoir area.

Step 3: Extract the coordinates in WKT (Well-Known Text) format from the Geometry panel and temporarily store them in memory.

Step 4: Open GraphBuilder, create a subset (selecting bands 4, 3, and 2), and save the graph.

Step 5: Edit the polygon in the graph using the coordinates corresponding to the reservoir area.

Step 6: Export the final output file in Mygraph.xml format.

The data files were downloaded from the Copernicus website and processed using the SNAP software. The polygon coordinates were applied to the reservoir area, after which the software automatically generated cropped images in PNG format. The resulting images correspond to the An Khê reservoir region and are illustrated in Figure 3.3.

Figure 3.3. Example of Sentinel-2 imagery of the An Khê Reservoir area.

Figure 3.4. Labeled image of the reservoir.

From the collected images, the next step involved annotating the water surface area of the An Khê reservoir using the CVAT tool within an object segmentation labeling environment with a single class (reservoir). Specifically, the satellite image files were uploaded and labeled according to the following principles: the region to be segmented corresponds to the water surface, characterized by relatively uniform

and continuous color features; points were placed to define the boundary between water and non-water areas; these points were then connected to form a polygon that most accurately represents the shape of the reservoir, as illustrated in Figure 3.4.

After completing the annotation process, the data were stored in a MongoDB database with two main fields: "**images**" and "**labels**".

3.3. Experimental data description

Sentinel-2 imagery for the An Khê and KaNak reservoir areas was collected from January 1, 2019 to December 31, 2022, following the procedure described in Section 3.2. A total of 18 images were obtained for the An Khê reservoir and 32 images for the KaNak reservoir.

Operational data for both reservoirs were also collected over the same period. The recorded variables include timestamp, reservoir water level, inflow discharge, and outflow discharge. These data were acquired on an hourly basis. Subsequently, the data were preprocessed to compute daily average water levels, resulting in 1,461 daily observations. (Illustrated in *Figure 3.5*)

	Hồ chứa	LVS/Hồ/Ngày/Giờ	LVS/Hồ chứa/Ngày	Mực nước hồ (m)	Lưu lượng đến hồ (m³/s)	Tổng lưu lượng xả (m³/s)[Thực tế]
2	An Khê	0	2/4/2022	428.37	40.15	49
3	An Khê	1	2/4/2022	428.36	40.23	49
4	An Khê	2	2/4/2022	428.35	40.32	49
5	An Khê	3	2/4/2022	428.34	40.26	49
6	An Khê	4	2/4/2022	428.33	40.19	49

Figure 3.5. Example Illustration of Reservoir Operational Data

3.4. Evaluation of the HOG-GRU Model Performance

The study compares the proposed HOG-GRU model with several advanced forecasting models commonly used for time-series data, including RNN, LSTM, GRU, HOG-RNN, and HOG-LSTM (Table 3.1).

Table 3.1. Performance evaluation metrics of the forecasting models applied to the entire test dataset of the An Khê reservoir.

Mô hình	MSE	RMSE	MAE	Tracking Signal	Phương sai
RNN	0.10637	0.32615	0.24107	-0.17156	0.10773
HOG-RNN	0.08613	0.29348	0.20533	0.16751	0.08745
LSTM	0.11056	0.33250	0.23836	-0.12391	0.10433
HOG-LSTM	0.08604	0.29332	0.20479	-0.05545	0.08608
GRU	0.10527	0.32445	0.23795	-0.18261	0.10108
HOG-GRU	**0.08060**	**0.28390**	**0.20446**	**-0.00032**	**0.08473**

In addition, the study evaluates the model performance on a dataset containing anomalous data points to verify its forecasting capability under nonlinear and extreme conditions. The corresponding results are presented in Table 3.2.

Table 3.2. Performance evaluation metrics of the models on the dataset containing anomalous data points of the An Khê reservoir.

Model	MSE	RMSE	MAE	Tracking Signal	Phương sai
RNN	0.11297	0.33611	0.24897	0.15842	0.12088
HOG-RNN	0.09136	0.30230	0.21256	-0.16571	0.09974
LSTM	0.11699	0.34203	0.24521	0.11173	0.11658
HOG-LSTM	0.09120	0.30199	0.21249	0.05765	0.10191
GRU	0.11133	0.33367	0.24472	0.17023	0.11204
HOG-GRU	**0.08546**	**0.29234**	**0.21044**	**-0.02169**	**0.09783**

The thesis also conducted experimental implementation and evaluation of the models on the Ka Nak reservoir dataset, with the results presented in Table 3.3.

Table 3.3. Performance evaluation metrics of the reservoir water level forecasting models applied to the entire test dataset of the Ka Nak reservoir.

Mô hình	MAE	MSE	RMSE	Tracking Signal	Phương sai
RNN	0.95277	1.11083	1.05396	-0.46232	19.33899
LSTM	0.65120	0.63653	0.79783	-0.02049	21.21868
GRU	0.62975	0.58228	0.76307	0.01026	22.01869
HOG-RNN	0.50916	0.39607	0.62934	-0.54665	23.75432
HOG-LSTM	0.45496	0.31161	0.55822	-0.43542	24.72102
HOG-GRU	**0.37937**	**0.20795**	**0.45601**	**-0.03985**	**22.69440**

In this section, the thesis conducts experimental comparisons between the two proposed multimodal forecasting models, VGG19-SARIMAX [CT4] and HOG-GRU. The experimental results are presented in Table 3.6.

Table 3.6. Comparison of performance evaluation metrics between the VGG19-SARIMAX and HOG-GRU models.

Model	MAE	MSE	RMSE	Tracking Signal	Train time (minutes)
VGG19-SARIMAX	0.26858	0.10492	0.32392	0.14551	2880
HOG-GRU	**0.20408**	**0.08355**	**0.28905**	**0.02625**	**2**

- **Remarks:**

The thesis conducted a comprehensive series of experiments to evaluate the performance of the HOG-GRU model, with the results demonstrating that the proposed model outperforms the comparative methods. These findings not only confirm the effectiveness of the proposed data fusion strategy but also show that using HOG for image feature extraction significantly reduces computational costs while maintaining forecasting accuracy, compared with approaches that rely entirely on deep learning networks.

In this evaluation, the thesis further analyzes and assesses the impact of parameters and components within the fusion function on the forecasting performance of the proposed HOG-GRU model.

- ***How does adding weights to the components of the fusion function $P(y_t, f_t)$ affect the forecasting performance of the prediction model?***

Formula of the fusion function $P(y_t, f_t)$ after introducing the weights w_1, w_2 and w_3, transformed into the modified functionc $P'(y_t, f_t)$ as follows:

$$P'(y_t, f_t) = w_1 \frac{y_t f_t}{1+f_t}\left[1 - S(y_t - a)\right] + w_2 \mu\left[S(y_t - a) - S(y_t - b)\right] + w_3\left(y_t + \frac{\mu}{f_t}\right)S(y_t - b) \quad (3.1)$$

By examining the formula $P'(y_t, f_t)$, it can be observed that when the weights are equal and all set to 1, the equation reduces to the original form $P(y_t, f_t)$. The case where $w_1 = w_2 = w_3 = 1$, is referred to as **Scenario 1.**

Next, additional experiments were conducted under three different scenarios, in which the weight values were adjusted as follows:

- **Scenario 2:** $w_1 = 2, w_2 = 1, w_3 = 1$
- **Scenario 3:** $w_1 = 1, w_2 = 2, w_3 = 1$
- **Scenario 4:** $w_1 = 1, w_2 = 1, w_3 = 2$

In these scenarios, w_1 is used to emphasize the lower bound, w_2 emphasizes the mean (normal) points, and w_3 emphasizes the upper bound. The experimental results are presented in Table 3.7.

Table 3.7. Performance evaluation results of the HOG-GRU model under different weighting scenarios of the fusion function.

	MAE	MSE	RMSE	Tracking Signal	Phương sai
Kịch bản 1	**0.20408**	**0.08355**	**0.28905**	**0.02625**	**0.08451**
Kịch bản 2	0.20619	0.08380	0.28955	0.10364	0.09358
Kịch bản 3	0.20901	0.08544	0.29230	0.18383	0.09358
Kịch bản 4	0.21381	0.08966	0.29944	-0.09678	0.09730

- ***Evaluation of the impact of the value α in the fusion function $P(y_t, f_t)$ on the model performance***

Table 3.8. Performance evaluation results of the HOG-GRU model with different values of α\alphaα in the fusion function.

	MAE	MSE	RMSE	Tracking Signal	Phương sai
α = 1.645(90%)	0.21223	0.08692	0.29481	-0.05273	0.09281
α = 1.960(95%)	0.20939	0.08349	0.28894	-0.05881	0.07864
α = 3.291(99.9%)	**0.20408**	**0.08355**	**0.28905**	**0.02625**	**0.08451**

Based on Table 3.8, the parameter α has a noticeable impact on the performance of the HOG-GRU model. When α decreases, the MAE, MSE, and RMSE errors slightly increase, indicating that a narrower water level confidence interval reduces model stability and prediction accuracy. At the same time, the Tracking Signal shifts from slightly positive to negative values, reflecting the presence of systematic bias in the forecasts. However, the mean value of the predicted series remains nearly unchanged, suggesting that α does not significantly affect the overall average trend.

- ***Evaluation of the contribution of each component in the fusion function to the performance of the forecasting model.***

This analysis aims to quantify the role and effectiveness of each factor in the fusion mechanism by sequentially removing each enhanced component and replacing it with the corresponding baseline representation.

- **Scenario 1:** The lower-bound emphasis component is removed; accordingly, the fusion function $P(y_t, f_t)$ is adjusted to $P_1(y_t, f_t)$.

$$P_1(y_t, f_t) = y_t\left[1 - S(y_t - a)\right] + \mu\left[S(y_t - a) - S(y_t - b)\right] + \left(y_t + \frac{\mu}{f_t}\right)S(y_t - b) \qquad (3.2)$$

- **Scenario 2:** The component emphasizing the mean water level is removed; accordingly, the function $P(y_t, f_t)$ is adjusted to $P_2(y_t, f_t)$.

$$P_2(y_t, f_t) = \frac{y_t f_t}{1 + f_t}\left[1 - S(y_t - a)\right] + y_t\left[S(y_t - a) - S(y_t - b)\right] + \left(y_t + \frac{\mu}{f_t}\right)S(y_t - b) \qquad (3.3)$$

- **Scenario 3:** The upper-bound emphasis component is removed; accordingly, the fusion function $P(y_t, f_t)$ is adjusted to $P_3(y_t, f_t)$.

$$P_3(y_t, f_t) = \frac{y_t f_t}{1 + f_t}\left[1 - S(y_t - a)\right] + \mu\left[S(y_t - a) - S(y_t - b)\right] + y_t S(y_t - b) \qquad (3.4)$$

- **Kịch bản 4:** Scenario 4: The fusion function $P(y_t, f_t)$ remains unchanged as the original formulation.

Table 3.9. Performance evaluation of the HOG-GRU model when individually removing each component from the fusion function $P(y_t, f_t)$.

	MAE	MSE	RMSE	Tracking Signal	Phương sai
Kịch bản 1	0.20611	0.08745	0.29572	0.20639	0.09589
Kịch bản 2	0.20746	0.08442	0.29055	0.03358	0.09525
Kịch bản 3	0.21508	0.08517	0.29184	-0.20966	0.08571
Kịch bản 4	**0.20408**	**0.08355**	**0.28905**	**0.02625**	**0.08571**

- **Remarks:**

The thesis conducted experiments across multiple scenarios for the components of the fusion function $P(y_t, f_t)$, the results consistently indicate that the original fusion function used in the proposed method delivers the best performance.

3.5. Performance Evaluation of the Colubrid-Net Model

Experiments were conducted on Sentinel-2 imagery and operational data of the An Khê, reservoir, with preprocessing procedures and configurations standardized according to the HOG-GRU framework to ensure fair comparisons. The results (Table 3.10) show that models without spatial feature integration exhibit large errors (RF: MAE = 0.7329; MLP: MAE = 4.7447), while incorporating spatial information significantly improves performance (VGG19+GRU: MAE = 0.2356).

Among advanced methods, the temporal transformer achieves MAE = 0.026 but shows a low R^2 whereas the study by Chau et al. reports MAE = 0.117. Colubrid-Net achieves the best performance with MAE = 0.0242, R^2 = 0.969 and KGE = 0.972,

representing improvements of 7.4% compared to the temporal transformer and 79.3% compared to Chau et al. These results confirm the critical role of spatial feature extraction using transformers combined with temporal interpolation in enhancing the accuracy and reliability of reservoir water level forecasting.

Table 3.10. Comparative results table

Model	Spatial Features	MAE	MSE	RMSE	R^2	KGE
Linear Regression	No	1.730	4.386	2.094	0.175	0.118
Ridge Regression	No	1.730	4.387	2.094	0.175	0.118
Lasso Regression	No	1.726	4.483	2.117	0.157	0.084
SV Regression	No	0.419	0.366	0.605	0.931	0.9972
Random Forest	No	0.732	0.8971	0.947	0.831	0.894
Gradient Boosting	No	0.493	0.442	0.665	0.916	0.959
XGBoost	No	2.941	18.280	4.275	-2.436	0.140
LightGBM	No	0.577	0.553	0.744	0.895	0.950
CatBoost	No	0.537	0.504	0.710	0.905	0.957
AdaBoost	No	0.200	0.447	0.306	0.911	0.954
Decision Tree	No	2.828	19.925	4.463	-2.746	-0.051
K-Nearest Neighbors	No	3.025	18.342	4.282	-2.448	0.205
MLP	No	4.744	45.808	6.768	-7.612	-0.665
VGG19 w/GRU	Yes	0.235	0.129	0.360	0.929	0.967
U-Net w/GRU	Yes	0.215	1.861	1.364	0.482	0.633
Proposed w/GRU	Yes	0.031	0.007	0.008	0.972	0.967
VGG w/ Bi-LSTM	Yes	0.026	0.002	0.047	0.936	0.965
U-Net w/Bi-LSTM	Yes	0.032	0.002	0.049	0.967	0.965
Temporal Transformers [101]	No	0.026	0.009	0.030	0.381	0.703
Chau et al. [CT4]	Yes	0.117	0.268	0.145	-	-
Colubrid-Net	Yes	0.024	0.002	0.046	0.969	0.972

To evaluate the impact of different temporal modeling approaches, Table 3.11 presents a comparison of various recurrent architectures within the proposed model. The results indicate that leveraging bidirectional temporal context enables the model to effectively capture both historical dependencies and “forward-looking” trends in reservoir system dynamics, thereby confirming the suitability of selecting Bi-LSTM in the Colubrid-Net architecture.

Table 3.11. Performance Comparison of Recurrent Architectures

Architecture	MAE	MSE	RMSE
MLP	0.1235	0.0168	0.1296
GRU	0.0236	0.0027	0.0523
Bi-GRU	0.0415	0.0035	0.0595
LSTM	0.0255	0.0023	0.0485
Bi-LSTM	**0.0242**	**0.0022**	**0.0464**

Table 3.12 further clarifies the important role of the snake activation function in the proposed architecture. The snake/snake configuration achieves superior performance compared to the remaining configurations.

Table 3.12. Comparison of activation functions in the encoder and decoder.

Activation Function (Encoder/Decoder)	MAE	MSE	RMSE
ReLU/ ReLU	0.2481	0.0714	0.2673
ReLU/ Snake	0.2848	0.0883	0.2971
Snake/ ReLU	0.3015	0.1046	0.3234
Snake/ Snake	**0.0242**	**0.0022**	**0.0464**

3.6. Multimodal Water Level Forecasting System

3.6.1. System Architecture

The system is designed based on a client–server architecture consisting of four main components (*Figure 3.7*): the web interface, backend, AI module, and database. The web interface enables user interaction with the system; the backend (Node.js) executes business logic, queries the MongoDB database, and sends inference requests to the AI model; the AI module performs data processing and forecasting; and the database stores satellite images, time-series data, and forecasting results. The inference outputs from the AI model are processed by the backend and returned to the web interface via HTTP responses, ensuring system scalability and practical applicability.

3.6.2. System Workflow

The system workflow (Figure 3.8) begins when users interact with the web interface, and the frontend sends HTTP requests to the backend. The backend validates input data, stores and retrieves the required data, and then forwards them to the AI module for processing, including: image segmentation using YOLOv9, image feature extraction using VGG19, numerical feature extraction using FCN, and water level forecasting using SARIMAX. The inference results are stored in the database and returned to the frontend in JSON format for visualization. The backend (Node.js) acts as the central coordinator, communicating with the AI model via Python and

managing MongoDB through RESTful APIs, thereby ensuring scalability, efficient processing, and real-world applicability of the system.

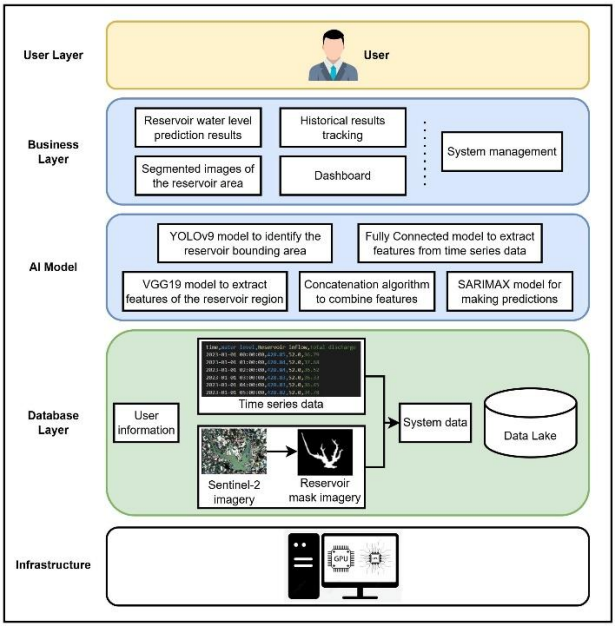

Figure 3.7. Architecture of the Reservoir Water Level Forecasting System

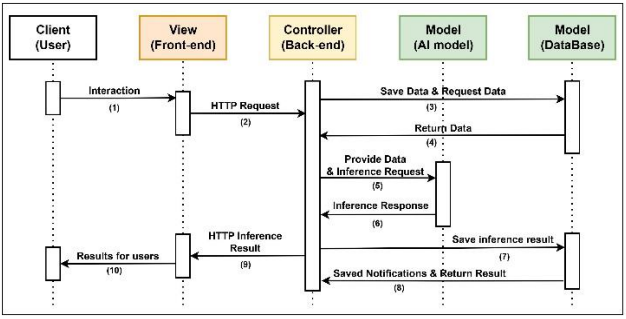

Figure 3.8. Workflow Diagram of the Reservoir Water Level Forecasting System

3.7. Chapter Conclusion

The results of this chapter have been published in works [CT4, CT5] listed in the "*List of Scientific Publications*" section.

CONCLUSION AND FUTURE DIRECTIONS

1) Main Contributions of the Thesis

In the context of climate change increasing the frequency and intensity of extreme weather events, improving the accuracy of reservoir water level forecasting plays a critical role in reservoir management and operation. Current forecasting approaches, ranging from traditional hydrological models to machine learning methods, still exhibit limitations in terms of accuracy, generalization capability, and detailed input data requirements.

The development of remote sensing technology has provided abundant, continuous satellite imagery with broad spatial coverage for this task. At the same time, deep learning models demonstrate advantages in feature extraction and modeling complex nonlinear relationships. Therefore, integrating remote sensing data with deep learning has become a promising approach to enhance reservoir water level forecasting performance under dynamic hydro-meteorological conditions.

Based on theoretical synthesis, model analysis, and experimental evaluations on reservoir datasets, the thesis achieves the following main contributions:

- Proposing the HOG-GRU multimodal forecasting model
- Proposing the Colubrid-Net multimodal forecasting model
- Developing a web-based multimodal reservoir water level forecasting framework, which enables users to visualize data, monitor reservoir water level variations, and effectively support monitoring, analysis, and decision-making processes for reservoir management authorities.

The outcomes of the thesis have strong potential for practical applications, particularly in the development of reliable multimodal reservoir water level forecasting systems that effectively support water resource management, reservoir operation, and decision-making under changing hydro-meteorological conditions.

2) Future Research Directions

In the future, the following directions may further improve the performance and applicability of reservoir water level forecasting models.

First, optimizing hyperparameters using intelligent search algorithms such as particle swarm optimization or genetic algorithms is a promising approach to enhance prediction accuracy and stability.

In addition, integrating more hydrological variables from diverse data sources-such as inflow, outflow, and other auxiliary observational data—can help the model better capture nonlinear relationships and complex physical mechanisms within the reservoir system.

Another notable direction is the application of Multi-Task Learning to simultaneously forecast multiple closely related hydrological variables. This approach enables effective exploitation of cross-variable dependencies, thereby improving overall forecasting performance.

LIST OF SCIENTIFIC PUBLICATIONS

[CT1]	**Hoang Thị Minh Chau**, Tran Thi Ngan, Nguyen Long Giang, Nguyen Nhu Son (2026), "HOG-GRU: a novel multi-modal water level forecasting model integrating satellite imagery and reservoir operation data". *Data Technologies and Applications*, in press. DOI: 10.1108/DTA-07-2025-0545 **(SCIE, Q2).**
[CT2]	Nguyen Duc Quang Anh, Nguyen Minh Anh, Tran Thi Ngan, **Hoang Thi Minh Chau** (2025). Colubrid-Net: A unified cross-modal framework for hydrological forecasting in An Khe Reservoir, Vietnam. in *IEEE Geoscience and Remote Sensing Letters*, vol. 23, pp. 1-5, 2026, Art no. 1500905. DOI: 10.1109/LGRS.2025.3645672 **(SCIE, Q1**).
[CT3]	**Hoang Thi Minh Chau,** Tran Thi Ngan, Nguyen Long Giang, Tran Manh Tuan, Tran Kim Chau. (2025). A Hybrid Deep Learning Method for Forecasting Reservoir Water Level from Sentinel-2 Satellite Images. *Computers, Materials & Continua*, *83*(3). DOI: 10.32604/cmc.2025.062784 (**SCIE, Q3**).
[CT4]	**Hoang Thi Minh Chau,** Tran Thi Ngan, Nguyen Long Giang, Tran Kim Chau, Hoang Duc Trung, Ton Nu Mai Khanh. (2024, December). An Integration of VGG19 and SARIMAX in Water Level Forecasting Using Satellite Imagery and Time Series Data. In *International Conference on Smart Objects and Technologies for Social Good* (pp. 16-28). Springer Nature Switzerland. **Scopus Q4**.
[CT5]	**Hoang Thi Minh Chau,** Tran Thanh Lam, Hoang Duc Trung, Tran Thi Ngan, Nguyen Long Giang, Nguyen Nhu Son, Phung The Huan. (2025). A new multi-modal forecasting system for water level estimation. In *International Conference on Advances in Information and Communication Technology*. Springer Nature Switzerland. **Scopus Q4**.