

		Labels
	Adjectives	Alive, dead, dirty, low, soft, beautiful, blind, deaf, sad, ugly, famous, nice, thin, cheap, curved, expensive, female, male, mean
	Places	Restaurant, Office, University, India,
	Greetings	Hello
	Days and Time	Friday, Saturday, Afternoon, Evening
	Electronics	Camera
	Pronouns	you

	DTW Distance Analysis Across ISL Vocabulary Used					
	Avg DTW Distance	Word	Video Count	Category	Rationale	Testing Video
	19.62	Nice	4	Adjectives	High consistency, common in ISL	14
	28.42	thin	4	Adjectives	High consistency, common in ISL	0
	29.89	alive	8	Adjectives	High consistency, common in ISL	0
	45.92	Beautiful	8	Adjectives	Distinct gesture, moderate variability	11
	38.33	Blind	8	Adjectives	High consistency, common in ISL	0
	45.84	cheap	8	Adjectives	Distinct gesture, moderate variability	0
	49.72	curved	8	Adjectives	Distinct gesture, moderate variability	0
	38.7	dead	8	Adjectives	High consistency, common in ISL	0
	45.64	Deaf	8	Adjectives	Distinct gesture, moderate variability	0
	45.01	dirty	8	Adjectives	Distinct gesture, moderate variability	0
	53.44	expensive	8	Adjectives	Distinct gesture, moderate variability	0
	43.9	famous	8	Adjectives	High consistency, common in ISL	0
	29.72	female	8	Adjectives	High consistency, common in ISL	0
	45.03	low	8	Adjectives	Distinct gesture, moderate variability	0
	46.26	male	8	Adjectives	Distinct gesture, moderate variability	0
	37.44	Mean	8	Adjectives	High consistency, common in ISL	0
	51.85	sad	8	Adjectives	Distinct gesture, moderate variability	0
	59.44	Ugly	8	Adjectives	Distinct gesture, moderate variability	0
	63.02	soft	8	Adjectives	Distinct gesture, moderate variability	0
	63.04	rich	8	Adjectives	Describes quality or wealth	0
	65.8	tight	8	Adjectives	Describes fit or closeness	0
	56.06	Afternoon	14	Days and Time	Temporal indicators; regularly signed	0
	60.45	Camera	14	Electronics	Concrete object, easily distinguishable gesture	0
	61.08	Evening	14	Days and Time	Temporal indicators; regularly signed	0
	55.09	Friday	14	Days and Time	Temporal indicators; regularly signed	0
	60.67	Saturday	14	Days and Time	Temporal indicators; regularly signed	0
	66.12	Yesterday	14	Adverbs	Refers to a past time	0
	66.9	clean	14	Adjectives	Describes cleanliness	0
	68.18	Night	14	Nouns	Refers to time of day	0
	72.47	bad	14	Adjectives	Describes something of poor quality	0
	73.11	Thursday	14	Nouns	Day of the week	0
	74.28	Morning	14	Nouns	Time of day	10
	75	Minute	14	Nouns	Unit of time	0
	47.74	India	20	Places	Represents a common noun in ISL vocabulary	0
	57.11	Office	20	Places	Represents a common noun in ISL vocabulary	0
	59.56	Restaurant	20	Places	Represents a common noun in ISL vocabulary	0

	46.82	University	20	Places	Represents a common noun in ISL vocabulary	0
	58.91	Hello	21	Greetings	Frequently used in communication	10
	60.49	you	21	Pronouns	Core grammatical element in ISL	0
	69.64	he	21	Pronouns	Refers to a male person	0
	71.98	it	21	Pronouns	Refers to an object or situation	0
	73.58	Pleased	21	Adjectives	Expresses satisfaction	0
	73.94	Monday	21	Nouns	Day of the week	0
	74.6	Good evening	21	Greetings	Common evening greeting	0
	66.61	Good afternoon	22	Greetings	Part of common time-based greetings	0

	Video Distribution before Augmentation		
	Vocabulary Size	Video Distribution	Total Videos
	6 Words	2×4, 2×8, 2×20	56
	14 Words	2×4, 4×8, 3×14, 5×20	182
	30 Words	2×4, 16×8, 5×14, 4×20, 4×21	514
	45 Words	2×4, 18×8, 12×14, 4×20, 9×21, 1×22	577
	Video Distribution After Augmentation		
	Vocabulary Size	Video Distribution after Augmentation	Total Videos
	6 Words	2×20, 2×20, 2×20	120
	14 Words	2×20, 4×20, 3×20, 5×20	280
	30 Words	2×20, 16×20, 5×20, 4×20, 4×21	624
	45 Words	2×20, 18×20, 12×20, 4×20, 9×21, 1×22	690

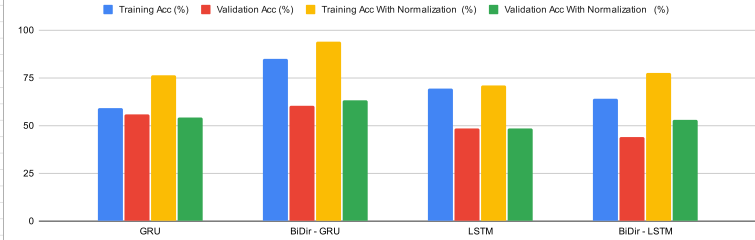
DTW Processing Time Summary per 30 Label Set					Augmentation Type vs Output Duration and Processing Time		
Initial Video Count	Total Labels	Avg DTW Processing Time per Label (mins)		Total Time with DTW	Augmentation Type	Avg Output Duration	Estimated Processing Time
4	2	27.5		55	Slow	2.16 sec	7.32 sec
8	16	9.594375		153.51	Fast	2.06 sec	6.82 sec
14	5	6		30	Interpolation	1.77 sec	49.3 sec
20	4	0		0			
21	2	0		0			
		Total Time with DTW check (mins)		238.51			
Non - DTW Processing Time Summary per 30 Label Set							
Initial Video Count	Total Labels	Avg Non-DTW Time per Label (mins)		Total Time with DTW			
4	2	15.25		30.5			
8	16	9.45		151.2			
14	5	3.25		16.25			
20	4	0		0			
21	2	0		0			
		Total Time w/o DTW check (mins)		197.95			
Comparison of Augmentation Types and Storage Size of Videos With and Without DTW Check for 30 Labels for							
Type of Augmentation	Slow	fast		interp	Size		
Without DTW (%)	33.08%		33.08%	33.84%	4840.63 MB		
With DTW	40.98%		39.85%	19.17%	837.55 MB		
Total Augmented Videos	266						

RNN Performance on Non-Augmented Data (Normalized vs. Without Normalized) – TPU x2 – 30 Labels								
Model	Without Normalization				With Normalization			
	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)
GRU	59.33	55.88	22.69	60.75	76.49	54.41	22.69	59.5
BiDir - GRU	85.07	60.29	51.37	107.875	94.03	63.24	51.37	107.25
LSTM	69.4	48.53	30.2	61.625	70.9	48.53	30.2	61.625
BiDir - LSTM	64.18	44.12	68.39	102	77.61	52.94	68.39	113.875

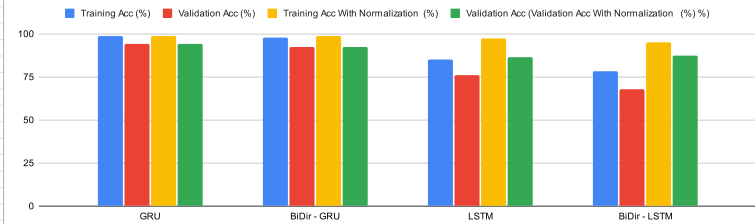
RNN Performance on Augmented Data w/o DTW (Normalized vs. Without Normalized) – TPU x2 – 30 Labels								
Model	Without Normalization				With Normalization			
	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)
GRU	98.96	94.21	22.69	101.375	98.75	94.21	22.69	101.25
BiDir - GRU	97.92	92.56	51.37	133	98.54	92.56	51.37	169
LSTM	85.03	76.03	30.2	80.625	97.3	86.68	30.2	112.5
BiDir - LSTM	78.36	67.77	68.39	154	95.22	87.6	68.39	203.625

RNN Performance on Augmented Data w/ DTW (Normalized vs. Without Normalized) – TPU x2 – 30 Labels								
Model	Without Normalization				With Normalization			
	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)	Training Acc (%)	Validation Acc (%)	Param Size (MB)	Avg Training Time (s)
GRU	98.96	90.08	22.69	90.875	99.79	89.26	22.69	100.125
BiDir - GRU	95.84	89.26	51.37	137.25	99.58	89.26	51.37	153.375
LSTM	77.15	72.73	30.2	86	96.88	86.78	30.2	103.75
BiDir - LSTM	83.37	66.94	68.39	130.5	91.04	80.17	68.39	163.75

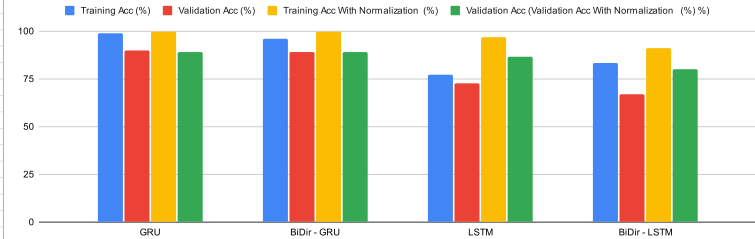
RNN Performance on Non-Augmented Data (Normalized vs. Without Normalized) – TPU x2 – 30 Labels



RNN Performance on Augmented Data w/o DTW (Normalized vs. Without Normalized) – TPU x2 – 30 Labels



RNN Performance on Augmented Data w/ DTW (Normalized vs. Without Normalized) – TPU x2 – 30 Labels



	Scalability of GRU Model with Normalization and without DTW							
	Label Count	Train Acc (%)	Val Acc (%)	F1-Score	Precision	Recall	AUC-ROC	Training Time (s)
	6	100	96	100	100	100	100	29
	14	89.29	90.18	89.17	90.71	89.29	99.55	52
	30	96.05	92.56	85	87.2	86	99.76	98.75
	45	99.86	91.76	88.99	91.85	89	97.74	145.75
	Scalability of GRU Model with Normalization and with DTW							
	Label Count	Train Acc (%)	Val Acc (%)	F1-Score	Precision	Recall	AUC-ROC	Training Time (s)
	6	100	100	100	100	100	100	30.125
	14	91.07	85.71	86.17	88.57	85.71	98.87	52.125
	30	98.96	90.91	90.24	93.49	90.83	99.87	98
	45	98.62	90.11	89.3	91.07	89.56	99.49	141.5

	Type of Normalization	Label Count	Train Acc (%)	Val Acc (%)	F1-Score	Precision	Recall	AUC-ROC
	Normalization to Chest with Bounding box	45	99.45	87.91	88.02	89.78	88.86	99.66
	Normalizing both position and scale: after bounding box translation, divide landmark coordinates by bounding box size (width, height, depth).	45	98.9	92.31	92.08	94.11	92.22	99.85
	Procrustes alignment	45	97.11	85.61	85.45	88.83	86.44	98.67