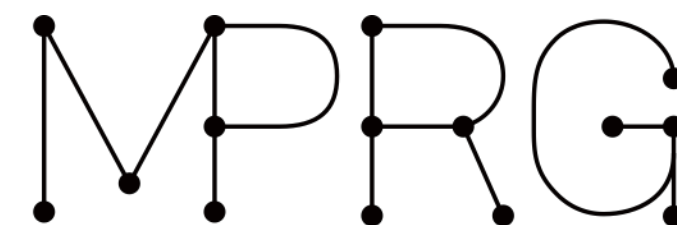


第15回ディスカッション

実験状況

ER20038 小林亮太

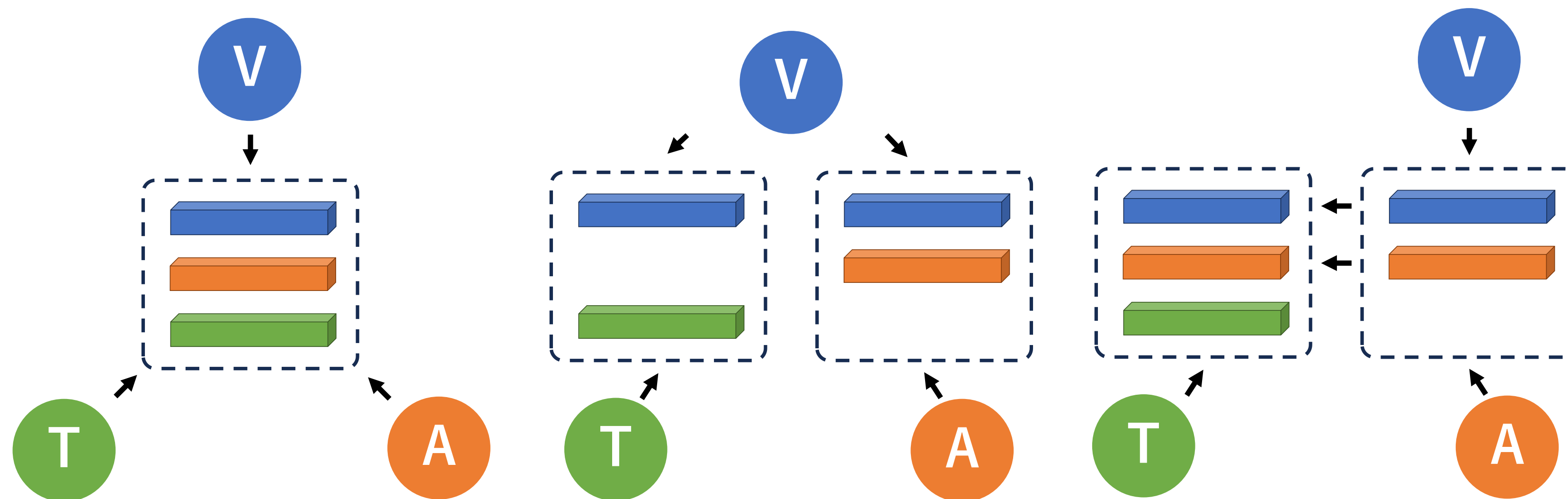
担当：鈴木雅★， 福井， 張



MACHINE PERCEPTION AND ROBOTICS GROUP

- 研究テーマ
- 実験状況

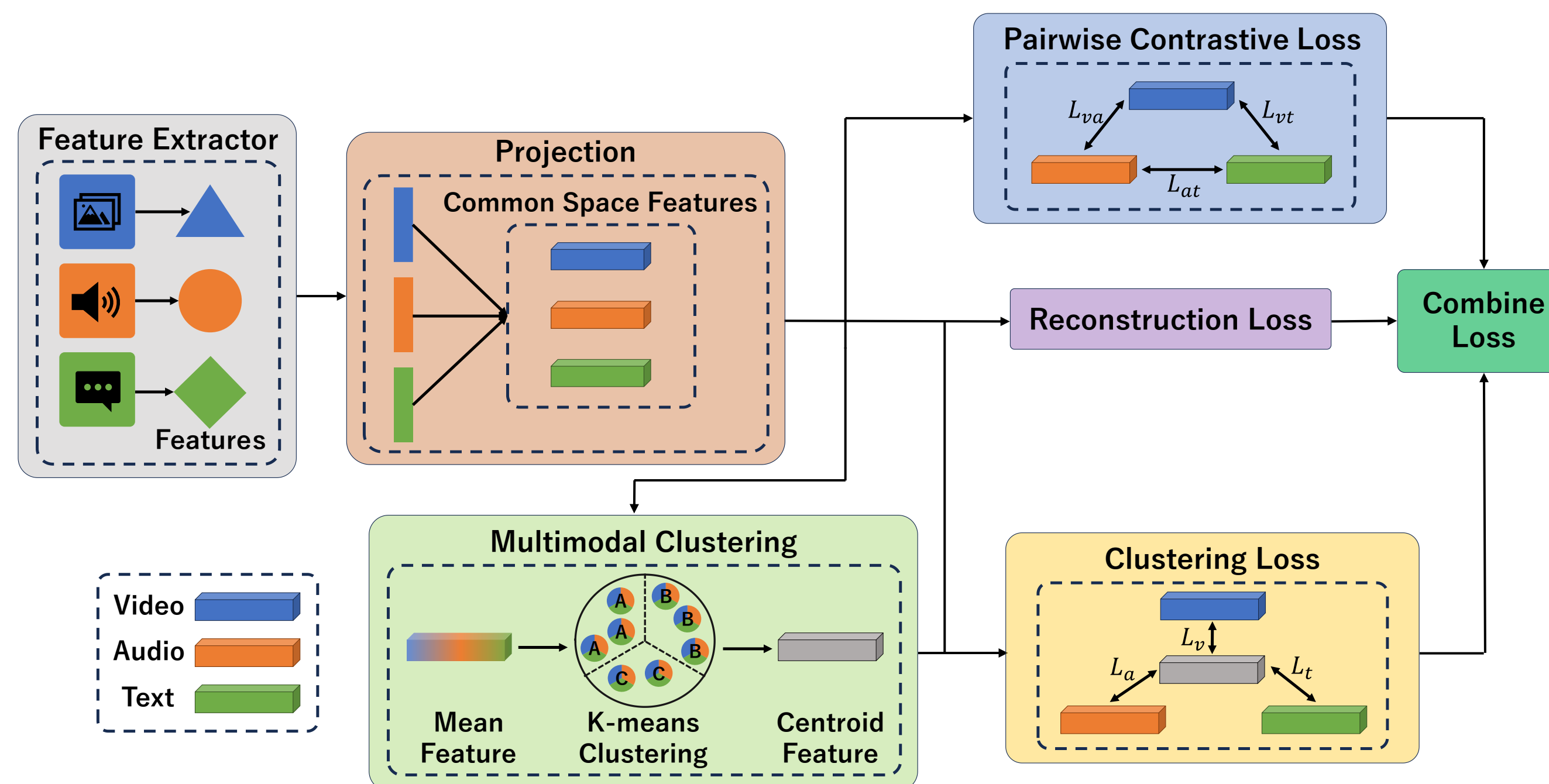
- 3モーダル（ビデオ，オーディオ，テキスト）のマルチモーダル自己教師あり学習
- テキストに比べビデオやオーディオにはノイズが多く存在
 - 各モーダルの組み合わせでノイズを抽出せずに学習ができる可能性
 - 近づけるモーダルの組み合わせによる学習効果への影響について調査



- 評価タスク：テキストからビデオの検索
 - 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
- 論文と比較して著しく低下

	R@1	R@5	R@10
論文	10.5	25.2	33.8
今回の実験	1.70	7.00	11.6

- Lossが減少していないことを確認
 - Contrastive Lossが原因
 - 他の2つは減少していることを確認
- 原因を調査中



- 実験 : 調査中
- 今後の予定 :
 - 卒論のための情報まとめ
 - プログラムの作成

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

