

EDA of the dataset

DATASET OVERVIEW

The ASL Sign Dataset is designed for ASL recognition from isolated sign videos captures using Mediapipe Holistic Pose estimations

Dataset Size

Train: 75581

Validation Samples: 18896

Total Classes: 250 ASL sign languages

Structure

Dataset is variable length sequence (padded / truncated to 64 length sequences)

75 key points per frame (33 pose landmarks, 21 left hand marks and 21 right hand marks)

Final tensor shape : (Batch size , 64, 75.3) -> 3 is not RGB but its spatial 3D positions

Data Format:

Each row represents a single Landmark joint

frame, row_id, type, landmark_index, x, y, z

Landmark vary by indices (0-32 for pose and 0-457 – for face and 0-20 for hands)

Sample classes

TV, after, airplane, all, alligator, animal, another, any, apple, arm, before, bird, blue, book, boy, brown, but, can, car, cat, chair, cold, cow, cut, dad, dog, down, drink, eat, elephant, fish, flower, go, green, help, home, horse, hot, how, like, mom, more, no, orange, out, play, please, red, run, sorry, stop, thank you, that, the, this, tree, water, what, when, where, white, who, why, yellow, yes...

Many more equivalent till 250

Difficulty Distribution Analysis

The dataset includes difficulty scoring based on three metrics:

Difficulty Metrics

1. Spatial Variance (30% weight)

- Measures how spread out the pose is
- Range: 0.11 - 0.46 (mean: 0.27)
- Higher variance = more complex spatial positioning

2. Temporal Difference (50% weight)

- Measures motion intensity across frames

- Range: 9 - 884 (mean: 207)
- Higher values = more dynamic movement

3. Shape Diversity (20% weight)

- Categorizes motion patterns
- Static (<0.3): 20% difficulty
- Moderate (0.3-0.7): 50% difficulty
- Dynamic (>0.7): 80% difficulty

Metric on sample of 500

Easy (<0.3): 142 samples (28.4%)

Medium (0.3-0.6): 338 samples (67.6%)

Hard (>=0.6): 20 samples (4.0%)

Mean difficulty: 0.378

Std difficulty: 0.114

Range: 0.155 - 0.680