

# Captioning Studies (all connections are residual)

## 1. Image Captioning

- ☐ LSTM, GRU tabanlı RNN'ler
- ☐ Attention mechanisms
- ☐ Auxiliary Logits
- ☐ Feature Injections (not includes attention mechanisms)
  - 1. Init-inject
  - 2. Pre-inject
  - 3. Par-inject
  - 4. Merge
- ☐ Datasets
  - 1. MSCOCO
  - 2. VizWiz

## 2. Video Captioning

- ☐ LSTM, GRU tabanlı RNN'ler
- ☐ Attention mechanisms (Frame ve zaman'a bağlı)
- ☐ Auxiliary Logits
- ☐ Feature Injections (not includes attention mechanisms)
  - 1. Init-inject
  - 2. Pre-inject
  - 3. Par-inject
  - 4. Merge
- ☐ Bidirectional RNN'ler (encoderde)
- ☐ Datasets
  - 1. MSVD
  - 2. MSR-VTT
- ☐ # of frames
  - 1. 2
  - 2. 4
  - 3. 8
  - 4. 16

## 3. Hyperparameters:

- ☐ Seed Numbers:
  - 1. 0-5

- ☐ RNN katman sayısı
  1. 3-10
- ☐ Optimizers:
  1. SGD
  2. Adam
  3. RMSProp
  4. Rprop
  5. ASGD
  6. LBFGS
  7. Adamax
  8. AdamW
  9. Adagrad
  10. Adadelata
- ☐ Learning Rates:
  1. 0.001
  2. 0.005
  3. 0.01
  4. 0.02
  5. 0.05
- ☐ Batch sizes:
  1. 16
  2. 32
  3. 64
  4. 128
- ☐ Embedding Sparse Tensors optimizers:
  1. SGD
  2. SparseAdam
- ☐ Pretrained word embeddings:
  1. charngram.100d
  2. fasttext.en.300d
  3. [fasttext.simple.300d](#)
  4. glove.42B.300d
  5. glove.840B.300d
  6. [glove.twitter.27B.25d](#)
  7. [glove.twitter.27B.50d](#)
  8. [glove.twitter.27B.100d](#)
  9. [glove.twitter.27B.200d](#)
  10. glove.6B.50d

11. glove.6B.100d
12. glove.6B.200d
13. glove.6B.300d