A Digital Twin Approach to Advancing River Emergency Response Systems in Smart Cities

Yueming Zhu
Georgia Institute of Technology
yzhu737@gatech.edu

Xiyu Pan
Georgia Institute of Technology
xyp@gatech.edu

Neda Mohammadi
Georgia Institute of Technology
The University of Sydney
nedam@gatech.edu
neda.mohammadi@sydney.edu.au

John E. Taylor
Georgia Institute of Technology
jet@gatech.edu

Abstract

Drowning incidents and floods in river cities have become a significant public safety concern worldwide. These incidents result in numerous deaths and injuries, highlighting the urgent need for effective monitoring and rescue systems. Traditional safety detection systems for bodies of water are primarily designed for controlled indoor environments, such as indoor pools, where conditions are stable and predictable; relying on sensors and wearable devices, which are not practical for the varied and challenging conditions of outdoor environments (e.g., distance, wider monitoring areas, and environmental factors such as waves). In response to this challenge, we propose a river emergency response system based on a digital twin model, supported by a human detection model, a water level prediction model, and related algorithms. This piloted system employs a single overhead camera as the primary hardware sensor for continuous real-time safety monitoring. We focus on the Chattahoochee River in the Columbus-Phoenix City area, where drowning incidents have surged in recent years. This system aims to improve rescue response time by generating multi-level of danger alerts based on varying real-time conditions.

Keywords: Digital Twins, Drowning Detection, Smart City, You Only Look Once (YOLO), Real-Time River Safety Monitoring

1. Introduction

Drowning is a serious problem in both indoor pools and outdoor open water environments (Jalalifar et al., 2024), responsible for approximately 7% of all injury-related deaths and ranked as the third leading cause of accidental injury-related death by the World Health Organization (World Health Organization, 2024). This issue is exacerbated in river cities, where natural bodies of water can pose significant risks to public safety. The combination of unpredictable water currents, varying depths, and potential for rapid flooding creates hazardous conditions for residents and visitors alike. During hot weather, people are naturally drawn to rivers and other bodies of water to cool off, increasing the likelihood of drowning incidents. High temperatures can lead to crowded riverbanks, where the lack of lifeguards and safety measures heightens the risk of accidents. The popularity of recreational activities such as swimming, kayaking, and tubing adds to the safety concerns, especially in areas where water levels can change suddenly and dramatically due to environmental factors or upstream water releases.

Summers in Columbus, GA, are characterized by intense heat. By 2050, people in Columbus are projected to experience an average of about 49 days per year with temperatures over 97.0ºF, compared to just 7 days in 1990 (ClimateCheck, 2024). This trend drives people to the Chattahoochee River to cool off. However, these activities are often overshadowed by the inherent risks posed by the river's frequent flooding. From 2017 to 2019, there have been a total of 54 rescue calls and 11 death cases in the Chattahoochee in the Columbus-Phenix City area (ChattVoice, 2024). The area shown in Figure 1 is an area with a significant number of accidents. In addition, first responders have rescued many people who have become stranded on the rocks when dams release water upstream. Both scheduled and unscheduled water releases from upstream dams lead to sudden and unpredictable rises in water levels, creating dangerous conditions that can trap individuals on the small islands scattered throughout the river. However, the current systems do not have

URI: https://hdl.handle.net/10125/109005
978-0-9981331-8-8
(CC BY-NC-ND 4.0)

ability to inform visitors of this danger in advance. This recurring hazard highlights the urgent need for an advanced solution to monitor and respond to these emergencies effectively.

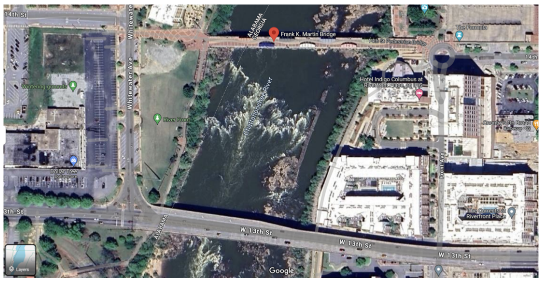

Figure 1. Satellite Image of the Chattahoochee River near Columbus-Phenix City (Maps, 2024)

A smart city digital twin is a dynamic digital replica of a city that is continuously updated with real-time data and analytics on interactions between humans, infrastructure, and technology (Mohammadi and Taylor, 2017). Its importance lies in the ability to provide continuous monitoring, predictive analytics, and optimization of operations, which are essential for managing complex environments, including river rescue systems. By enabling macroscopic observation of situations in a virtual realm, digital twins enhance situational awareness. Additionally, by integrating different types of data into a unified system, digital twins maximize the utility of heterogeneous information and extract valuable insights for improved decision-making.

Current water rescue systems are predominantly designed for indoor swimming pools or rely on sensor-based detection. The controlled environment of indoor pools, with sufficient and consistent lighting, makes it more conducive to the installation and implementation of cameras above or under the water. Additionally, sensor-based detection systems, including wearable devices and environmental sensors such as sonars, generate data in various modalities, thereby enhancing detection capabilities from multiple dimensions. However, in the context of rescuing individuals in outdoor river environments, it is impractical to mandate that every individual to wear a detection device or to implement and maintain such extensive systems. The complexities and dynamic conditions of outdoor environments, such as varying water currents, present significant challenges for traditional detection systems.

In this paper, we present the development of a digital twin-based river emergency response system, implemented for the Chattahoochee River in Columbus, GA. Employing a camera positioned above the water as the primary sensor for detection and situational awareness, the system is designed to enhance real-time monitoring and improve recuse response times to prevent tragedies in such high-risk environments.

2. Prior Studies

2.1. Drowning detection studies

Recent advancements in technology have significantly enhanced public safety in aquatic environments (Cepeda-Pacheco and Domingo, 2022; Kałamajska et al., 2022). These advancements can be broadly categorized into two primary approaches: Image processing-based approaches use computer vision techniques to monitor and detect signs of drowning through video surveillance systems, and sensor-based approaches employ various sensors, such as wearable devices, to monitor physiological parameters and environmental conditions to identify potential drowning incidents.

In the image processing-based approach, a popular Drowning Detection System, called *Poseidon* (Tech, 2024), is based on multiple overhead and underwater cameras to capture the actions of swimmers. (F. Wang et al., 2022) presents a method for early detection of dangerous conditions in the deep-water zone of swimming pools based on video surveillance. (Chan et al., 2020) collected their own dataset in an indoor swimming pool with an underwater camera and then trained a drowning prevention system based on deep learning. Aquatic cameras perform well in pools with high-contrast walls that effectively contrast with all skin tones, making detection easier. In addition, the pool location facilitates the use of underwater cameras and provides a suitable mounting location for such equipment (Eng et al., 2003). In contrast, river environments face challenges such as earthen seabeds and varying water clarity. Sediments such as mud and debris often reduce visibility to unusable levels, thus compromising the effectiveness of these systems (Eng et al., 2008).

In the sensor-based approach, physiological sensors focus on parameters like heart rate, blood oxygen saturation, and so on. (Chaudhari et al., 2016) proposed a remote control-based drowning detection system by comparing heart rate to a threshold. Oxygen saturation is the amount of oxygen circulating in the blood (Clinic, 2024). (Kulkarni et al., 2016) developed a system for detecting drowning using non-invasive sensors for oxygen saturation, respiration, and water presence. Investigations have shown (Shehata et al., 2021) that the accuracy of sensor-based methods can be significantly improved if more parameters are monitored simultaneously. An example is the waterproof device proposed by (Jalalifar et al., 2022) for the detection of drowning. The device includes heart rate, oxygen saturation, body temperature and depth sensors. Each sensor operates autonomously to improve the drowning prevention capability of the system. However, all of this equipment needs to be worn on the swimmer's body, which makes it difficult for outdoor enthusiasts to take advantage of it, as well as to feed information detected by personal devices into the rescue system.

Environmental Sensors include sonar, camera, and so on. (Rooz and Ben-Sira, 1991) first proposed using active sonar for drowning detection, scanning the pool area to differentiate between people and inanimate objects based on the sonar images. (He et al., 2022) introduced the first practical drowning detection system based on underwater sonar, which employs an active ultrasonic sonar and features a novel sonar scanning strategy that balances time and accuracy. Another commonly used environmental sensor is the camera, which is a necessary sensor for image processing-based methods. Among them, the cameras are categorized as overhead and underwater. Overhead cameras are convenient to install and maintain. However, since a drowning person will quickly sink below the water surface, especially in a turbulent current, the overhead camera cannot always capture the drowning incidents. Underwater cameras are widely used for their ability to capture human behavior beneath the water's surface. Nevertheless, they are susceptible to interference from underwater sediments and are more difficult to maintain.

2.2. Smart City Digital Twins

Urban systems are complex and have interdependent components, including human, natural landscapes, buildings, and technological facilities. The earliest modeling of urban spatial data dates back to the emergence of geographic information system (GIS) (Shen et al., 2007) technology in the 1960s. This made possible for the first time the mapping of real-world data into the virtual world. The growing popularity of Global Positioning System (GPS) technology in smartphones and wearable devices, combined with the rapid growth of social media, which enables real-time online information sharing, has brought the relationship between humans and the environment closer and closer.

Topographic LiDAR technology is a remote sensing method that uses laser light to measure distances and create high-resolution maps of the Earth's surface, enabling researchers to generate topographic maps of an area within minutes. Despite the vast, diverse, and expansive data sources available today for virtual modeling, research on integrating and deriving insights from this heterogeneous information remains limited. Integrated virtual city models, are thus, increasingly recognized as a much-needed approach to mining and extracting information from complex urban systems.

Traditional methods of creating virtual urban environments have been constrained to static, physical depictions of the city (Batty, 2008). This static data serves as the foundational framework and provides the backdrop for the scenarios to be simulated. Dynamic parameters, which are our primary focus and the elements we intend to simulate within the environment, are typically sourced from sensors. These dynamic elements are integrated into the framework, representing entities that are monitored and investigated during research. This combination of dynamic and static data accurately and validly projects real-world scenarios into the virtual world, while the degree of statics in these models is associated with the level of abstraction in modeling. Smart city digital twins expand from complete real to complete virtual environments. Although numerous studies have explored virtual modeling of urban information, most have focused on simple visualization of the obtained data and lack comprehensive combined analysis to derive deeper insights (Mohammadi and Taylor, 2020). For example, (Cristie et al., 2015) proposed CityHeat, a system using colo-mapped cubes to relate heat to traffic, and (Sedor, 2017) developed Connected Cities VR, a city maintenance system.

Although underwater cameras and various sensors are widely used for aquatic safety monitoring, as previously mentioned, these devices are more suitable for indoor swimming pools. In outdoor rivers, the accumulation of sediment at the riverbed, rapid water flow, and complex surrounding environments make the installation and maintenance of underwater cameras or other sensors challenging. Additionally, it is impractical to ensure that every visitor wears a detection device. However, overhead cameras can be installed on bridges, making installation and maintenance more feasible. These cameras can capture a continuous stream of real-time images of riverbanks and surrounding areas. High-resolution cameras can mitigate the limitations of unclear top-down images, enhancing outdoor aquatic safety monitoring. Therefore, this paper uses data from overhead cameras to monitor the Chattahoochee river environment. This data is integrated into our smart city digital twin model, which combines multi-dimensional information such as water levels, topographic maps, and computer vision data. By employing advanced algorithmic models, as described in the following section, it can detect and predict potential hazardous scenarios and generate warnings.

3. System Development

Our river emergency response system and the digital twin model consist of four components. As shown in Figure 2, human detection, water level prediction, submerging prediction, and drowning alert cards. For the human detection part, You Only Look Once Version 7 (YOLOv7) (C.-Y. Wang et al., 2023) is used to process live stream from an onsite overhead camera, identifying bounding boxes for people and boats for each frame of the video. The water level prediction part uses a recurrent neural network to predict the future 15-minute and 30-minute water level values to monitor the significant water level rise in the next 30 minutes. The submerging prediction generates alerts when individuals are standing on islands that are predicted to be submerged. It compares the bounding boxes given by the human detection part with the predicted changes to the island areas. Based on the information from the above components, drowning alert cards (yellow, orange, and red) will flash according to the assessed conditions and danger levels by the system.

3.1. Human Detection

We installed an onsite overhead camera provided by Verkada(Verkada, 2024) on the 14th Street bridge, facing south. This camera ensures security through a continuously changing token over time. To maintain uninterrupted access to the live stream, our system requests a new token from the central controller every hour, continuously updating the credentials. This approach enables us to receive a continuous 24/7 live stream of real-time data. Every raw image is resized before being fed into the You Only Look Once (YOLO) network to achieve high inference speed. Balancing resolution and inference speed through a trial-and-error process, we found that resizing the original images from 3840 x 2160 pixels to 1600 x 900 pixels maintains nearly the same accuracy while doubling the processing speed, perfectly meeting the 24fps frame rate requirement for video. Meanwhile, to further enhance the processing speed, we set up two Real Time Messaging Protocol (RTMP) servers using **Nginx** (Reese, 2008). The first server receives the raw video then pushes it to the second server where **YOLOv7** is running. The parallel server strategy reduced the computation time by half compared to a single server setup strategy.

As shown in Figure 3, the result of this step is the bounding boxes around people and boats in the image. For each frame, the backend records the coordinates of the four corners of these bounding boxes, counts the number of detected objects, and logs the information. To enhance accuracy, a sliding window with a length of 96 frames (equivalent to a total duration of 4 seconds) is applied to smooth the data, compensating for false positives and false negatives that may occur during frame-by-frame detection.

3.2. Water Level Prediction

The river data used in this study was obtained from the United States Geological Survey (USGS) and used for model training. The water level is a crucial indicator of river conditions. Its fluctuations can significantly impact the areas of islands within the river and the overall safety of the region. The base elevation of the gauge located at the targeted river segment is 190 feet. By adding 190 feet to the gauge readings, we obtain the elevation of the river's water surface. Precipitation and discharge data were also collected from USGS. The USGS records these data every 15 minutes, which defines the interval between our data points. We use a continuous sequence of 28 data points (equivalent to 7 hours) as the input for each prediction. Thus, each input is represented as a 3x28 matrix. Long Short-Term Memory (LSTM) networks (Wikipedia, 2024a), a type of recurrent neural network that excels in modeling sequential data by retaining long-term dependencies through their unique cell state and gating mechanisms, are ideal for tasks like time series prediction. In this study, the aforementioned 3x28 matrix is used as the input to the LSTM, which outputs water level predictions for the next 15 and 30 minutes, represented as a 1x2 vector. Since the USGS API data has a 15-minute delay, the prediction for 30 minutes ahead corresponds to actual data 15 minutes in the future. The predicted values are then compared to the current values to assess changes. An increase of more than 0.5 feet is defined as significant, triggering a warning, because such an increase can already lead to notable changes in the island areas. Conversely, an increase of less than 0.5 feet is considered a normal state. This study utilized data collected from July 12, 2017, to June 20, 2022. This dataset was divided into three subsets: 70% for training, 15% for validation, and 15% for testing. The final LSTM model achieved a mean squared error of 0.1013 feet, a mean absolute error of 0.2344 feet, and an R-squared value of 0.978, indicating high predictive accuracy.

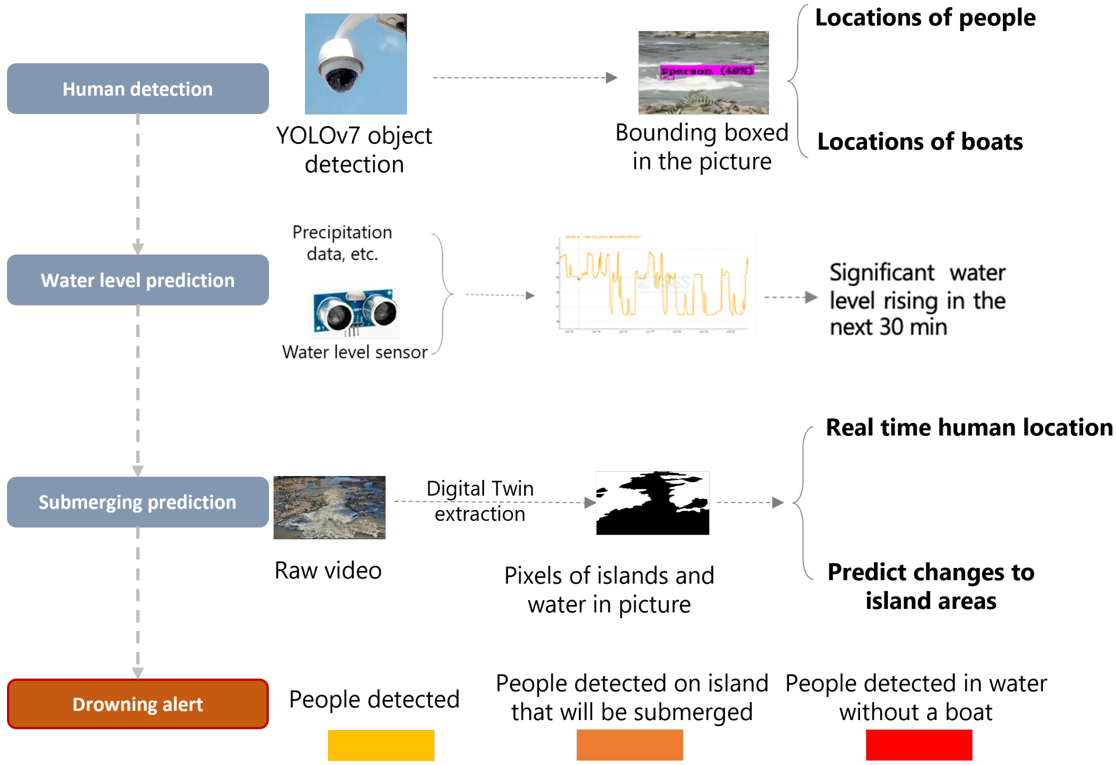

Figure 2. Components and workflow of the smart city digital twin for river safety design; including human detection with YOLOv7, water level prediction using neural networks, submerging prediction, and generation of color-codded drowning detection alerts (yellow, orange, and red).

Figure 3. Bounding boxes for detected people and boats on the Chattahoochee River using YOLOv7.

3.3. Submerging Detection

Since 2020, the water level of the Chattahoochee River has fluctuated between 204 and 225 feet, with the majority of levels clustering between 205 and 211 feet. This range also corresponds to the most significant changes in the river's island regions. When the water level falls below 205 feet or rises above 211 feet, the river's morphology remains largely unchanged. We manually extracted island contours within the 205 to 211 feet range at 0.5-foot intervals. By manually tracing and binarizing the images—representing the islands in white and the water in black (as shown on the right side of Figure 4)—we can distinguish different areas. In this way, if the water level rises, the area of water, which corresponds to the black area in the image will increase. On the contrary, the area of islands, which corresponds to the white area in the image will decrease, and verse versa.

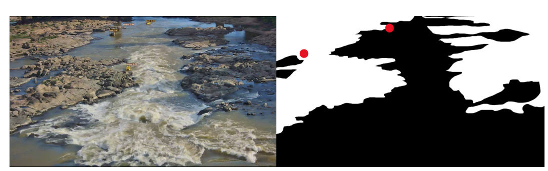

Figure 4. Identification of island areas by comparing the raw video frame with the extracted contour of the islands.

From Section 3.1, we can obtain the bounding box coordinates of people in each frame. By comparing these coordinates with the contour map corresponding to the current water level (shown in Figure 4, where the two red dots indicate the positions of persons on the left in the image), we can determine whether a person is on the island. The system will also store the number of people on the island and in the water in the backend.

3.4. Visualization

To visualize the geographical context, we acquired the topographic map of the area from the Website (Map, 2024) and used graphing library of **Plotly** (Plotly, 2024) to construct a 3D-model of the entire region. The water level data stream, extracted from Section 3.2, is visualized by adjusting the Z-axis value to reflect the current water surface elevation. To project an individual's position from the camera's 2D plane into the digital twin model's 3D space, we used the current water level as the Z-axis coordinate of the target position. As shown in Figure 5, both images are oriented from north to south. In the 3D model, the origin is set at the intersection of the bridge base and the riverbank in the top left corner. A coordinate system is established along the bridge and riverbank directions. The position of the person in the left image is projected onto the right image through a linear transformation. The specific formula is as follows:

$$\begin{cases} w = \frac{x}{X}W \\ h = \frac{y}{Y}H \end{cases}$$

Where x and y are the coordinates of the person in the left image, X and Y represent the overall width and height of the image, w and h are the coordinates of the person in the right model, and W and H represent the length of the bridge and the corresponding length of the river in the model.

Every two seconds, the digital twin model updates the current water level and displays the positions of people in the river as red dots. Additionally, the digital twin model allows for zooming in and out, panning up and down, and changing the viewing angle, enabling users to closely examine the positions of people in the environment and the condition of the river segment, providing situational awareness.

3.5. Alerting Cards

We define three different levels of alerting cards based on data streams from the previous sub-sections, as shown at the top of Figure 6. A Yellow alert indicates that people have been detected in the frame. In this context, the presence of people in the frame suggests that tourists are active in the area, which implies a potential risk. Although there is no obvious hazardous behavior observed, the detection of people warrants a Yellow alert to indicate a heightened awareness of possible danger.

The Orange alert indicates that people have been detected on islands that are predicted to be submerged. The system aims to predict danger in advance. When individuals are observed lingering on an island rather than just rowing through the river, and Section 3.2 has detected a significant increase in water level (i.e., more than 0.5 feet) in the near future, it becomes highly probable that they will be trapped on the island and unable to escape in time, posing a severe threat to their safety. This situation warrants an Orange alert to prompt immediate attention and action.

A Red alert indicates that people have been detected in the water without a boat. This situation suggests that the individual is no longer on a boat. Regardless of whether the person is drowning or not, this situation is extremely dangerous and urgent. Therefore, this scenario warrants a Red alert to signal immediate danger and the need for prompt intervention.

3.6. System Integration and Dashboard

In the aforementioned sections, each component operates independently in the backend. To integrate all these components into a unified system, we utilize Redis (Wikipedia, 2024b) as an intermediary database. Redis's high read and write speeds are sufficient to support the operation of a real-time system. Each independent component writes its output to the database. For example: Section 3.1 (human detection) writes the smoothed positions of people and boats. Section 3.2 (water level prediction) writes the predicted water levels for the next 15 and 30 minutes. Section 3.3 (submersion detection) writes the comparison information of people's positions relative to the islands, indicating how many people are staying on the islands. Subsequently, Section 3.4 (digital twin) reads the water level information and people's positions from the database. This integration ensures that all parts of the system work together seamlessly to provide real-time monitoring, and situational awareness. Section 3.5 (alerting cards) analyzes the information written by the previous sections and generates the corresponding alerts based on the processed data.

The dashboard is developed using Plotly (Plotly, 2024), a Python library for creating interactive, high-quality graphs and dashboards. Plotly integrates with Dash for easy web application development, supporting various chart types and interactive features. The final result, shown in Figure 6, allows users to drag, move, and zoom in or out of each chart to observe details. Currently, the system runs on a local computer and can be accessed and interacted with through a web

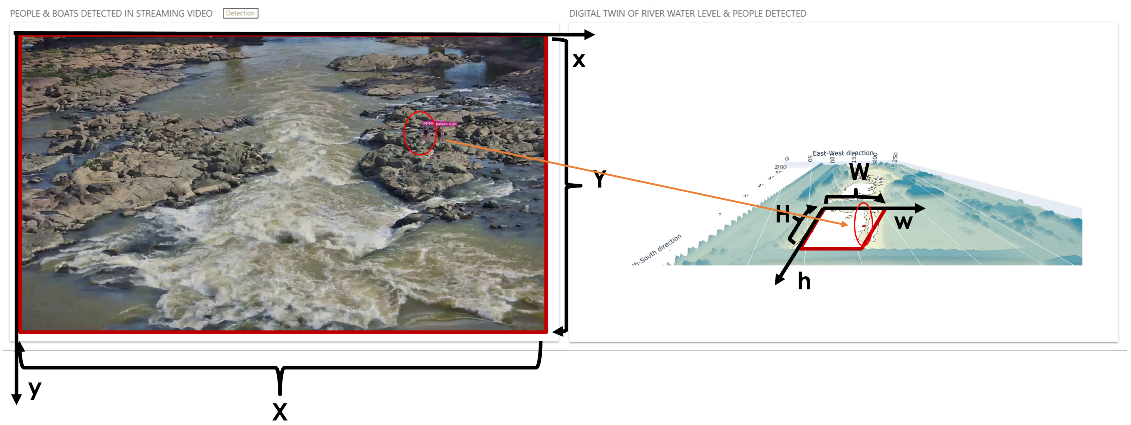

Figure 5. Integration of digital twin with real-time river monitoring and detection. Real-time video feed (left) detecting people and boats, integrated with the digital twin model (right) showing water levels and spatial coordinates (x, y) of detected entities.

browser by entering the corresponding local address. This system can also be ported to other platforms or run on online platforms, offering flexibility and scalability.

4. Validation

Through long-term observation and collaboration with the local government agency, the Columbus Consolidated Government (CCG) and Columbus Fire & EMS (Emergency Medical Services), we simulated a range of possible scenarios and collected empirical data to validate the system's performance. This dataset comprises video footage capturing 7 incidents that should be necessitating Yellow alerts, 4 incidents warranting Orange alerts, and 9 triggering Red alerts. These incidents were used to test the system's alert generation capabilities. As detailed in Table 1, the system demonstrated high accuracy for Yellow and Orange alerts successfully responding to all identified threats in these categories. However, the system only detected 33.3% of the simulated Red alert incidents.

The performance of the system in handling real-time video streams is also evaluated. By separating the resizing and inference processes, although the video stream has a delay of about 5 seconds, the system can recognize objects and perform the risk assessment in each frame at 28fps, which is a higher speed than the original video of 24fps. It means our system can keep up with the real-time operation of the system.

In order to simulate a Red alert scenario, we conducted an isolated scenario experiment involving 9 firefighters standing on an island and sequentially jumping into the water. The objective was to determine whether the system could generate an alert as they were swept away by the stream. This scenario was intended to simulate a potential drowning situation, which is inherently difficult to replicate accurately due to the unpredictable and complex nature of such events. The results revealed two primary factors contributing to the system's performance.

Table 1. Validation results of the system

Alerts		Rate
	Yellow	7/7
False Negative	Orange	4/4
	Red	3/9
False Positive	All included	2/hour

First, the effectiveness of the Red alert relies on the ability to detect individuals within the scope of the river. When subjects entered the water, most of their bodies were immediately submerged, leaving only their heads above the water surface. This limited visibility significantly reduced the detection capability of the YOLO algorithm, which relies on visible body parts to identify individuals. In the three instances where the system successfully generated a Red alert, the subjects were taller and jumped higher, remaining visible for a longer duration before completely submerging. This increased visibility allowed the YOLO algorithm sufficient time to accurately detect them. Apart from this, the chaotic nature of the river and the presence of waves and rapids, also make the trace of people much harder.

Second, the human detection component as detailed in Section 3.1 often misidentifies rocks or trees as boats. Since Red alerts are only triggered when no boats are detected, these false positives hindered the

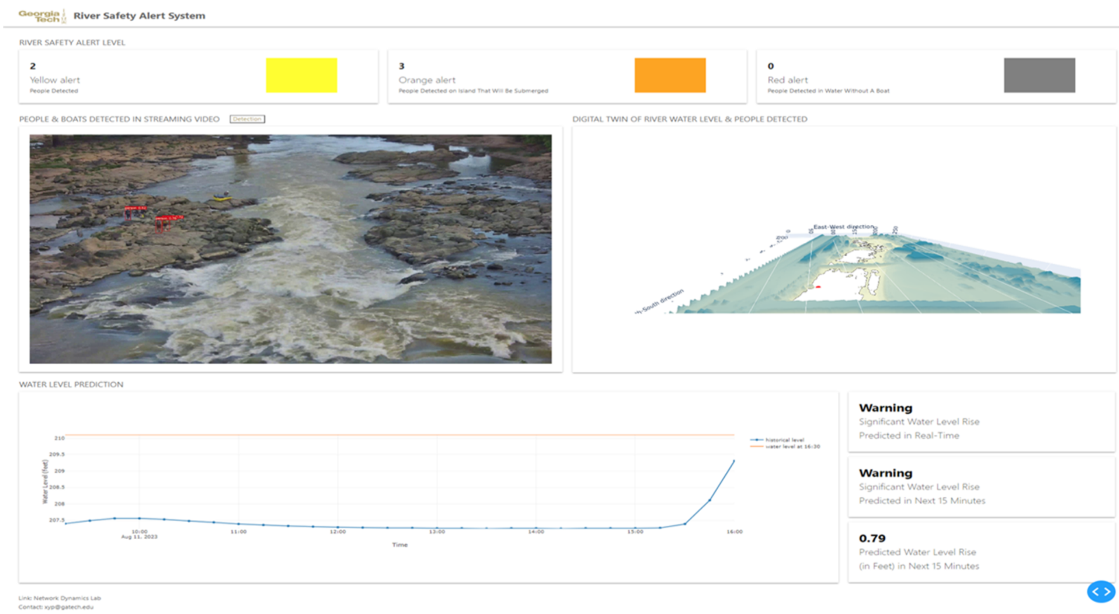

Figure 6. Screenshot of the river safety system dashboard, displaying key components: real-time alerts (yellow, orange, red), live video feed from detected people and boats, digital twin model of the river environment, and water level predictions.

system's ability to generate timely alerts. To mitigate this problem, we added a filter for the data sequence for each read from the database and also fine-tuned the parameters for the YOLO network, reducing the false positive rate for detecting boats to 2 times per hour. This adjustment improved the accuracy of red alerts by minimizing misidentifications.

These observations highlight the challenges of accurately simulating the drowning scenarios and the need for further refinements in the detection algorithms. Future enhancements of our ongoing research could involve integrating thermal imaging or other sensor technologies to improve detection accuracy, especially in scenarios where individuals are partially submerged. In a, hopefully rare, real-world drowning situation, the system's layered alert approach ensures that individuals are monitored from the moment they enter the high-risk area of the river and the YOLO algorithm's performance would ideally benefit from the earlier detection stages (Yellow or Orange alerts) before escalating to a Red alert.

5. Discussion and Conclusions

This paper presents the development of a digital twin-based river emergency response system. The system integrates multiple dimensions of information, including water level values, human detection, and real-time dynamic modeling. By leveraging this integrated approach, the system establishes an alert mechanism for outdoor river environments that not only detects and alarms for ongoing dangers but also provides early warnings for potential threats.

This study faces several limitations, particularly in the detection of drowning incidents indicated by Red alerts. Additionally, the current version of the system relies on several software packages and libraries that are not fully optimized, impacting the system's stability and portability. It currently operates independently and has not yet been integrated with any Emergency Management System or 911 services. In addition, the user interface of this the digital twin has not been designed from a human-computer-interaction perspective. The user-friendliness of the information such as readability, emphasis, and interaction need to be improved.

Future directions can be divided into two main categories. The first is improving model performance. One of the most advanced human detection models currently available at the time of this study is YOLOv10 (A. Wang et al., 2024). YOLOv10 demonstrates superior performance compared to YOLOv7 in terms of COCO Average Precision (AP) across both latency and parameter ranges. This indicates that YOLOv10 is more efficient and scalable, achieving higher accuracy with lower latency and a better utilization of parameters. This makes YOLOv10 a more suitable choice for applications where both speed and model size are critical factors, such as those discussed in this study. The second direction involves collaboration with downstream EMS and other relevant departments. Integrating the system with current practices in place will ensure that effective measures are taken promptly once a danger is detected, thereby maximizing the chances of saving lives. This collaboration will enhance the system's real-world applicability and improve overall emergency response efficiency. Furthermore, to better design the user interface of the digital twin, human-computer interaction principles and techniques need to be applied. For example, user research should be conducted to understand the user's needs, goals, and preferences, and the visual information should be organized more logically, with the most critical data easily accessible.

6. Acknowledgments

This material is based upon work supported by the Georgia Smart Communities Challenge program and the National Science Foundation Grant No. 1837021. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors alone and do not necessarily reflect the views of the National Science Foundation or the Georgia Smart Communities Challenge program. The authors would like also thank the Columbus Consolidated Government (CCG) for their collaboration and support.

References

Batty, M. (2008). Virtual reality in geographic information systems. In J. P. Wilson & A. S. Fotheringham (Eds.), *The handbook of geographic information science* (pp. 317–334). Blackwell Publishing.

Cepeda-Pacheco, J. C., & Domingo, M. C. (2022). Deep learning and 5g and beyond for child drowning prevention in swimming pools. *Sensors*, *22*(19). https://www.mdpi.com/1424-8220/22/19/7684

Chan, Y., Hou, T., Huang, Y., Lan, W., Wang, P., & Lai, C. (2020). Implementation of deep-learning-based edge computing for preventing drowning. *Proceedings of the 8th IIAE International Conference on Industrial Application Engineering*.

ChattVoice. (2024). *Chattahoochee river safety becomes top concern after numerous deaths*. https://chattvoice.com/chattahoochee-river-safety-becomes-top-concern-after-numerous-deaths-p1914-1.htm (accessed: 26.05.2024).

Chaudhari, T., Pandit, G., Gupta, P., & Kumar, M. (2016). Anti drowning system using remote control. *IOSR J. Eng.*, *1*, 38–42.

ClimateCheck. (2024). *Columbus, ga climate risk report*. https://climatecheck.com/georgia/columbus (accessed: 26.05.2024).

Clinic, C. (2024). Blood oxygen level: What it is & how to increase it [Accessed: 2024-06-01]. https://my.clevelandclinic.org/health/diagnostics/22447-blood-oxygen-level

Cristie, V., Berger, M., Bus, P., Kumar, A., & Klein, B. (2015). Cityheat: Visualizing cellular automata-based traffic heat in unity3d. *SIGGRAPH Asia 2015 Visualization in High Performance Computing*, p. 6.

Eng, H., Toh, K., Kam, A., Wang, J., & Yau, W. (2003). An automatic drowning detection surveillance system for challenging outdoor pool environments [[Google Scholar]]. *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)*, 532–539.

Eng, H., Toh, K., Yau, W., & Wang, J. (2008). Dews: A live visual surveillance system for early drowning detection at pool [[Google Scholar]]. *IEEE Transactions on Circuits and Systems for Video Technology*, *18*, 196–210.

He, L., Hou, H., Yan, Z., & Xing, G. (2022). Demo abstract: An underwater sonar-based drowning detection system. *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 493–494. https://doi.org/10.1109/IPSN54338.2022.00047

Jalalifar, S., Belford, A., Erfani, E., Razmjou, A., Abbassi, R., Mohseni-Dargah, M., & Asadnia, M. (2024). Enhancing water safety: Exploring recent technological approaches for drowning detection. *Sensors (Basel, Switzerland)*, *24*(2). https://doi.org/10.3390/s24020331

Jalalifar, S., Kashizadeh, A., Mahmood, I., Belford, A., Drake, N., Razmjou, A., & Asadnia, M. (2022). A smart multi-sensor device to detect distress in swimmers. *Sensors*, *22*, 1059. https://doi.org/10.3390/s22031059

Kałamajska, E., Misiurewicz, J., & Weremczuk, J. (2022). Wearable pulse oximeter for swimming pool safety. *Sensors (Basel, Switzerland)*, *22*(10), 3823. https://doi.org/10.3390/s22103823

Kulkarni, A., Lakhani, K., & Lokhande, S. (2016). A sensor based low-cost drowning detection system for human life safety. *Proceedings of* *the International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*.

Map, T. (2024). Topographic map of columbus [Accessed: 2024-06-01]. https://en-in.topographic-map.com/map-2gddrr/Columbus/?center=40.01868%5C%2C-82.79469%5C&zoom=11

Maps, G. (2024). Map view at coordinates (32.4719635, -84.9962562) [Accessed: 2024-06-01]. https://www.google.com/maps/@32.4719635,-84.9962562,445m/data=!3m1!1e3?entry=ttu

Mohammadi, N., & Taylor, J. E. (2017). Smart city digital twins. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–5. https://doi.org/10.1109/SSCI.2017.8285439

Mohammadi, N., & Taylor, J. E. (2020). Knowledge discovery in smart city digital twins. *Hawaii International Conference on System Sciences*. https://api.semanticscholar.org/CorpusID:211552408

Plotly. (2024). Plotly python graphing library [Accessed: 2024-06-01]. https://plotly.com/python/

Reese, W. (2008). Nginx: The high-performance web server and reverse proxy. *Linux J.*, *2008*(173).

Rooz, E., & Ben-Sira, I. (1991). *Method and system for detecting a motionless body in a pool* (U.S. Patent 5043705A).

Sedor, A. (2017, May). Vive debuts future of vr city planning [URL].

Shehata, A., Mohamed, E., Salem, K., Mohamed, A., Abdul Salam, M., & Gamil, M. (2021). A survey of drowning detection techniques. *Proceedings of the International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*.

Shen, T.-y., Wang, W.-d., Min, H. O. U., Guo, Z.-c., XU, L., & Yang, K.-z. (2007). Study on spatio-temporal system dynamic models of urban growth. *Systems Engineering-Theory & Practice*, *27*(1), 10–17.

Tech, P. (2024). *Poseidon tech official website*. https://poseidon-tech.com (accessed: 29.05.2024).

Verkada. (2024). Try verkada - security cameras and access control [Accessed: 2024-06-01]. https://www.verkada.com

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection [Accessed: 2024-06-01]. *arXiv preprint arXiv:2405.14458*. https://arxiv.org/abs/2405.14458

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2023). Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7464–7475. https://doi.org/10.1109/CVPR52729.2023.00721

Wang, F., Ai, Y., & Zhang, W. (2022). Detection of early dangerous state in deep water of indoor swimming pool based on surveillance video. *Signal, Image and Video Processing*, *16*. https://doi.org/10.1007/s11760-021-01953-y

Wikipedia. (2024a). Long short-term memory [Accessed: 2024-06-01]. https://en.wikipedia.org/wiki/Long_short-term_memory

Wikipedia. (2024b). Redis [Accessed: 2024-06-01]. https://en.wikipedia.org/wiki/Redis

World Health Organization. (2024). *Drowning*. https://www.who.int/news-room/fact-sheets/detail/drowning (accessed: 29.05.2024).