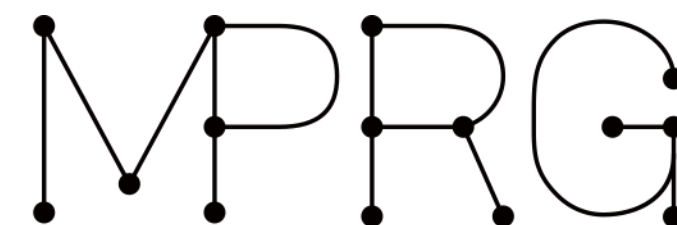


第20回ディスカッション

実験

ER20038 小林亮太

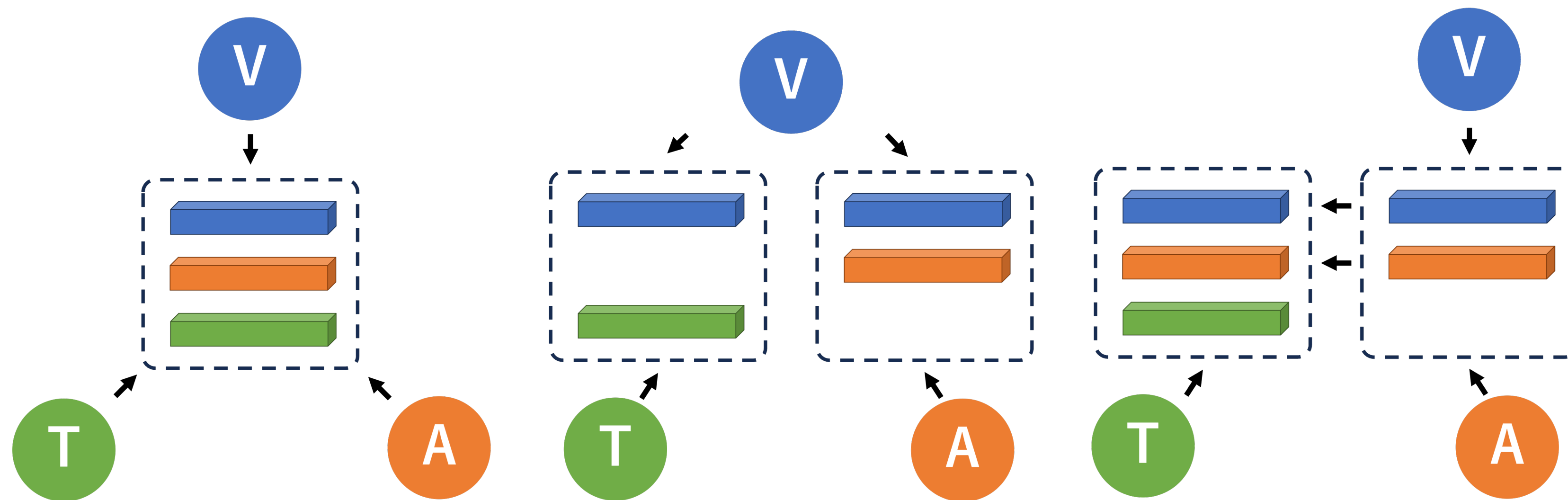
担当：鈴木雅★， 福井， 張



MACHINE PERCEPTION AND ROBOTICS GROUP

- 研究テーマ
- 実験状況
- 実験結果

- 3モーダル（ビデオ，オーディオ，テキスト）のマルチモーダル自己教師あり学習
- テキストに比べビデオやオーディオにはノイズが多く存在
 - 各モーダルの組み合わせでノイズを抽出せずに学習ができる可能性
 - 近づけるモーダルの組み合わせによる学習効果への影響について調査



- デフォルトの条件のまま実行
 - Num_workerの値のみ変更
- CUDAのエラーは解決
 - RuntimeError: CUDA error: CUBLAS_STATUS_EXECUTION_FAILED
 - 原因
 - Vfs_fuse.c:281 UCX ERROR inotify_add_watch(/tmp) failed: No space left on deviceが原因？
 - サーバの再起動で解消
- 3モデル使用の学習は完了
- 現在は2モデルのみの学習を実行中
 - 1epochごとの所要時間に変化なし

- 3モデルを使用した学習の結果
- 評価タスク：テキストからビデオの検索 on YouCook2
 - R@k
 - R : Recall（再現率）
 - K : 各クエリの上位k個の予測のうちの正解数の総正解数に対する割合

	R@1	R@5	R@10
論文	18.1	35.5	45.2
実験	14.7	34.1	44.7
著者配布の重み	17.3	32.9	41.6

- 実験 : 実行中
- 今後の予定 :
 - 中中間の資料作成
 - 卒論の作成

Multimodal Clustering Network (MCN) [B. Chen+, ICCV'21]

- ラベル付けされていないナレーション付きビデオから学習
 - テキストからビデオの検索, 時系列行動検出が可能
- テキスト, オーディオ, ビデオの3つのモーダルを使用

