SwimVision: Automated Swimmer Tracking and Activity Classification in Competitive Swimming

Task 2: Business Understanding

Background

Competitive swimming is a highly technical sport where small improvements in speed, technique, and strategy can make the difference between victory and loss. Currently, swimming performance analysis relies heavily on manual observation and timing, which is both labor-intensive and prone to human error. With advancements in computer vision and artificial intelligence, there is an opportunity to automate this process, offering real-time, precise, and actionable insights for athletes, coaches, and event organizers.

Business Goals

- Automate Swimming Performance Analysis: Develop an AI-powered system to track swimmers in real time, classify their activities (start, freestyle, underwater), and measure their speeds.
- Improve Coaching Efficiency: Provide coaches with detailed visual and numerical data on swimmer performance, reducing manual workload and enabling data-driven training.
- Enhance Athlete Performance: Help swimmers identify specific areas for improvement in their technique and race strategy.

Business Success Criteria

- Accuracy: Achieve over 80% accuracy in swimmer detection, tracking, and classification during competitive events.
- Usability: Positive feedback from coaches and event organizers reporting ease of use and meaningful insights from the system.

Assessing My Situation

Inventory of Resources

Human Resources:

- Expertise in computer vision, deep learning, and software engineering.
- Domain knowledge.

Technological Resources:

- Video footage of swimming events for training and testing the system.
- Hardware: GPUs for model training.

- Software: Pre-trained YOLO models, PyTorch framework, and ByteTrack for tracking.

Requirements, Assumptions, and Constraints

Requirements:

- High-resolution video input to ensure detection accuracy.
- A lot of data.

Constraints:

- Limited availability of labeled swimming-specific datasets.
- Real-time processing requirement limits model complexity.
- Complex environment for video analysis (moving cameras, wide pools, constant changing of position of swimmers)

Risks and Contingencies

- Risk: Difficulty in obtaining diverse video data due to privacy concerns.
  - Contingency: Partner with swimming organizations to gain access to anonymized footage.

Terminology

- Bounding Box: A rectangular box used in object detection to define the location of a swimmer in the video frame.
- Homography: A mathematical transformation to map 2D video coordinates onto a flat pool representation.
- Track ID: A unique identifier assigned to each swimmer for tracking across video frames.

Costs and Benefits

Costs:

- Development costs, including computing resources and dataset acquisition.
- Time investment in testing and debugging the system.

Benefits:

- Increased efficiency and accuracy in performance analysis.
- Enhanced training and competition outcomes for swimmers.
- Potential revenue from licensing the system to swimming organizations.

Data-Mining Success Criteria

- Swimmer Detection and Tracking: Achieve at least 80% precision and recall in detecting and tracking swimmers across video frames.

- Activity Classification: Classify swimmer activities with at least 80% accuracy.
- Appoint correctly simmers to the lanes using homography.

Task 3: Data Understanding

1. Gathering Data

Outline Data Requirements

Swimmer Detection Dataset: Images with bounding boxes around swimmers to train and validate the object detection model. Link

- Attributes: Image files annotated with bounding box coordinates and class labels (e.g., swimmer).
- Quantity: At least 4000 images for sufficient model generalization.

Activity Classification Dataset: Images of swimmers performing specific activities (e.g., start, underwater, freestyle) with corresponding labels. Link

- Attributes: Image files annotated with activity classes.
- Quantity: At least 3559 images evenly distributed across activity classes.

Verify Data Availability

The datasets consist of about 8000 images in total, split evenly between swimmer detection (4000 images) and movement classification (3559 images).

Images are labeled with bounding boxes (for detection) and activity classes (for classification), confirming the availability of both required datasets.

All the data has been composed by me.

Define Selection Criteria

Swimmer Detection:

- Include images with swimmers in diverse pool environments (e.g., lighting conditions, pool sizes).
- Ensure bounding boxes accurately represent swimmers' positions.
- Exclude blurred or low-quality images.

Activity Classification:

- Select images with clearly visible swimmers performing distinct activities.
- Balance activity classes to avoid class imbalance in the dataset.
- Exclude ambiguous or mislabeled images.

2. Describing Data

Swimmer Detection Dataset

- Data Size: 4000 images.
- Annotations: Bounding boxes with class labels ("swimmer").
- File Format: JPEG/PNG for images; JSON/YOLO annotation format for labels.

Activity Classification Dataset

- Data Size: 3559 images evenly distributed across activity classes.
- Classes: Start, freestyle, underwater.
- File Format: JPEG/PNG for images; CSV or JSON for class labels.
- Image Dimensions: Normalized to 232x80 pixels for MobileNetV2 input.

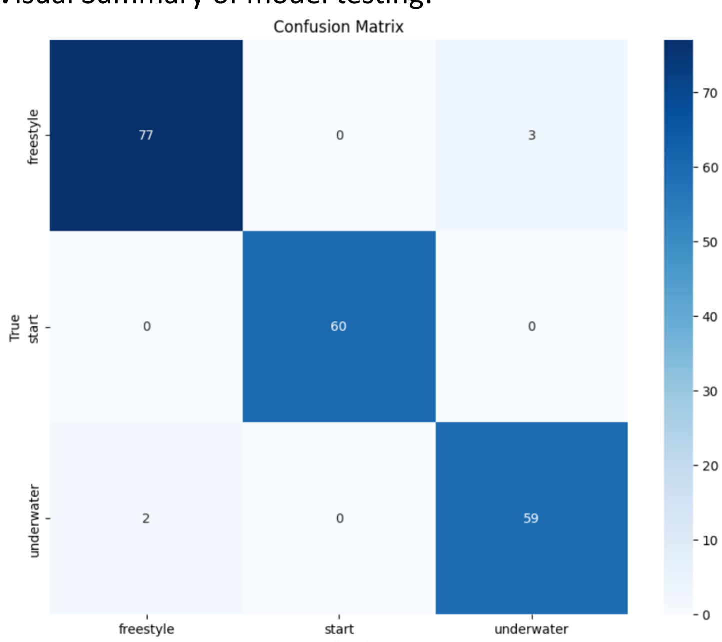
3. Exploring Data

Swimmer Detection Dataset

- Class Distribution: Since only "swimmer" class exists, this dataset is class-uniform.
- Object Size Distribution: Bounding box sizes vary with swimmer proximity and camera angles.

Activity Classification Dataset

- Class Distribution: Dataset includes approximately 1200 images per activity class (start, freestyle, underwater).
- Image Quality: Variations in lighting, focus, and swimmer orientation noted.
- Potential Issues: Some images may contain occlusions or ambiguous movements.
- Visual Summary of model testing:



-

4. Verifying Data Quality

Swimmer Detection Dataset

Completeness: No missing annotations or images; bounding boxes cover all visible swimmers.

Consistency: Annotation format follows YOLO standards.

Activity Classification Dataset

Completeness: All images are labeled with a single activity class.

Note: There is also a swimming pool key point detection model in development, but for this I have to gather a lot of information and data about this as I am still not sure about the format of keypoints. As far as I know there are no publicly available datasets nor any testing made in this direction. Right now, there are multiple testing done. But not anything worth mentioning in comparison to other datasets.

Task 4: Planning Your Project

Step 1: Creating datasets. (80 hours)

Tasks:

- Compose datasets from the scratch.
  - Gather and organize images (10 hours).
  - Annotate images for swimmer detection and activity classification (70 hours).
  - Verify data quality and double-check other things

Step 2: Model developments and Training(Estimated 100 hours)

Tasks:

- As I have trained over 20 versions of both datasets and each training modification and improvement takes a lot of time and effort. It is hard to estimate the total number of hours
- Train YOLO model for swimmer detection (at least 6 hours for one version)
- Train MobileNetV2 for activity classification (at least 2 hours for one version).

Step 3: Creating reusable code for visualization and combining all the components (30 hours)

Tasks:

- Implement code to integrate all models (YOLO, classification, pool mapping) into a unified pipeline (10 hours).
- Test and refine visualization tools for both video and pool representation outputs (5 hours).
- Watch a lot of tutorials regarding coding.

Tools and Methods

Tools:

- YOLO and PyTorch for model training.
- OpenCV for image and video processing.
- Python libraries (e.g., Matplotlib, NumPy) for visualization and analysis.
- Modular programming for reusable code.

Summary

The SwimVision: Automated Swimmer Tracking and Activity Classification in Competitive Swimming project aims to revolutionize performance analysis in competitive swimming using advanced computer vision and AI techniques. By automating swimmer detection, activity classification, and performance tracking, this project seeks to improve coaching efficiency, enhance athlete performance, and streamline event management.

The project followed the CRISP-DM methodology to structure the development process. In the Business Understanding phase, goals such as automating analysis, achieving usability, and delivering actionable insights were established. Success criteria, including achieving 80% accuracy in detection and classification, were defined to guide the project toward practical outcomes.

During the Data Understanding phase, 8,000 high-quality images were composed and annotated to create datasets for swimmer detection and activity classification. The datasets were verified for completeness, consistency, and balance to ensure robust model training.

The Project Planning phase outlined three key steps: (1) composing the datasets, (2) developing and training models, and (3) creating reusable code for visualization and integration. These steps were detailed with tasks, time estimates, and tools used, reflecting the project's complexity and commitment to quality. Over 160 hours were invested in composing datasets and training models, with additional efforts dedicated to developing a modular and reusable codebase for visualization and analysis.

In conclusion, the SwimVision project demonstrates the potential to transform competitive swimming analysis, providing accurate, automated, and actionable insights for athletes, coaches, and organizers. This effort represents a significant contribution to the intersection of sports and technology.