

Tsubame2.5:

- CPU: Xeon X5670
- GPU: Tesla K20X

	Feature		LSTM Training	LSTM Testing	Total
	Frame Extraction	Feature Extraction			
#CPU	1	1	1	3	3
CPU Mem	920 MB	1.6 GB	2.3 GB	7.5 GB	7.5GB
#GPU	N/A	1	1	3	3
GPU Mem (batchSize)	N/A	1.1 GB (1)	4.25 GB (5)	4.78 GB (20)	4.78 GB
Time	50 hrs	150 hrs	20 hrs	2.33 mins	220.04 hrs
Storage	1.2TB	32 GB	407 MB	1.1 MB	1.23 TB

Tsubame3.0:

- CPU: Xeon E5-2637
- GPU: Tesla P100

Improvement:

- Speed
 - Feature: 10.53X
 - LSTM Training: 28.12X
 - LSTM Testing: 14.68X
 - Total: **11.16X**
- Storage
 - Feature: 39.4X
 - Total: **38.87X**

	Feature	LSTM Training	LSTM Testing	Total
	Direct Feature Extraction			
#CPU	4	1	4	4
CPU Mem	5.6 GB	3.22 GB	18.44 GB	18.44 GB
#GPU	4	1	4	4
GPU Mem (batchSize)	1.1 GB (1)	9.19 GB (128)	9.68 GB (512)	9.68 GB
Time	19 hrs	42.67 mins	9.54 secs	19.71 hrs
Storage	32 GB	407 MB	1.1 MB	32.40 GB

Analysis of improvement

- Feature
 - Speed: Faster GPU speed + more GPU parallelism (10.53X)
 - Storage: Combine frame extraction and feature extraction (39.4X)
- LSTM training speed:
 - Fewer epochs required to converge due to adopting larger batch (100 epochs -> 40 epochs, 2.5X)
 - Less time required to run one epoch due to
 - Faster GPU speed and better parallelism for larger batch (9.8X)
 - Code optimization including applying asynchronous data copying and avoiding unnecessary data copying (1.15X)
- LSTM testing speed:
 - Faster GPU speed + more GPU parallelism (12.18X)
 - Code optimization including applying asynchronous data copying and avoiding unnecessary data copying (1.20X)