

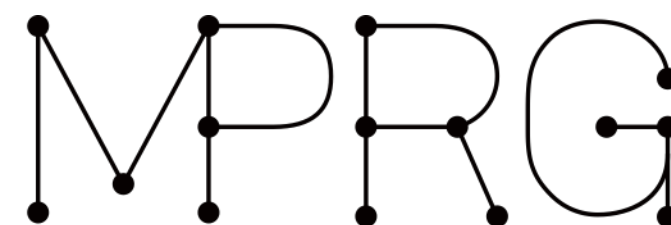
第19回ディスカッション

# バッチサイズ

---

ER20038 小林亮太

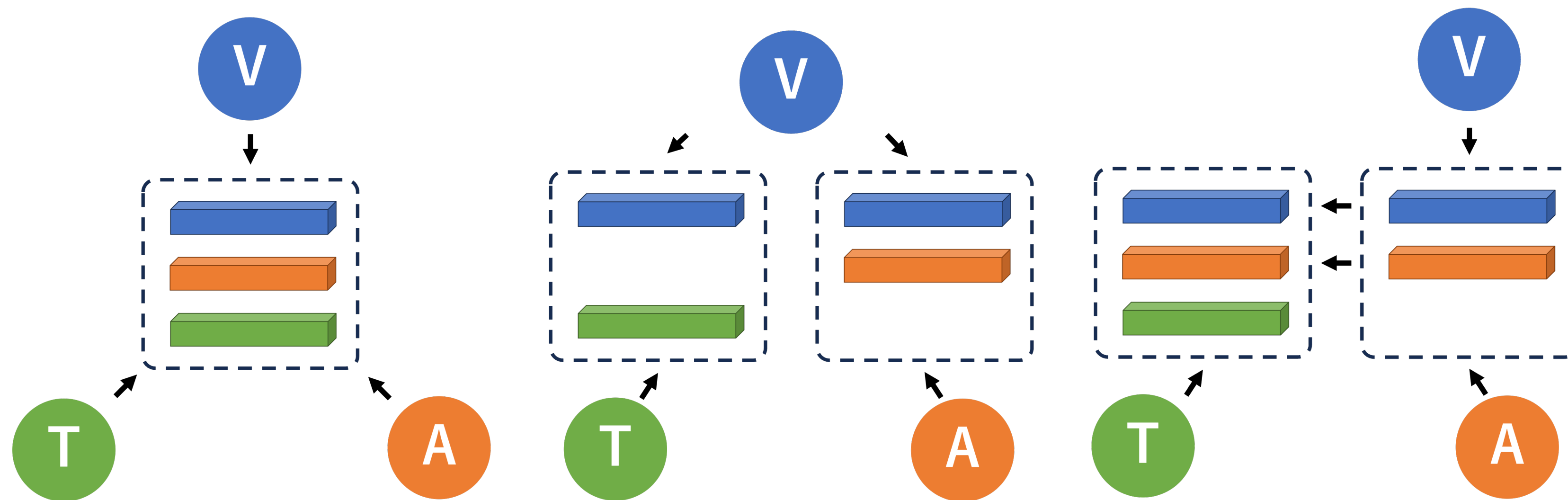
担当：鈴木雅★， 福井， 張



MACHINE PERCEPTION AND ROBOTICS GROUP

- 研究テーマ
- 実験条件について
- 条件batch\_size=128, n\_pair=32の場合
- 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




- 実験条件を論文と揃える目的
  - 論文内の表記を参照しバッチサイズを4096に設定
    - メモリ不足エラーが出現
    - バッチサイズ2048, 1024, 512, 256でもエラーが出現
    - 解決が難航
  - 論文内の表記は直接的なバッチサイズではない値
    - $\text{Batch\_size} \times \text{n\_pair}$ の値
      - Batch\_size : 動画の数
      - N\_pair : 動画1本あたりのclip数
      - それぞれ初期値はbatch\_size=128, n\_pair=32
      - 2つの積は4096

# 条件batch\_size=128, n\_pair=32の場合

- A 100 × 4で0番の使用量が13~15GB
  - 著者の環境のV 100が16GBモデルだとしても動作可能
  - 情報が合致
- 学習時間
  - 著者環境で2日程度
    - CPU : 80スレッド
    - Num\_worker : 74
  - 自分の環境では1週間程度
    - CPU : 32スレッド
    - Num\_worker : 4
      - 大きくしすぎるとRAMが不足
  - CPU性能およびRAM搭載量の差やHDDの速度が原因の可能性

- デフォルトの条件のまま実行
  - Num\_workerの値のみ変更
- CoreDumpで停止→解決
  - Docker runのオプションipc hostが原因
  - Shm-sizeオプションのコンテナでは動作
- CUDAのエラーで停止
  - RuntimeError: CUDA error: CUBLAS\_STATUS\_EXECUTION\_FAILED
  - いつものOutOfMemoryとは別物
    - CUDA error: CUBLAS\_STATUS\_EXECUTION\_FAILED
  - 原因を調査中
    - Vfs\_fuse.c:281 UCX ERROR inotify\_add\_watch(/tmp) failed: No space left on deviceが原因？
    - Epoch11時点での停止のため再現に時間が必要
      - 原因の迅速な特定が困難

### Download



Data

P

↓ README

↓ S3D features

↓ All-in-One zip

Bibtex:

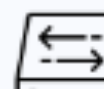
```
@inproceedings{miech2019howto100m,  
  title={HowTo100M: Learning  
  author={Miech, A., et al.,  
  booktitle={ICCV},  
  year={2019},  
}
```

If you wish to download the video files from our private server, please fill this form: [here](#). We are hosting the videos rescaled so that  $\min(\text{height}, \text{width}) = 256$  and we have removed the audio.

We note that the distribution of identities and activities in the HowTo100M dataset is not representative of the general population. Please be careful of unintended societal, gender, racial

以前

### Download



Data

↓ README

↓ S3D features

↓ All-in-One zip

Bibtex:

```
@inproceedings{miech2019howto100m,  
  title={HowTo100M: Learning  
  author={Miech, A., et al.,  
  booktitle={ICCV},  
  year={2019},  
}
```

We note that the distribution of identities and activities in the HowTo100M dataset is not representative of the general population. Please be careful of unintended societal, gender, racial

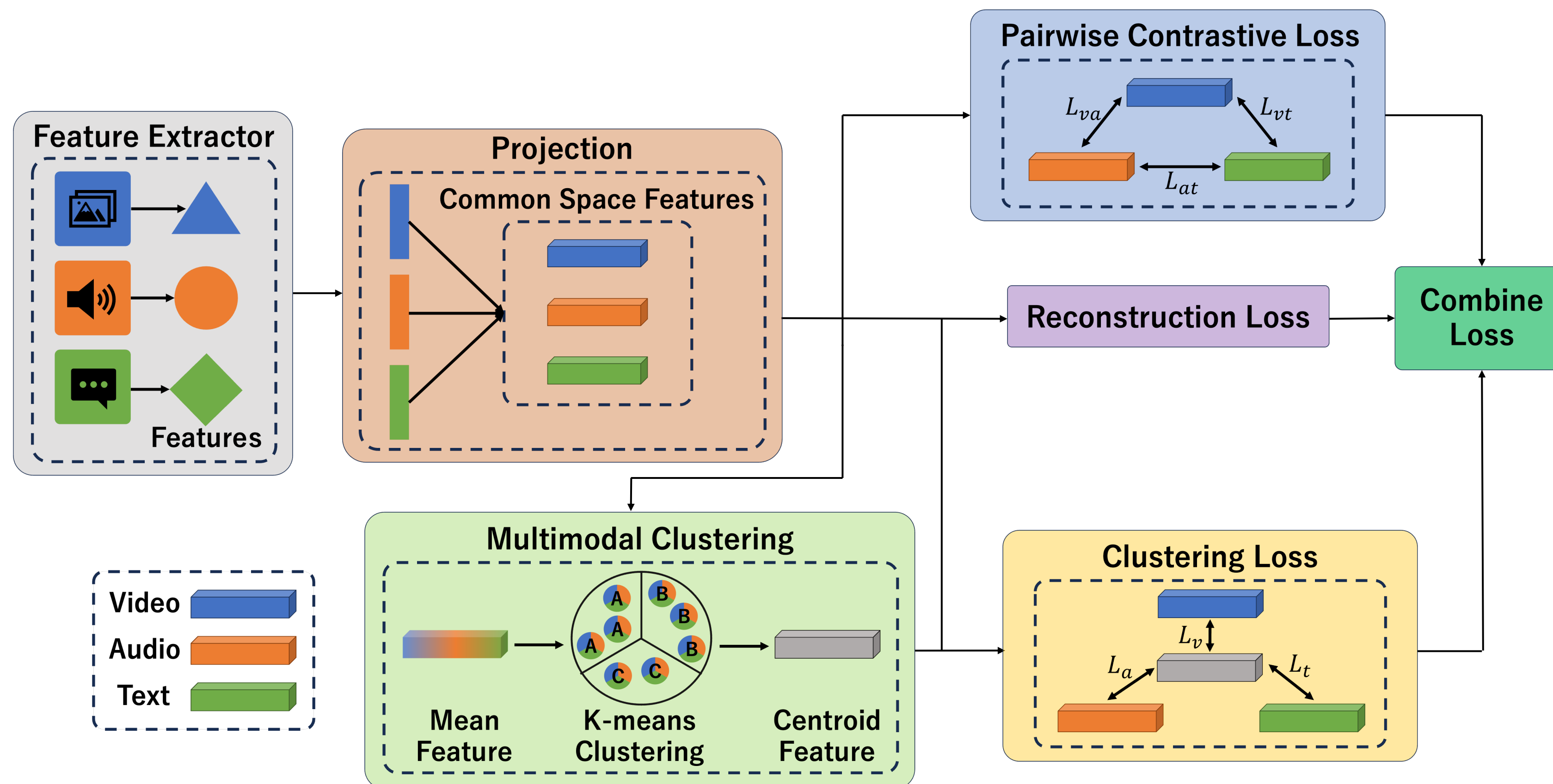
現在

- 実験 : エラー調査中
- 今後の予定 :
  - 中中間の資料作成
  - 他の環境での実験
    - 7TBのデータの置き場がなく断念



# 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



- 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
- 平均を取る処理の追加で精度向上
- バッチサイズを論文の条件と揃えることが必要

	R@1	R@5	R@10
論文	18.1	35.5	45.2
前々回の実験	1.70	7.00	11.6
前回の実験	5.29	13.2	18.4