Let Me See Sea: A Vision-Based Smart Surfing Area Analysis System

Syuan Liao, Chih-Hao Yang, Yi-Hsuan Shan, and Tzu-Hsin Wang

Project ID: PRJ-NTPUCSIE-114-010

Project Duration: From September, 2025 to June, 2026

***Abstract*—With the growing popularity of marine recreation, surfing has become a booming sports industry, yet a good surfing experience depends heavily on suitable ocean conditions such as wave height and breaker type. Surfers and surf-related businesses today judge conditions mainly through forecast applications or manual visual inspection of coastal cameras; forecasts give only coarse, region-level predictions, while manual observation lacks quantitative data and is time-consuming, often causing wasted trips or missed waves. This project develops a Vision-Based Smart Surfing Area Analysis System that upgrades video monitoring into a data-driven decision-support tool. An RT-DETR model detects waves in each frame while a BoT-SORT tracker links them across frames to give every wave a consistent ID, and a Support Vector Classification (SVC) texture model separates "spume" from "sunlight reflection" for stability under strong light. We integrate camera attitude with projective geometry for monocular depth and wave-height estimation, and a Decision Tree trained on expert labels classifies each wave as "plunging" or "spilling." A spatial Segmentation Decider then divides the sea area into high/low and plunging/spilling zones, presented through a web application so surfers can quickly decide when and where to enter the water, reducing wasted trips and decision time.**

I. Introduction

SURFING has grown into a globally significant industry: around the turn of the century the worldwide surfing population was estimated at 18–50 million and growing 12%–16% per year, sustaining a multi-billion-dollar market in equipment, apparel, and travel, and its inclusion as an Olympic sport has further expanded public attention [1]. In this context, "surf amenity"—wave height, breaker type, wave frequency, wind direction, and tide—has been recognized as a key factor in the industry [2], since these factors directly affect the safety, challenge, and enjoyment of a session and determine how attractive a spot is to surfers of different skill levels.

However, the monitoring tools surfers rely on today (e.g., Surfline, Windy) are based on buoy data or global weather models and provide only macroscopic information such as average wave height and wind direction over a large sea area; results are often inconsistent across platforms, so surfers frequently find that conditions on arrival do not match expectations. Moreover, wave requirements vary by board type—shortboarders chase steep plunging waves, whereas longboarders prefer gentle, slowly collapsing spilling waves—so on arrival surfers still spend several minutes on manual observation, and a mismatch means wasted travel and time.

From a technical perspective, computer-vision research has focused mainly on rigid land objects such as vehicles and pedestrians; for non-rigid, highly dynamic ocean waves, automated quantification remains underdeveloped, and it is especially hard to separate breaking spume from sea-surface reflections under strong sunlight or to convert 2D image cues into physically meaningful wave heights. This project therefore builds a surf-amenity analysis system that combines computer vision with expert experience: using monocular video and inertial-sensor data, it quantifies and visualizes wave features, translating a professional coach's judgment logic into standardized indicators that help surfers plan their sessions and can also inform local coastal-recreation management.

II. Related Work

A. Wave Peel Tracking

The Wave Peel Tracking (WPT) framework of Thompson *et al.* [2] is a pioneering effort that uses computer vision to track waves one by one and defines indicators such as Ride Length and Ride Rate, showing that surfing data can be quantified from video. Its pipeline, however, leaves two issues we address. First, in the detection stage, WPT uses background subtraction, a 20×20 blur kernel, and blob detection, followed by a CNN classifier on low-resolution 52×52 crops that reports about 12.5% error on an external test set, with failures on spume, strong reflections, and boat wakes; moreover, because it can only crop fixed-size patches, its boxes contain much background water and miss parts of the wave. Second, in the tracking stage, it relies on a complex hand-tuned cost function with many parameters that the authors themselves note require sensitivity analysis. To overcome both, we replace its entire detection, tracking, and classification pipeline with a fine-tuned RT-DETR detector paired with a BoT-SORT tracker and SVC texture classification, and adopt a tighter bounding box that keeps background to a minimum (Section III-B).

B. Object Detection

While object detection has traditionally been dominated by CNN-based models such as the YOLO series, the field has increasingly embraced Transformer-based architectures [3]. Building on this paradigm, RT-DETR (Real-Time DEtection TRansformer) [4] achieves real-time efficiency comparable to the YOLO family while offering superior robustness in complex, dynamic scenes. Crucially, its attention mechanism effectively captures global features. This capability is highly

advantageous for detecting targets with unstable shapes and boundaries distorted by light and reflections. Consequently, we adopt RT-DETR as our per-frame wave detector.

C. Multi-Object Tracking

Because an object detector only localizes targets within a single frame, a separate tracking stage is required to maintain wave identities over time. We adopt BoT-SORT [5], a tracking-by-detection algorithm that associates per-frame detections across consecutive frames and assigns each wave a persistent ID, allowing a wave to be followed through its full life cycle from build-up to collapse.

D. Texture Classification

Distinguishing sunlight reflection, spume, and water surface is critical, yet traditional color-segmentation methods are easily affected by illumination. We instead use Support Vector Classification (SVC), which handles non-linear features with strong noise tolerance, good generalization, and a small memory footprint [6], [7].

E. Monocular Depth Estimation

Under a pinhole camera model and the ground-plane assumption, the distance between the target and the camera can be derived from the target's vertical pixel position given the camera mounting height H_{cam} and pitch angle θ. This is the basis of Inverse Perspective Mapping, which maps image coordinates into world coordinates; we adopt the distance-derivation formula of [8] and additionally use gyroscope (attitude) data to correct for wind-induced view-angle changes, improving height-estimation accuracy.

III. Methodology

We develop an automated system that integrates video processing with physical computation; the overall architecture is shown in Fig. 1. The pipeline comprises video input, object detection and tracking, SVC feature classification, physical height estimation, breaker-type classification, and sea-area surf-condition partitioning.

A. Data Collection and Wave Classification Criteria

Our data come from videos captured on-site at a fixed location with the same recording device. The number of frames a wave persists between the "Start Peeling" and "Collapsed" stages is a key feature of its breaking behavior, and a wave is labeled high or low according to whether its system-estimated height at "Start Peeling" exceeds a threshold. The estimated height and the persistence in frames are then fed to a Decision Tree to determine breaker type. Ground-truth breaker-type labels are provided by a professional surf coach who reviews the videos wave by wave, so that the classification aligns with domain knowledge and gains credibility and practical applicability.

B. Wave Detection and Tracking with RT-DETR and BoT-SORT

This stage stably detects and continuously tracks each wave in a dynamic, noisy ocean environment, providing the basis for later texture analysis, height estimation, and breaker-type classification. The system collects several minutes video and runs a fine-tuned RT-DETR model for per-frame detection, followed by a BoT-SORT tracker for cross-frame association. Compared with traditional CNN detectors, RT-DETR's attention mechanism [3] attends to both local and global features, which matters for targets whose shapes are highly unstable and whose boundaries are blurred by light and reflections.

1) Dataset and annotation: We build a dedicated wave-detection dataset of 592 training, 169 validation, and 84 test images, all manually annotated with wave bounding boxes covering varied wave and illumination conditions. Unlike the fixed-size patches used by Thompson et al. [2], we label each wave with a bounding box that tightly encloses the wave to keep background to a minimum (Fig. 2). The box is then drawn according to the following rules:

1) Start labeling when the wave begins to show spume.
2) Stop including the wave wall once it shrinks to below roughly $1/3$ of the label-box height.
3) Stop labeling when the spume on the left/right of the breaking wave exceeds the remaining wave wall height.
4) Stop labeling once the box would include rocks, or once the wave goes out of frame.
5) When two waves collide:
   a) Spume vs. spume: stop labeling both.
   b) Spume vs. wave wall: decide whether to continue based on whether the wave-wall wave still exists.
   c) Wave wall vs. wave wall: split into two if possible, otherwise label as one.
6) Minimize boxes around spume areas to avoid false detections.

2) Training and tracking: We fine-tune RT-DETR with its native loss design—bounding-box regression (L1 and Generalized IoU) plus a Varifocal classification loss handling the wave/background imbalance—which already suits dynamic targets with fast-changing shapes and indistinct boundaries. At inference we set $\mathrm{conf} = 0.7$ and $\mathrm{iou} = 0.8$ to filter background noise and overlapping boxes. RT-DETR outputs a bounding box per wave in each frame, and a BoT-SORT tracker [5] then assigns each per-frame detection a consistent track ID, and each tracked box is cropped from the frame to remove irrelevant context (sky, beach), yielding clean wave crops as input to the SVC stage.

C. SVC Local Texture Classification and Post-processing

We use SVC to classify local image patches into three classes—background water and others, sunlight reflection, and spume—which is critical for identifying the different stages of wave breaking.

1) Feature engineering and dataset: Instead of classifying individual pixels, we extract a 3×3 RGB neighborhood around each pixel into a 27-dimensional feature vector, sliding

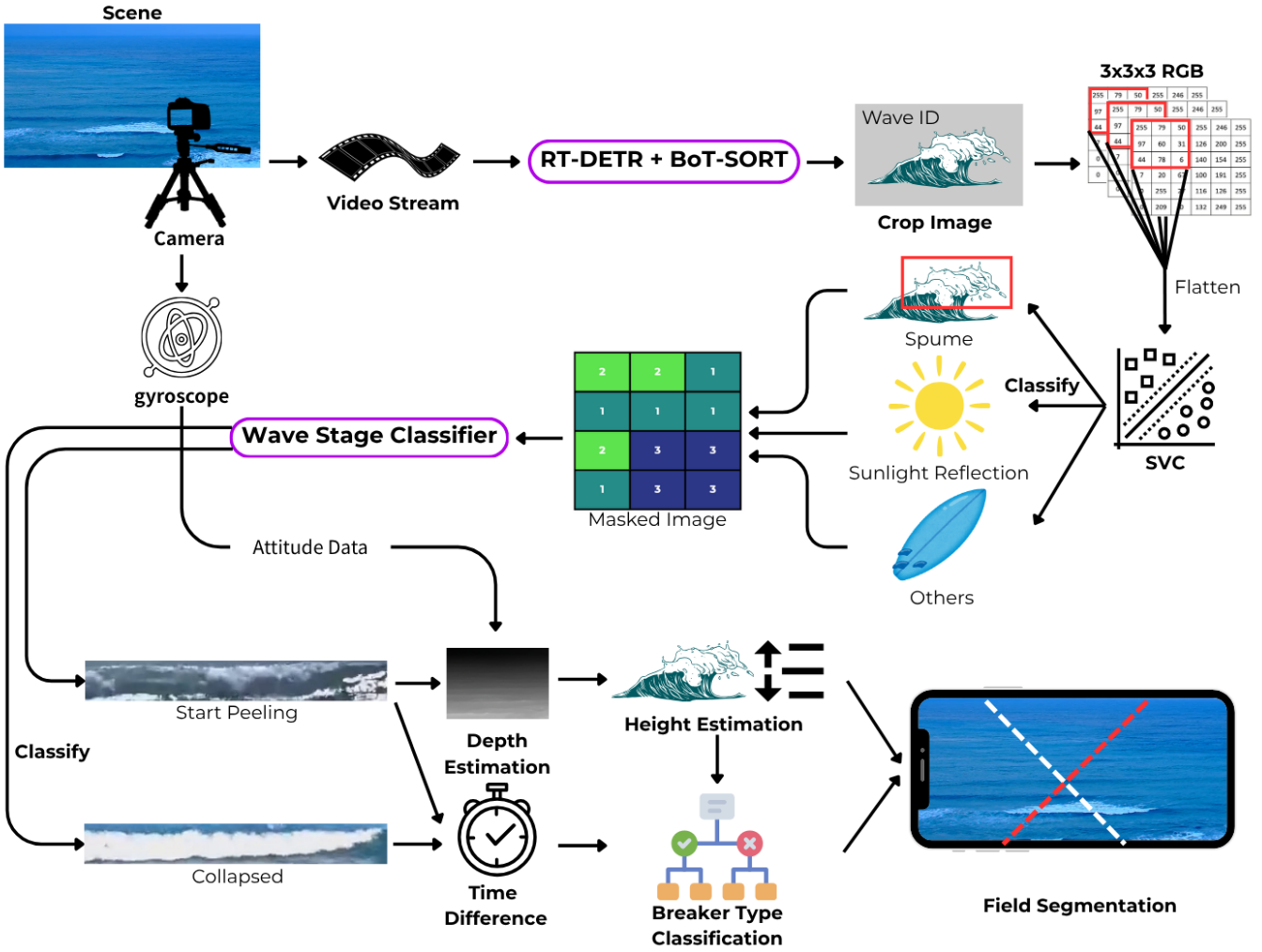


Fig. 1. Overall design of the proposed system.

Fig. 2. Our wave annotation box.

TABLE I
SVC KERNEL FUNCTION COMPARISON.

Kernel	Precision	Recall	F1	Time (s)
Linear	0.83	0.84	0.81	0.1332
Polynomial	0.86	0.86	**0.86**	0.1338
RBF (Gaussian)	**0.87**	**0.87**	**0.86**	**0.1325**
Sigmoid	0.48	0.69	0.57	0.1630

the window with a stride of 2 to reduce computation while preserving the spatial distribution of textures. To ensure stability under different illumination, we manually labeled 17,223 samples from images captured in both the morning and the afternoon, which differ in sun angle and brightness.

2) Model and kernel selection: Because we process a large number of RGB patches, we adopt the GPU-accelerated ThunderSVM library [9]. Benchmarking kernels on 899×54 images (Table I), the Gaussian/RBF kernel achieved the best accuracy and the lowest average prediction time, which suits our near-real-time, high-throughput pipeline at low hardware cost. Running the trained SVC on each crop yields a downsampled prediction mask that preserves the spatial layout and separates strong sunlight reflection from genuine spume, solving a common pitfall of color-based segmentation.

3) Glare-mask post-processing: Because SVC works on local 3×3 patches, reflections may still be misread as spume, so we add a Glare Mask step. The SVC output is reshaped into a 2-D prediction grid; the crop is converted to HSV color space, and the brightness V, saturation S, and local 5×5 grayscale standard deviation are sampled per cell. Since reflections are usually bright, low in saturation, and smoother than breaking spume, such cells become glare-like candidates. Correction is adaptive—broader for obvious reflection crops, with relaxed thresholds for small or dim ones,

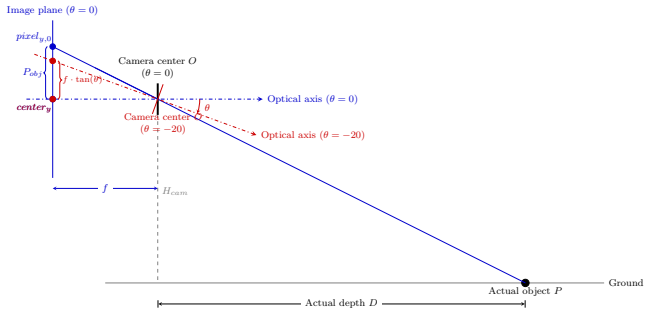


Fig. 3. Geometric model for depth estimation.

and otherwise limited to cells near the dilated SVC sunlight mask to avoid over-correction; connected-component analysis further reclassifies thin, small, bright, low-saturation regions still labeled as spume into reflection. The refined mask is more stable under changing light and gives more reliable input to the downstream wave-stage and breaker-type analysis.

D. Depth and Height Estimation

We use projective geometry to convert 2D pixel coordinates into 3D physical measurements, assuming the camera intrinsics are calibrated and the object–ground contact point can be correctly detected.

1) Depth estimation: From the pinhole imaging principle and the derivation in [8], given the camera height H_{cam}, focal length f, and pitch angle θ, the depth D is

$$D = \frac{f \cdot H_{\text{cam}}}{\text{pixel}_y - \text{center}_y + f \cdot \tan(\theta)}, \tag{1}$$

where pixel_y is the vertical pixel coordinate where the object meets the sea surface and center_y is the image-center row (Fig. 3); gyroscope attitude data are used for correction.

Early experiments optimized this formula with Curve Fit (non-linear least squares), but at long range the tiny pixel change sits in the denominator and makes the far-field error explode. We therefore algebraically linearized the relation, verifying that pixel_y is strongly linear in the inverse depth:

$$\frac{1}{D} = \left(\frac{1}{f \cdot H_{\text{cam}}}\right) \text{pixel}_y + \left(\frac{1}{f \cdot H_{\text{cam}}}\right) \tan\theta - \frac{\text{center}_y}{f \cdot H_{\text{cam}}}, \tag{2}$$

and switched to Linear Regression, which removed the numerical instability caused by a vanishing denominator. A clear data gap appears near 70 m, where pixel variation reaches the camera's optical-resolution limit and causes quantization error; we therefore use a piecewise strategy split at 70 m, comparing Curve Fit, Linear, and Ridge Regression below 70 m and Linear vs. Ridge above it (Section IV-C).

2) Height conversion: With D known, perspective projection and similar-triangle geometry give the object's real-world height from its vertical pixel extent P_{obj}:

$$H_{\text{obj}} = D \times \frac{P_{\text{obj}}}{f}, \tag{3}$$

where H_{obj} is the real wave height (m), D is the regression-estimated distance (m), f is the equivalent focal length (pixels), and P_{obj} is the wave's vertical pixel extent.

E. Breaker-Type Classification and Wave Stage Analysis

We analyze the SVC classification mask by computing the spume ratio in the upper and lower halves of the image, dividing breaking into "Start Peeling" and "Collapsed." When the upper-half spume ratio exceeds 0.05, the system has caught the initial breaking pattern and marks the wave as "Start Peeling." During collapse, spume appears in the lower half and tends to spread sideways, so we divide the lower half vertically into 5 equal-width regions and scan a sliding window of size 3; for window position i, the wave is marked "Collapsed" when

$$\sum_{j=i}^{i+2} W_j \;\geq\; T_{\text{Base}} - \alpha \sum_{j=i}^{i+2} R_j, \tag{4}$$

where W_j and R_j are the spume and sunlight-reflection ratios of region j, $T_{\text{Base}} = 0.95$ is the base threshold, and $\alpha = 0.1$ is a compensation coefficient; the adaptive threshold lowers as reflection interference rises, maintaining stability under strong illumination. We found that breaker type (plunging vs. spilling) correlates with both persistence (the frames between the two stages) and estimated height, and trained a Decision Tree classifier on these two features using 37 professionally labeled samples (29 for training, 8 for testing).

F. Field Segmentation and Visualization

Surfers usually care about the overall spatial distribution of conditions rather than single waves, so we design a Segmentation Decider that maps each wave's center, height, and breaker type into 2D image space and draws representative boundaries from geometric centroids and linear separation. Over a time window T (a few minutes of video) we record each wave's break-start center $C_i = (x, y)$, estimated height H, and breaker type, then split the data into opposing sets by threshold: G_{high} (if $H > 0.8$ m) vs. G_{low}, and G_{plunging} vs. G_{spilling}. For each pair we compute the group centroid

$$\mu = \frac{1}{N} \sum_{i=1}^{N} C_i. \tag{5}$$

Taking the high/low pair as the example, the split line passes through the midpoint $M = (\mu_{\text{high}} + \mu_{\text{low}})/2$ with normal vector $\vec{N} = \mu_{\text{high}} - \mu_{\text{low}}$ (and likewise for μ_{plunging} vs. μ_{spilling}), dividing the frame into two half-planes; if $N_x > 0$ the right half is the high (or plunging) zone, otherwise the left half is.

IV. Results

A. Wave Detection and Tracking

For detection we focus on precision, mAP50, and giou_loss: the precision reaches 82%, the mAP50 reaches 0.85, and the training/validation giou_loss settle around 0.3 and 0.5, respectively. The higher validation localization error is expected because waves are non-rigid objects with no clear boundaries, making the annotation rules somewhat subjective; nonetheless an mAP50 of 0.85 shows the model is capable enough to support the downstream computations. A single frame may contain multiple waves with unstable shapes and reflection

(a) Second 1 (b) Second 2 (c) Second 3

Fig. 4. Wave tracking across consecutive frames: (a) second 1; (b) second 2; (c) second 3.

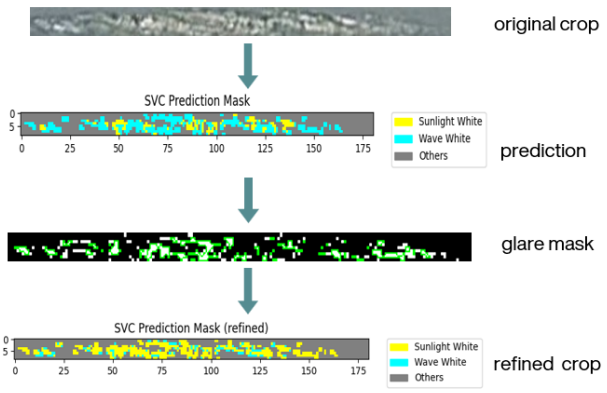
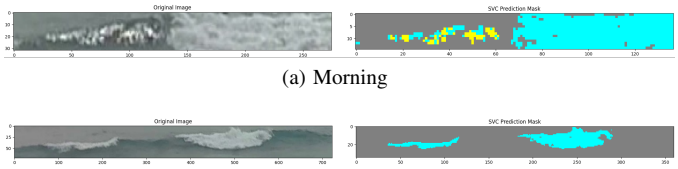


Fig. 5. SVC classification and Glare Mask post-processing flow. The original crop is classified by SVC to generate a raw prediction mask, which a Glare Mask then refines into the final prediction.



(a) Morning

(b) Afternoon

Fig. 6. Original crop vs. SVC prediction mask under different illumination: (a) morning; (b) afternoon. In each panel the left is the original crop and the right is the SVC mask, where cyan marks wave spume, yellow marks sunlight reflection, and gray marks other sea-surface regions.

interference, yet BoT-SORT keeps the same ID for each wave across consecutive frames (Figs. 4a–4c), following each through its movement and life cycle—including appearance and collapse—and providing a stable data source for the height estimation and breaker-type classification.

B. SVC Classification and Post-processing

As shown in Fig. 5, the system first applies SVC to generate a raw prediction mask and then refines it with the Glare Mask post-processing step, so sunlight reflections and true wave spume are separated more clearly. Evaluating the model under different illumination conditions (Fig. 6) confirms that SVC together with Glare-Mask post-processing can effectively distinguish wave spume, sunlight reflection, and other sea-surface regions across varying lighting.

TABLE II
DEPTH-ESTIMATION REGRESSION MODEL COMPARISON.

Range	Method	MAE (m)	MAPE (%)	R^2
< 70 m	Curve Fit	3.693	10.95	0.938
	Linear	1.261	2.90	0.991
	Ridge	**1.231**	**2.86**	**0.992**
70–188 m	Ridge	7.273	5.33	0.878
	Linear	**7.060**	**5.08**	**0.884**

TABLE III
DECISION TREE TRAINING DATA.

	Train	Test
Total Samples	29	8
Plunging	10	5
Spilling	19	3
AUC	0.779 (CV)	0.800

C. Depth and Height Estimation

Using piecewise regression split at 70 m, Table II compares the regression methods across both ranges. Within 70 m, Ridge Regression provides the best stability (MAE 1.231 m, $R^2 = 0.992$), absorbing human-labeling noise better than Curve Fit or Linear Regression. Above 70 m, where pixel variation nears the camera's optical-resolution limit, Linear Regression is slightly better (MAE 7.060 m, MAPE 5.08%), likely because Ridge's L_2 penalty over-constrains a noise-dominated signal. The height conversion is also accurate: MAE is only 0.013 m within 70 m and rises to 0.081 m in the 70–188 m range as depth errors compound, i.e. centimeter-level even at long range, confirming the geometric conversion is reliable for wave-height observation.

D. Breaker-Type Classification

Fig. 7 shows the learned decision boundary separating plunging waves (green) from spilling waves (red) in the feature space spanned by wave height and persistence, revealing that a wave is more likely to be plunging when its persistence falls within a certain range. Under 5-fold cross-validation on the training set the Decision Tree reaches an AUC of 0.779, and it achieves 0.800 on the test set (Table III).

E. Field Segmentation

Integrating detection, tracking, height estimation, and breaker-type classification results within the 2D image coordinate system, the system renders each wave as a marker whose color encodes its height and whose shape encodes its breaker type (Fig. 8). It thus aggregates spatial trends across

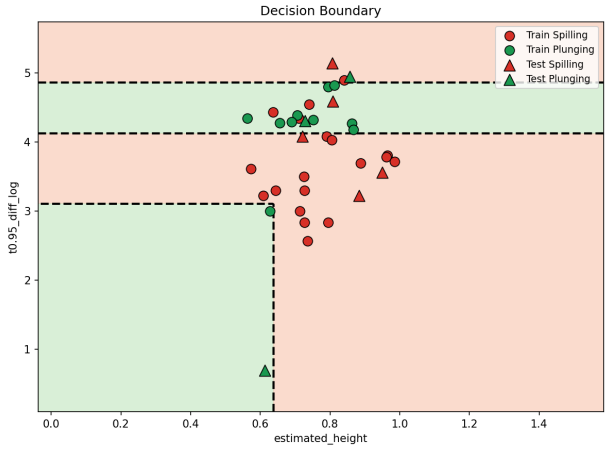


Fig. 7. Decision tree decision boundary.

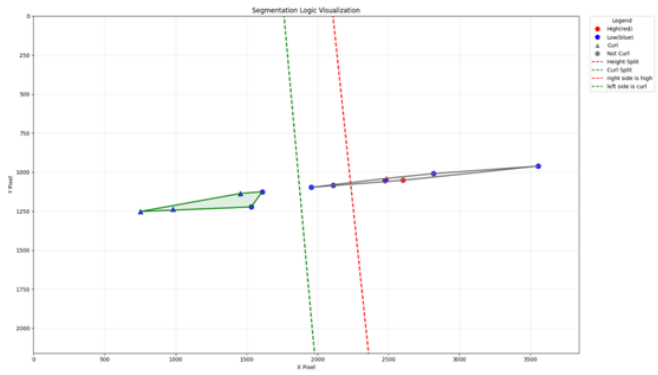


Fig. 8. Height split line (red dashed, right side is higher) and breaker-type split line (green dashed, left side is plunging). Red/blue points indicate high/low waves; triangles/circles indicate plunging/spilling.

multiple waves, turning scattered detections into region-level surf-condition information so users can intuitively identify the locations of high-wave and plunging-wave zones, supporting downstream surf-environment assessment.

F. Web Application

We integrate the back-end wave-analysis results into a web application so surfers can read the surfing-area status at a glance (Fig. 9): the main view overlays the segmentation result on the sea-area footage with an ocean-data panel beneath it, and the app also reports today's tides and a 24-hour forecast of swell period, wave direction, wind speed, and air and water temperature. Building on the high/low and plunging/spilling zones produced by the segmentation stage, its core feature maps each board type to a target combination of these conditions and highlights only the matching surf zone—a shortboard to the plunging-wave zone, a longboard to the low, spilling zone—so a user simply selects a board to see where to surf; users can also view the current high/low and plunging/spilling zones directly. In Fig. 9a, for instance, selecting "Short board" shades its suitable plunging zone over the footage.

V. Conclusion and Future Work

We developed a Vision-Based Smart Surfing Area Analysis System that addresses the lack of fine-grained wave-condition information in traditional, manual-observation-only surfing. On the technical side, we integrate a fine-tuned RT-DETR detector with a BoT-SORT tracker for stable wave detection and continuous tracking in dynamic, complex marine backgrounds. To address strong-light interference at the sea surface, we use SVC local texture classification combined with a Glare Mask post-processing step that separates sunlight reflection from true wave spume, producing a refined mask from which the system detects each wave's breaking stage via an adaptive spume-ratio threshold that compensates for sunlight reflection.

For physical quantification, we built a monocular depth and height estimation model based on projective geometry that handles the camera's resolution bottleneck through range-segmented regression, keeping wave-height errors at the centimeter level even at long range. Combining breaking-stage persistence with estimated height, the system classifies breaker types with a Decision Tree, and a Segmentation Decider aggregates per-wave detections into region-level high/low and plunging/spilling zones. The final output is packaged into a web application that, through intuitive visual boundary lines and ocean data, gives surfers a practical decision-support tool.

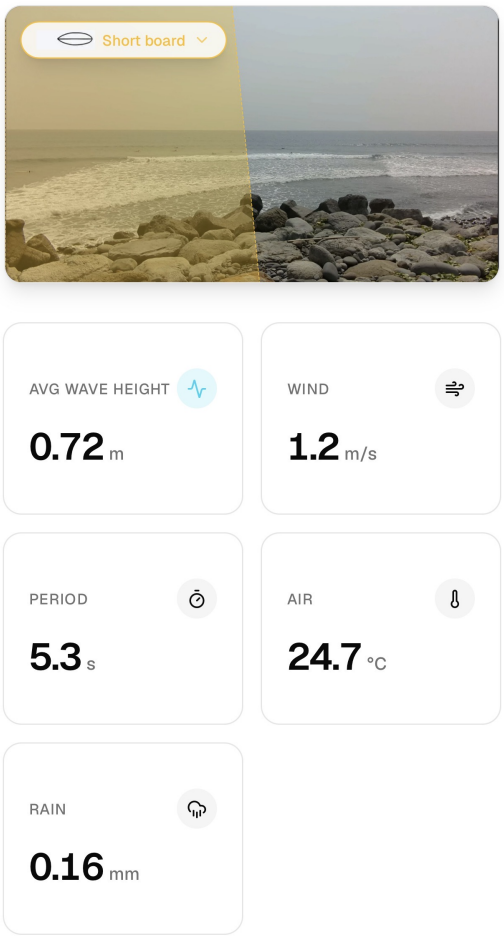
Building on these results, we see two directions for future work. First, scaling data and infrastructure: our data are currently captured only briefly on-site with our own equipment, so we plan to deploy permanent, robust coastal installations and extend the system to more surf spots, yielding continuous, large-scale wave data across regions and seasons. Such data would let us refine the detection and estimation algorithms, improve environmental adaptability, and ultimately evolve the system from reporting current conditions to forecasting how surf conditions will develop, enabling broad deployment across major surf hot-spots worldwide.

Second, commercialization through a Software-as-a-Service (SaaS) model, in which surfers subscribe to decide when and where to surf while local surf clubs, equipment-rental businesses, and tourism operators use the same automated analytics to schedule lessons, advise on rentals, or warn visitors.

Acknowledgment

We would like to express our sincere gratitude to our professor for helping us shape the initial idea of this project and for providing invaluable guidance whenever we encountered difficulties in designing our methods. Our professor's insightful discussions and constructive feedback were instrumental in steering the direction of our system and in refining our report.

We are also grateful to the professional surfing coach for providing the expert wave annotations that served as the ground truth for our breaker-type classification model, and for sharing her insights into the real difficulties of judging surf conditions, which motivated and informed the design of this system. We further thank the surfers who gave us valuable feedback on the actual wave conditions during our on-site recordings. Finally, we thank Pei-Ti Shi and Hsin-Fei Wu for their assistance in conducting the on-campus experiments.



(a) Short-board view with ocean-data panel

24-hour forecast

Hourly sea-state and weather prediction for the monitored shore.

Time	Period	Dir	Wind	Air	Water
00:00	5.6 s		1.8 m/s	21.7°C	28.2°C
01:00	5.6 s		1.5 m/s	21.5°C	28.2°C
02:00	5.6 s		1.2 m/s	21.4°C	28.2°C
03:00	5.7 s		1.2 m/s	20.9°C	28.2°C
04:00	5.7 s		1.2 m/s	20.5°C	28.2°C
05:00	5.7 s		1.3 m/s	20.1°C	28.2°C
06:00	5.8 s		1.3 m/s	20.1°C	28.3°C
07:00	5.8 s		1.4 m/s	20.0°C	28.3°C
08:00	5.8 s		1.5 m/s	20.0°C	28.3°C
09:00	5.8 s		1.6 m/s	21.9°C	28.3°C
10:00	5.8 s		1.6 m/s	23.7°C	28.3°C

(b) Tides and 24-hour forecast

Fig. 9. Web application: (a) main view with "Short board" selected, shading the matching plunging zone over the sea-area footage, with the ocean-data panel beneath; (b) today's tides and the 24-hour ocean forecast.

References

[1] B. E. Scarfe, T. R. Healy, and H. G. Rennie, "Research-based surfing literature for coastal management and the science of surfing—A review," *J. Coast. Res.*, vol. 25, no. 3, pp. 539–557, May 2009, doi: 10.2112/07-0958.1.

[2] M. Thompson, I. Zelich, E. Watterson, and T. E. Baldock, "Wave peel tracking: A new approach for assessing surf amenity and analysis of breaking waves," *Remote Sens.*, vol. 13, no. 17, p. 3372, Sep. 2021, doi: 10.3390/rs13173372.

[3] A. Vaswani *et al.*, "Attention is all you need," Jun. 2017, arXiv:1706.03762. [Online]. Available: https://arxiv.org/abs/1706.03762

[4] Y. Zhao *et al.*, "DETRs beat YOLOs on real-time object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 16965–16974, doi: 10.1109/CVPR52733.2024.01605.

[5] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust associations multi-pedestrian tracking," Jun. 2022, arXiv:2206.14651. [Online]. Available: https://arxiv.org/abs/2206.14651

[6] Scikit-learn developers, "1.4. Support Vector Machines." [Online]. Available: https://scikit-learn.org/stable/modules/svm.html

[7] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm

[8] H. A. Mallot, H. H. Bülthoff, J. J. Little, and S. Bohrer, "Inverse perspective mapping simplifies optical flow computation and obstacle detection," *Biol. Cybern.*, vol. 64, no. 3, pp. 177–185, Jan. 1991, doi: 10.1007/BF00201978.

[9] Z. Wen, J. Shi, Q. Li, B. He, and J. Chen, "ThunderSVM: A fast SVM library on GPUs and CPUs," *J. Mach. Learn. Res.*, vol. 19, no. 21, pp. 1–5, 2018. [Online]. Available: http://jmlr.org/papers/v19/17-740.html