Your Smart Guide to Surf Conditions

SwellSight

ELIYA ZAKAY, SHALEV ATSIS, SAM SOTIL

Project Review

Project Motivation and Core Specifications

- Goal: Provide objective surf metrics from 2D images.
- Core Task: Multi-task computer vision model predicting:
  - Wave height (regression: 0.1-10.0)
  - Wave breaking style (4-class classification: A_FRAME, CLOSEOUT, BEACH_BREAK, POINT_BREAK)
  - Wave direction (3-class classification: LEFT, RIGHT, BOTH)

Project Review

Changes from Proposal

- Originally: Model predicts parameters, then a separate system converts them into a “Surf Score”.
- Now: We do <u>not</u> build the scoring system, our project output is parameters only.
- <u>Deliverable (Model):</u> Image → {Height, Type, Direction}.
- <u>Data generation:</u> parameters → depth map (beach camera) → ControlNet → photorealistic images with automatic labels.

Project Review

Novelty & Contribution:

- Multi-task Architecture: Single model handling regression + dual classification tasks
- Synthetic Data Pipeline: Depth map-guided wave generation using ControlNet + Stable Diffusion.
- Real-Synthetic Integration: Mixed training approach (25% real, 75% synthetic data).
- Domain-Specific Innovation: First application of depth-controlled diffusion models for oceanographic data.

Previous Work

- Paper 1: Adding Conditional Control to Text-to-Image Diffusion Models
- Paper 2: Deep Learning Object Detection Application to Surfing Wave Quality
- Paper 3: AdaBins: Depth Estimation Using Adaptive Bins

Paper / Year	**Task**	**Methods**	**Data**	**Relation to SwellSight**
Zhang et al. (ControlNet, 2023)	Conditional Image Generation	Adding spatial conditioning to Diffusion models	Large Scale (LAION)	Foundation for our Data Gen: We use this to turn Depth Maps into Waves.
Wouters et al. (SurfNet, 2019)	Wave Classification	CNN for classifying surf height ranges	Real images (scraped)	Baseline Task: Similar classification goal, but we add precise regression & synthesis.
Bhat et al. (AdaBins, 2021)	Depth Estimation	Transformer-based depth estimation from 2D	NYU Depth V2	Inverse Problem: They go → Depth; We go Depth → 2D to create training data.

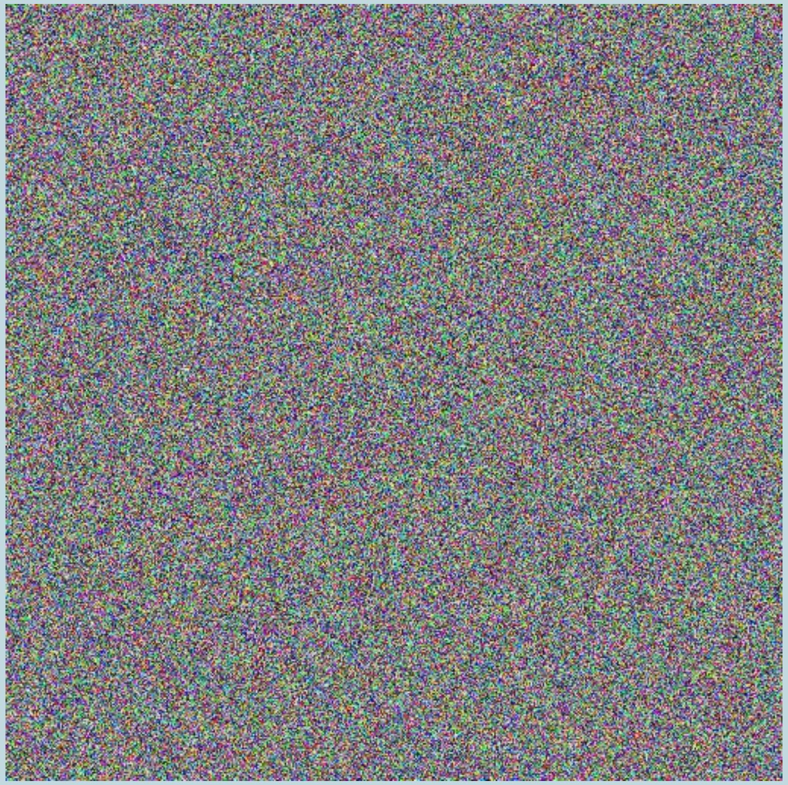
Dataset

Dataset Description:

- Synthetic Data: Automated labeling from depth map parameters
  - Depth maps define exact 3D wave structure (height, shape, angle)
  - ControlNet + Stable Diffusion renders photorealistic textures
  - Ground truth labels extracted with 100% accuracy from generation parameters

- Small real-world dataset used only for domain adaptation and validation.

Dataset



Dataset

Quality Assessment per Task:

- Parametric Geometry: generate a depth map from controlled wave parameters + beach camera setup (perspective, noise, occlusions, run-up).
- Texture Rendering: The Depth Map is fed into Stable Diffusion + ControlNet with prompts (for example- "Sunny day, blue ocean").
- Automatic Labeling: The parameters used in step 1 become the Ground Truth labels (No manual labeling needed).

Dataset

Exploratory Data Analysis (EDA) Summary:

Planned Dataset Composition:

- Target Size: 10,000 synthetic images + 729 real images
- Training Split: 70% train, 15% validation, 15% test
- Real:Synthetic Ratio: 1:14 (6% real, 94% synthetic)

Class Distribution (Planned):

Wave type	Target
A_FRAME	25%
CLOSEOUT	25%
BEACH_BREAK	25%
POINT_BREAK	25%

Direction	Target
LEFT	33%
RIGHT	33%
BOTH	34%

Dataset

Exploratory Data Analysis (EDA) Summary:

Wave Height Distribution:

- <u>Range:</u> 0.5m - 4.0m (log-normal distribution)
- <u>Mean:</u> 2.5 m

Input Specifications:

- <u>Image Size:</u> 512×512 pixels (arbitrary input resolution supported)
- <u>Format:</u> RGB images (JPG/PNG)
- <u>Preprocessing:</u> ImageNet normalization, aspect ratio preservation

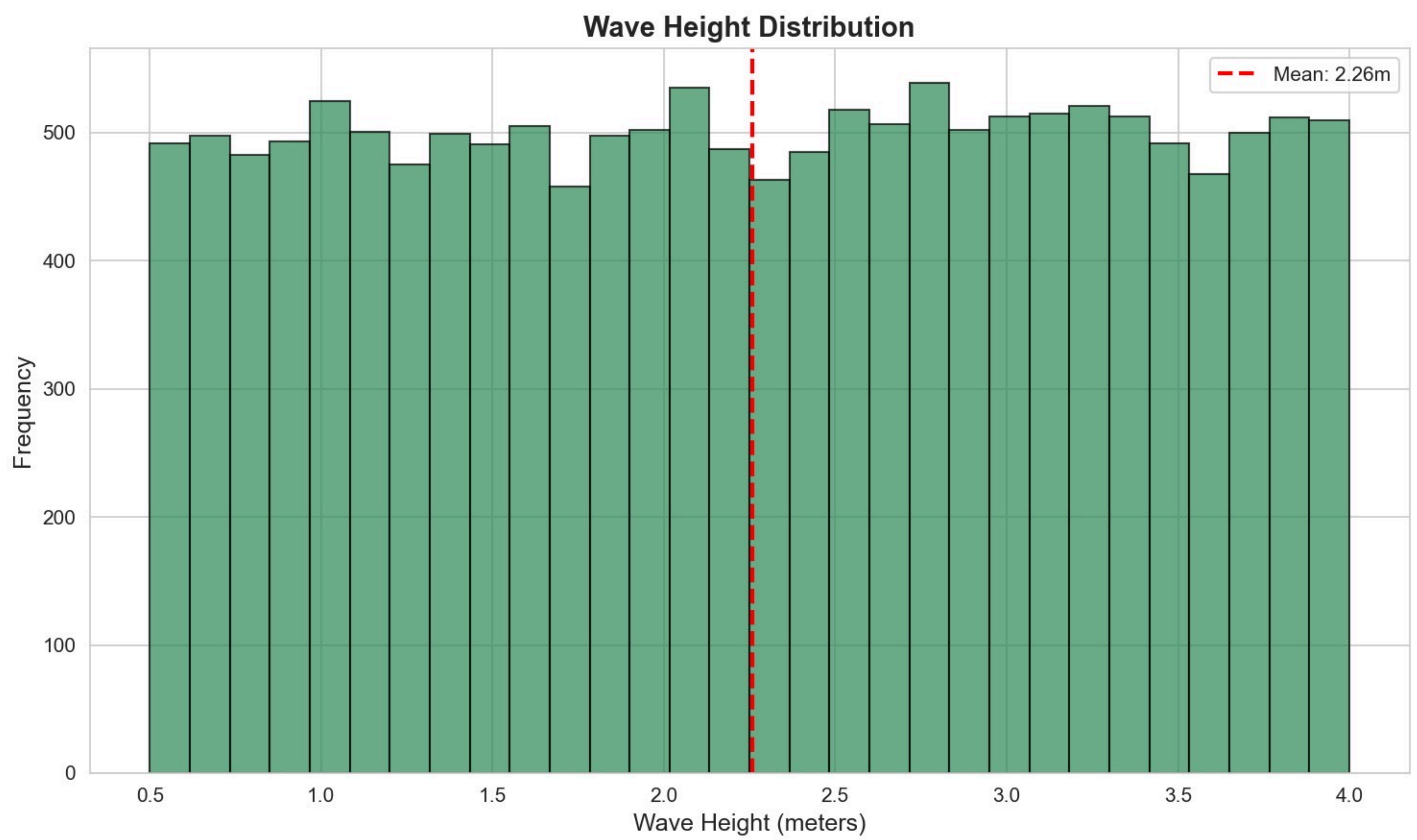
Baseline solution and results

Baseline Model Architecture:

- Backbone: ResNet-50 with ImageNet pre-trained weights.
- Architecture: Shared encoder + task-specific heads
  - Height Head: Linear(2048→512→1) with ReLU + Dropout(0.3)
  - Type Head: Linear(2048→512→4) with ReLU + Dropout(0.3)
  - Direction Head: Linear(2048→256→3) with ReLU + Dropout(0.3)
- Loss Function: Multi-task weighted loss (Smooth L1 + CrossEntropy)
- Training: AdamW optimizer, cosine annealing scheduler

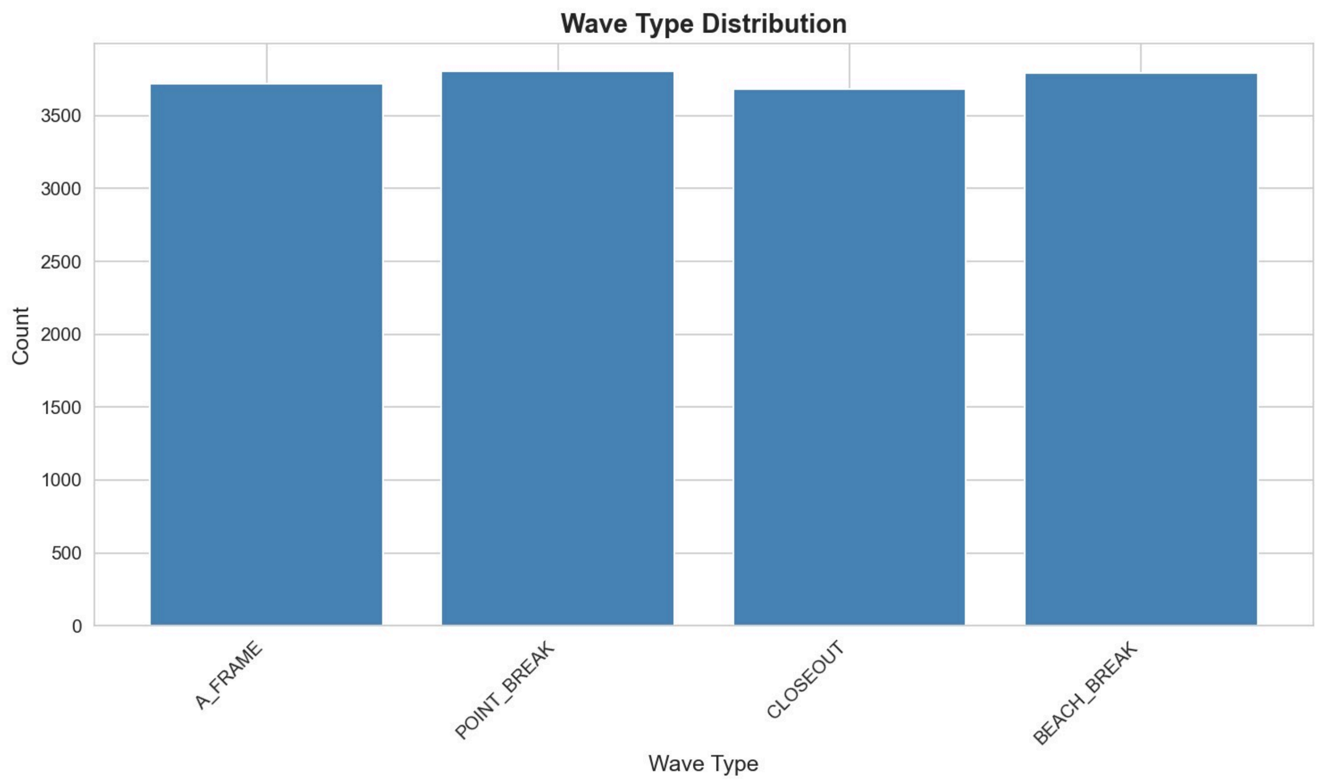
Baseline solution and results

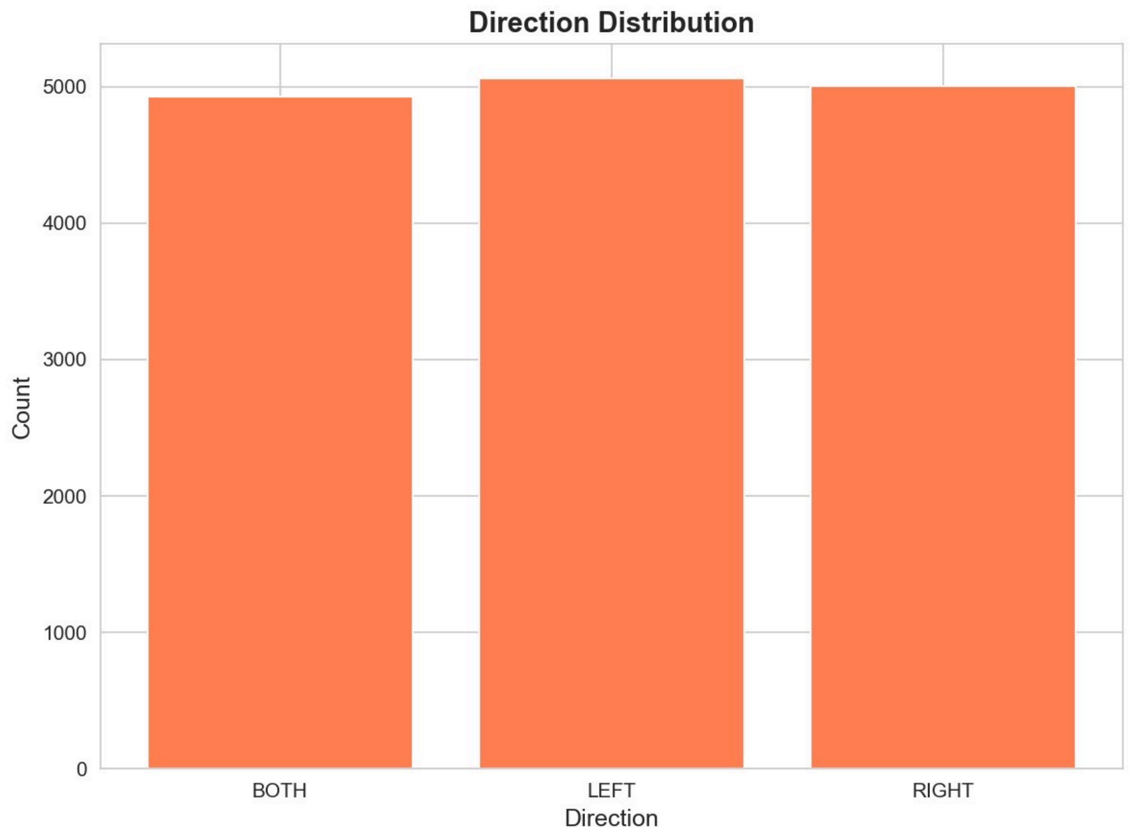
Synthetic data results:



Baseline solution and results

Synthetic data results:





Baseline solution and results

Preliminary Results (Evaluation):

- Synthetic data generation: synthetic beach cam images quality and accuracy are insufficient.

Error Analysis:

- Issue: The model analysis of the real dataset does not function well. it leads to bad synthetic data generation.
- Action item: understanding RCA to solve the issue.

Plan

Step	Scope/Models Used	Expected Outcome	Due Date
Model Tuning	Multi-Task Backbone	Improved F1-Scores	Week 9
Domain Adaptation	Noise/Lighting Augmentation	Better real-world accuracy	Week 10
Final Testing	Real-world Validation Set	Visual consistency report	Week 11
Final Wrap-up	Documentation	Prepare Final Presentation	Week 12

Feedback

- Create depth maps from your real dataset
- Use MiDaS model in order to create the depth maps
- Preform data augmentation from the depth maps
- Data augmentation should be preformed with different attributes
- Use controlNet to create synthetic images from the depth maps

Actual Implementation

- Built a labeled real beach-image dataset and split it into train, validation, and test sets.
- Generated depth maps for the real images using a DPT depth model, then applied depth augmentations to increase variation.
- Used SDXL + ControlNet Depth to synthesize photorealistic wave images from the augmented depth maps, and also generated extra parameter-based depth maps to cover more wave types, directions, and heights.