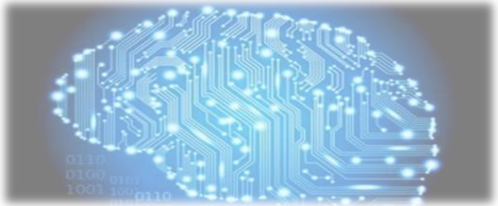
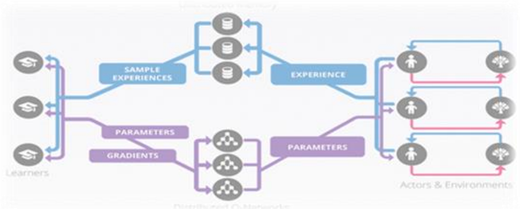




기계학습자동화 연구실

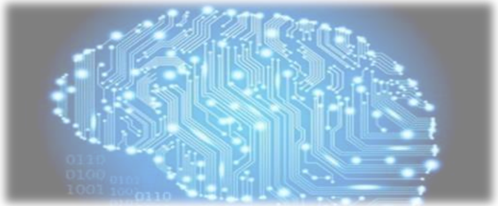
2차 정기 세미나

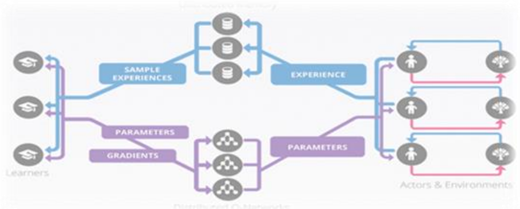
문아성, 이태원, 윤건일, 이한용

2022.03.29

Outline

- Individual Studies
  - Multimodal Learning Research, A-Seong Moon
  - Stock Price Prediction, Tae-Won Lee
  - 2-Way Inference on Embedded Board (DOOSAN), Geonil Yun
  - Restoration of Homoglyph Attack with BERT, Hanyong Lee





Multimodal Learning Research

A-Seong Moon

TO DO List

- Last week
  - ☑ Multimodal Image Classification: Image + @
  - ☑ Master's thesis: An efficient neural network specialized for sparse CT Slice images
- This week
  - ☑ Multimodal Learning Research

Traditional Multimodal Fusion Methods

Three typical multi-modality fusion strategies: Early fusion, Model-level fusion, Late fusion method

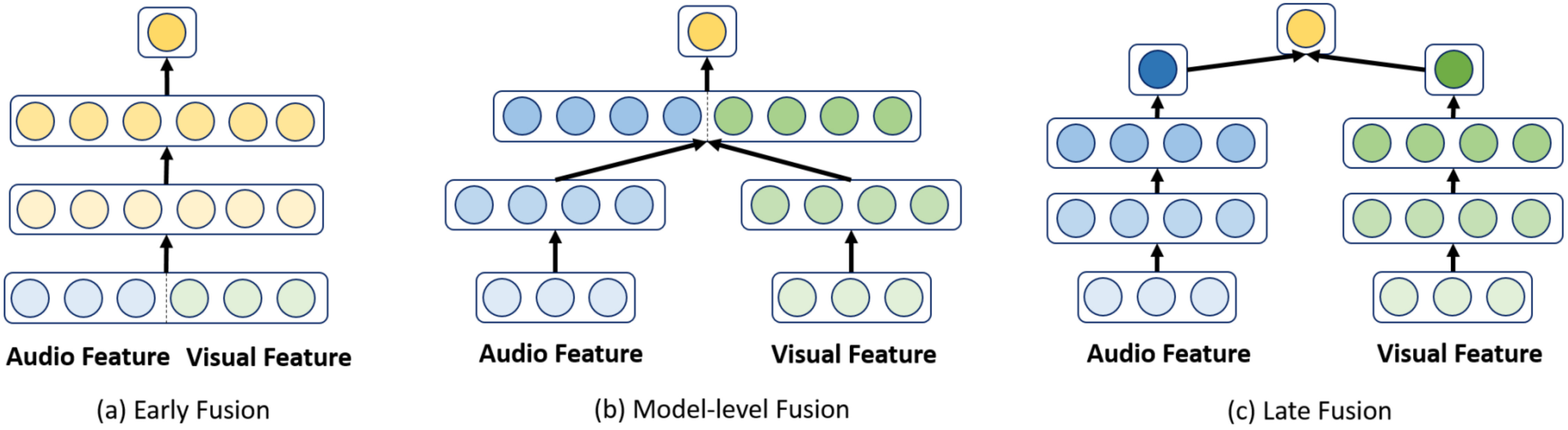


Figure 1: Three typical multi-modality fusion strategies. (a) Early fusion: concatenation of features from different modalities. (b) Model-level fusion: concatenation of high level feature representations from different modalities. (c) Late fusion: fusion of predictions from different modalities.

MMTM: Multimodal transfer module for CNN fusion

Joze, Hamid Reza Vaezi, et al. "MMTM: Multimodal transfer module for CNN fusion." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

- **Multimodal Transfer Module**
  - *Squeeze and Multimodal Excitation*

    To recalibrate channel-wise features in each modality

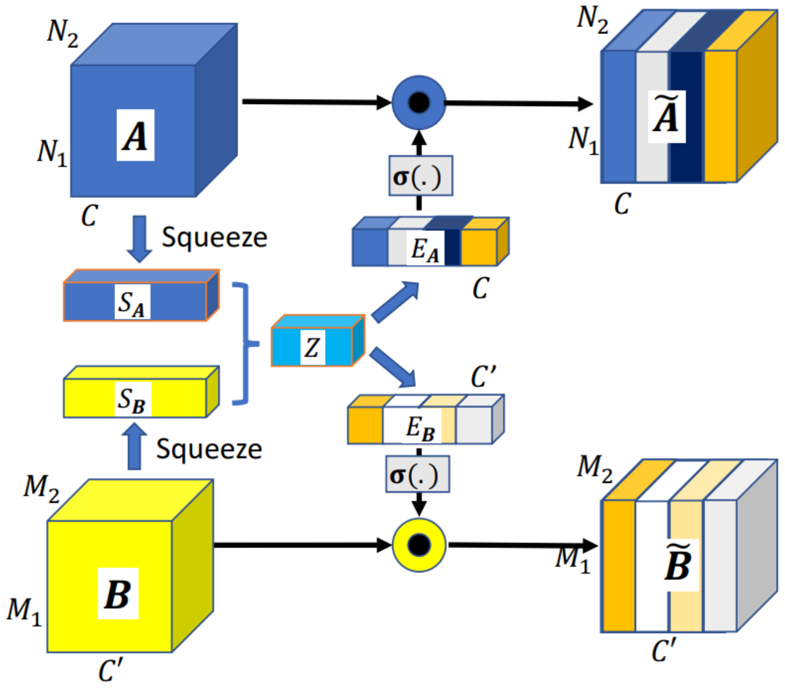


Figure 2. Architecture of MMTM for two modalities. A and B, that represent the features at a given layer of two unimodal CNNs, are the inputs to the module.

Deep Multimodal Fusion by Channel Exchanging

Wang, Yikai, et al. "Deep multimodal fusion by channel exchanging."
Advances in Neural Information Processing Systems 33 (2020): 4835-4845.

- **Channel Exchanging Module**
  - *Replace sparse channels with different modalities*

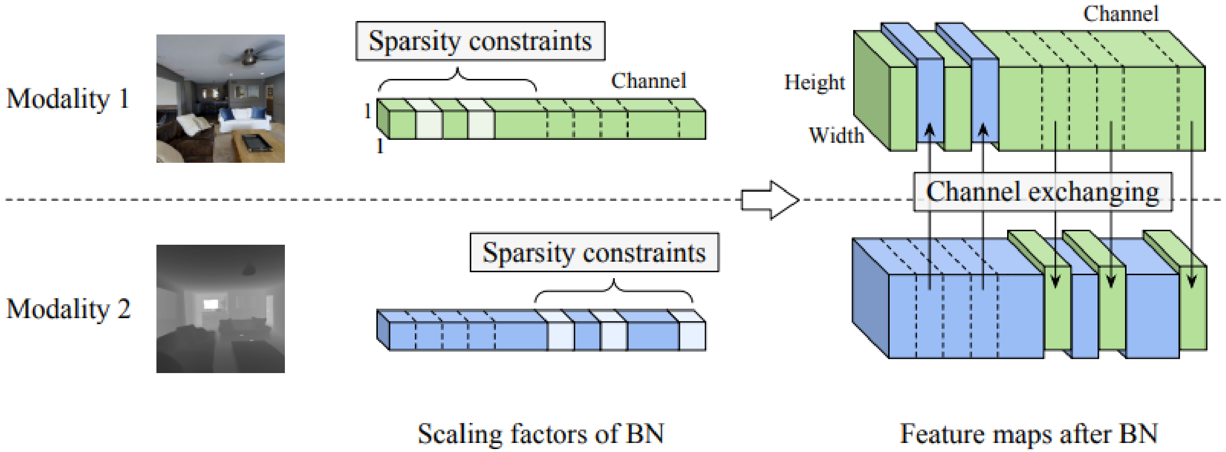


Figure 3: An illustration of our multimodal fusion strategy. The sparsity constraints on scaling factors are applied to disjoint regions of different modalities. A feature map will be replaced by that of other modalities at the same position, if its scaling factor is lower than a threshold.

Efficient Low-rank Multimodal Fusion with Modality-Specific Factors

Liu, Zhun, et al. "Efficient low-rank multimodal fusion with modality-specific factors."
arXiv preprint arXiv:1806.00064 (2018)., 인용 수: 208

- **Low-rank Multimodal Fusion**
  - *To increase efficiency by removing low-ranking modality*

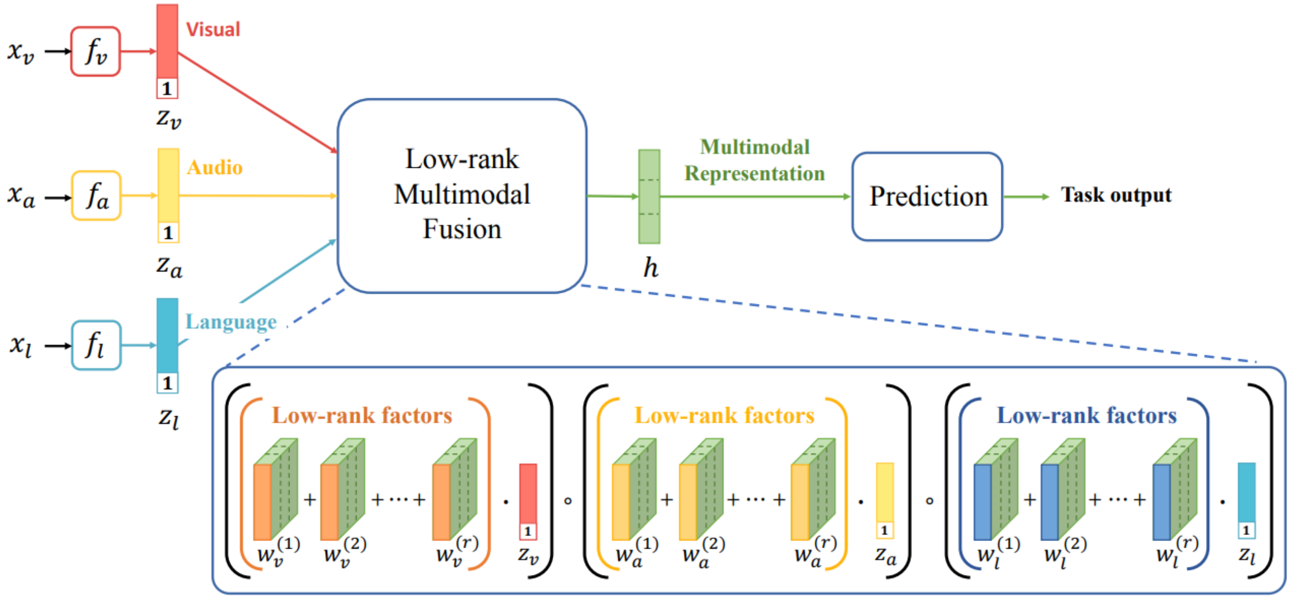
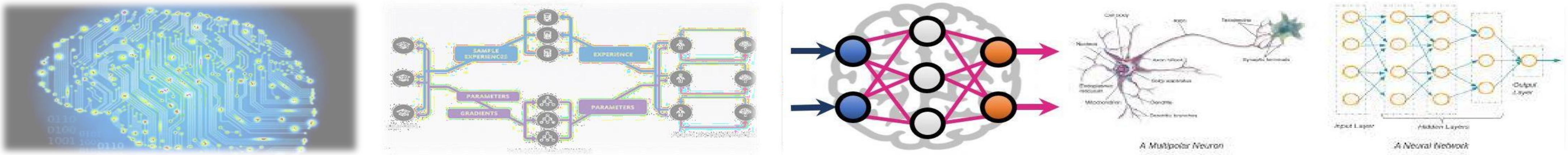


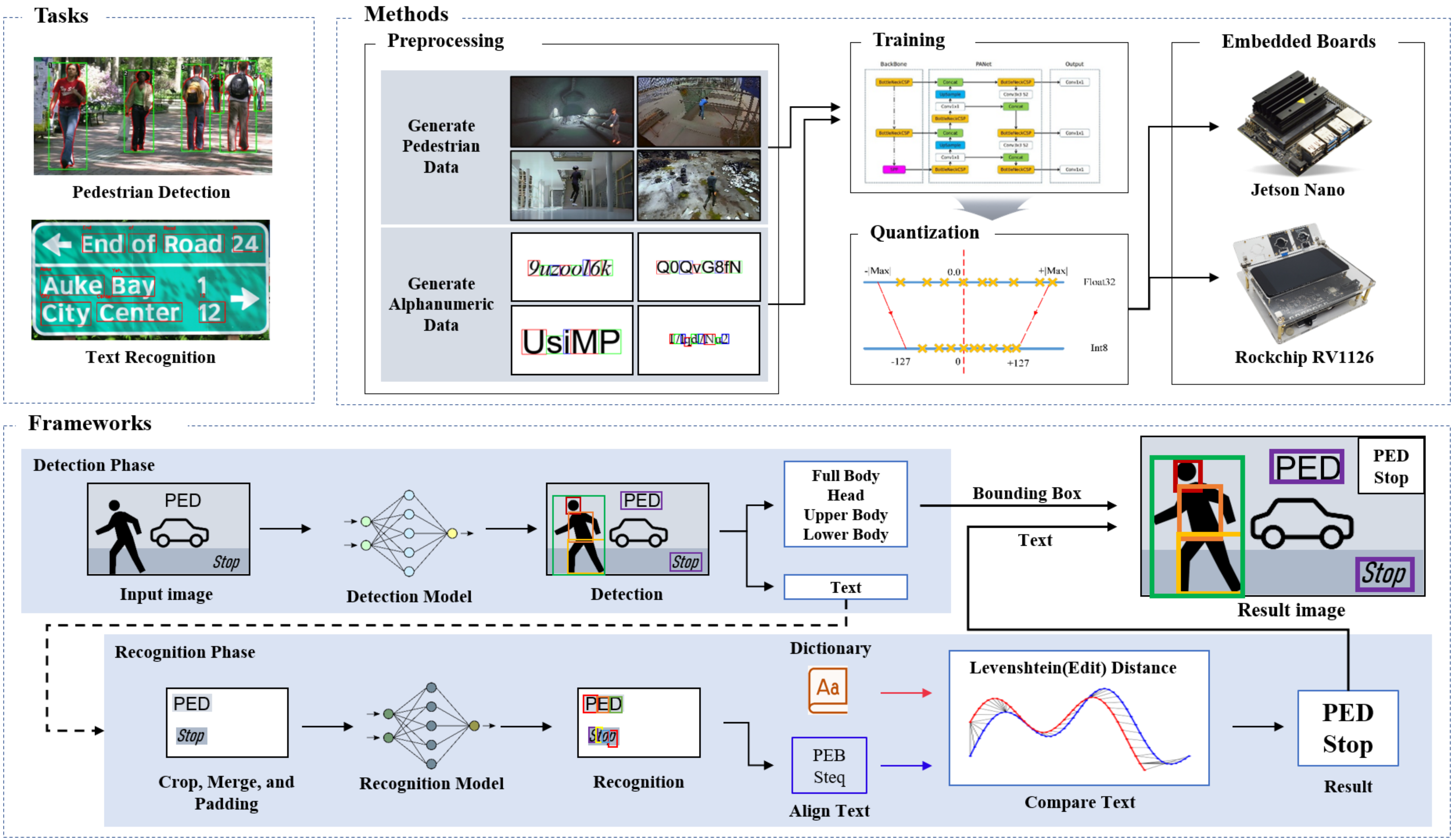
Figure 1: Overview of our Low-rank Multimodal Fusion model structure: LMF first obtains the unimodal representation z_a, z_v, z_l by passing the unimodal inputs x_a, x_v, x_l into three sub-embedding networks f_v, f_a, f_l respectively. LMF produces the multimodal output representation by performing low-rank multimodal fusion with modality-specific factors. The multimodal representation can be then used for generating prediction tasks.



Text Summarization

Geonil Yun

2-Phase(Pedestrian Detection & Text Recognition) Inference on Embedded Board



Text Summarization

- **Text Summarization approaches**

  - **Extractive text summarization**
    - Select the most important sentences in the documents and concatenate them.

  - **Abstractive text summarization**
    - Generate the summary with sentences.

  - **Hybrid text summarization**
    - Combine both the extractive and abstractive text summarization.

Text Summarization

- **To do**

  - **Survey state-of-the-arts(or implementation)**
    - BART(2020), PEGASUS(2020)

  - **Collect datasets & Evaluation Methods**

Restoration of Homoglyph Attack with BERT

Hanyong Lee

Restoration of Homoglyph Attack with BERT

✓ To Do List

- Last Week
  1. Making homoglyph dictionary for test dataset
  2. Annotating for homoglyph dataset.
  3. Removing junk data.

- This Week
  1. Defining Metric
  2. Evaluating the Baseline Algorithms
  3. Pretraining a BERT Masked Language model

1. Defining Metric

- Using **Accuracy**

- $$accuracy = \frac{n\left(\left\{char_{predicted}^{(i)} \middle| char_{predicted}^{(i)} = char_{label}^{(i)}\right\}\right)}{n\left(\left\{char_{original}^{(i)} \middle| char_{original}^{(i)} = homoglyph\right\}\right)}, i = index\ of\ character$$

i

Original Sentence: … access to yoür operatíng system …

$char_{original}^{(i)}$

Label Sentence: … access to your operating system …

$char_{label}^{(i)}$

Predicted Sentence: … access to yorr operatíng system …

$char_{predicted}^{(i)}$

2. Evaluating the Baseline Algorithms

- Baseline Algorithms

1. **OCR**
   Reference: Sawabe, Yuta, et al. "Detection method of homograph internationalized domain names with OCR." Journal of Informati on Processing 27 (2019): 536-544.

2. **SimChar**
   Reference: H. Suzuki, D. Chiba, Y. Yoneya, T. Mori, and S. Goto, "ShamFinder: An Automated Framework for Detecting IDN Homographs," Proceedings of the 19th ACM Internet Measurement Conference (IMC 2019), pp. 449--462, October 2019

3. **Manually Labeled Mapping Table**
   Labeling 318 non-ASCII characters manually.

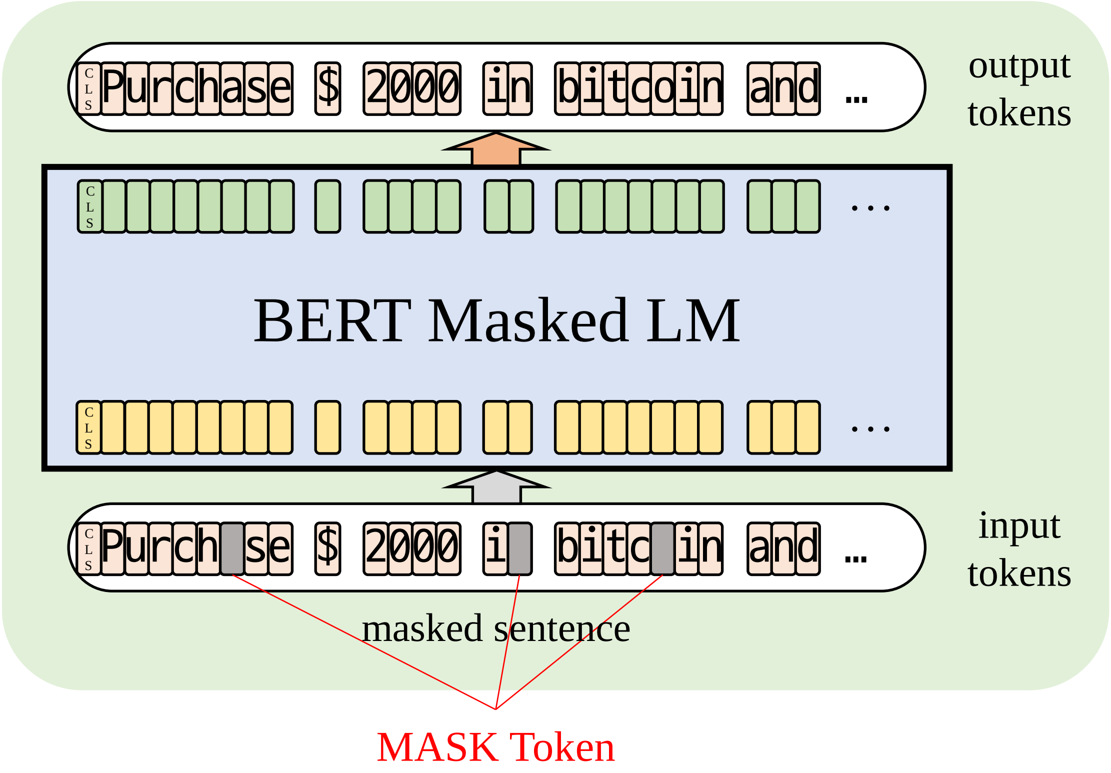
2. Evaluating the Baseline Algorithms

- Evaluating algorithm using **OCR, SimChar Database, Manually Labeled Mapping Table**

Algorithm	**Accuracy**
OCR	0.725
SimChar Database	0.635
Manually Labeled Table	0.949
BERT MLM	Training...

3. Pretraining a BERT Masked Language model

- Pre-training BERT Masked ML
- Mask Probability: 0.15
- Train / Dev set: 95% / 5%
  (total 93910 examples)
- **Work in progress**



4. Plan For Next Week

- Evaluating a BERT MLM.
- Pretraining a BERT MLM with more corpora. (if needed)