Few Shot Image Classification

Aman Butoliya 203070001
Pasunuri Prathiba 203070024
Piyush Singh Pasi 213050069

Contributions and conflict of interest

- Quantitative:
  - Aman : 100%
  - Prathiba : 100%
  - Piyush : 100%
- Qualitative:
  - Aman : Data processing, Creation of episodes
  - Prathiba : ResNet, Auto-encoder
  - Piyush : Self-Attention, Contrastive Loss, Euclidean Similarity

Problem Statement

Few-Shot Learning:

- ○ FSL means N-way-K-shot-classification, N stands for number of classes and K for the number of samples in each class
- ○ Few Shot Learning is characterized as a Meta Learning Problem. In Meta-learning model tries to learn how to classify things/images **(Learning To Learn).**
- ○ FSL promotes actual learning with fewer annotations !!

Such a model that learns to classify images for **unseen classes.** Model should takes a few samples per class (support set) and samples that need to be classified into these classes (query set). For the model to truly it should be able to classify on training disjoint class.

Related Work:

- There are models for image classification using Few-Shot Learning.
- Few methods are on two stage approach, first training and then fine tuning.
- Others are using Metric based Prototypical Networks, Matching Networks based on attention mechanism, LSTM Meta Learner..etc.
- We are using simple structured gradient based model and achieving comparable results with existing models

Model Architecture

- FSL is challenging **(MAML & Prototypical network).**

- _Our goal_: understand FSL with simple structured gradient based model.

- Used ResNet18 & AutoEncoder as feature extractor

- Self-Attention for better representation.

- Contrastive loss to bring support set images representation closer.

- Classifier with Euclidean Similarity measure

- Input: An episode (support images, support labels, query images, query labels)

- Output: predicted query labels

Model Architecture
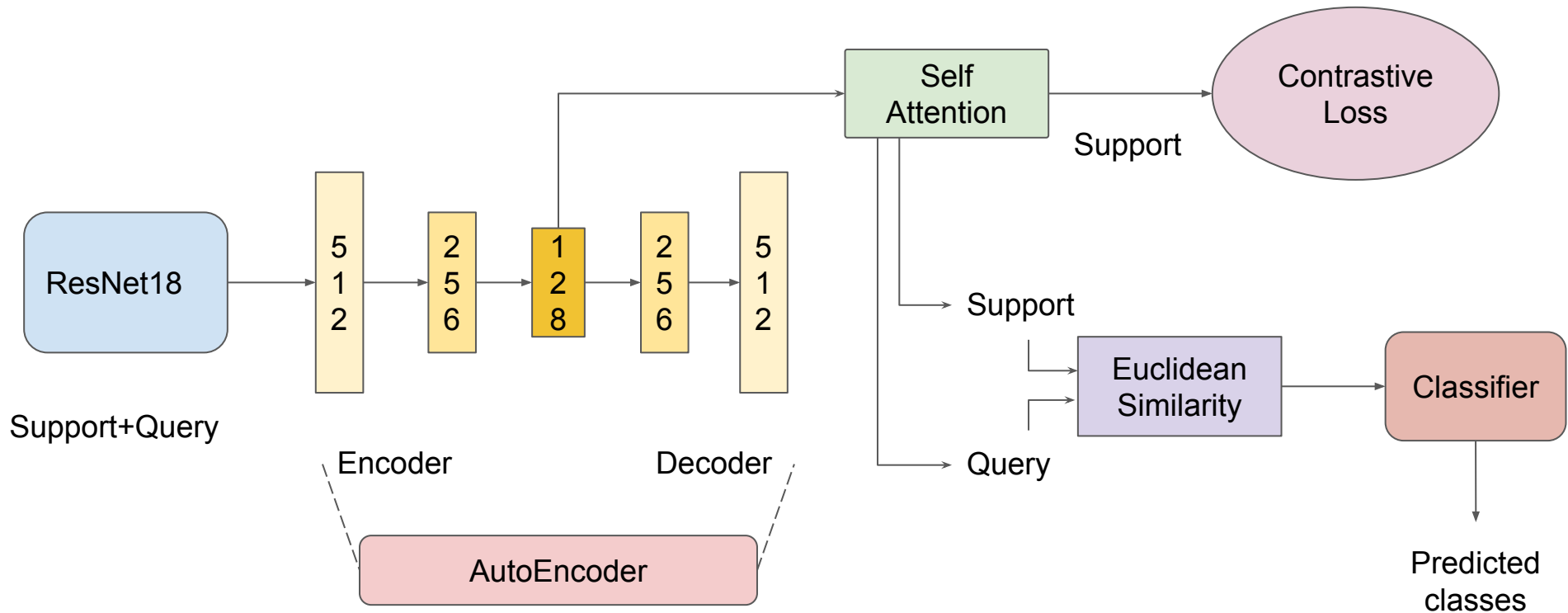
Dataset:

- We are using Caltech-UCSD Birds-200-2011 (CUB-200-2011) dataset
- Details of the dataset:
  - Number of Classes: 200
  - Number of images: 11,788
  - 70%, 10%, 20% split for train, val, test sets (class disjoint)


Count of classes

No. of Images
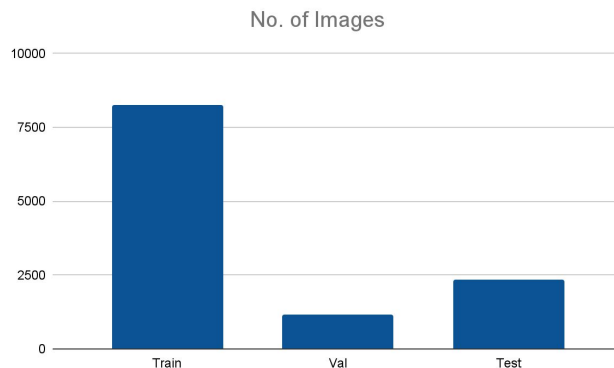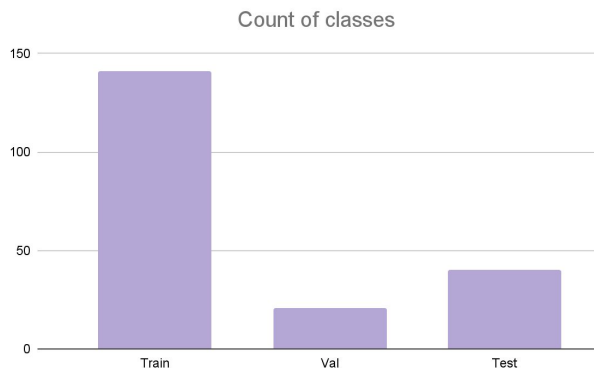
Techniques

- We perform a union of different techniques

- ResNet18, AutoEncoder, Contrastive Learning, Attention

- Euclidean similarity measure borrowed from prototype networks

- Losses:

  - Reconstruction Loss : MSE

  - Contrastive Loss

  - Cross entropy loss

- Metrics: Accuracy, F1-score aggregated over all episodes:

- Best Acc: 62%, F-score:61% trained over 20 epochs with early stopping 5

Results

Varying N & K

● N   ◆ K
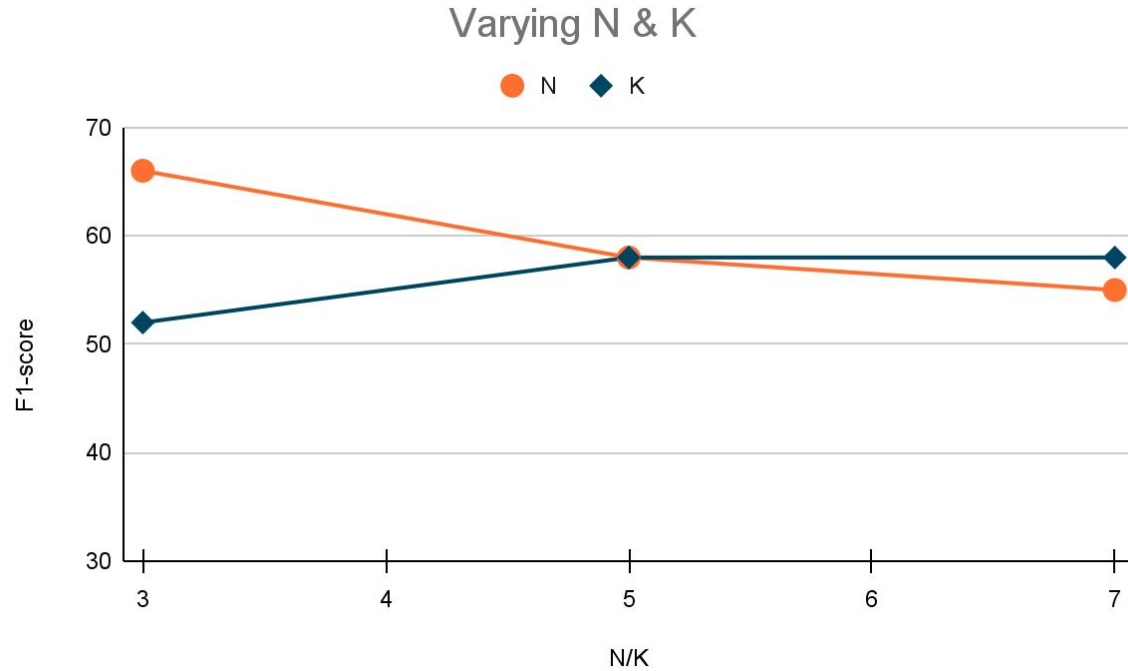
Trained over 10 epochs to study the effect of N & K (other is kept constant)

Conclusion

- We propose a simple FSL model that is an amalgamation of different techniques we came across.
- Our model performs well on unseen classes (training disjoint).
- We perform quantitative analysis on the key hyperparameters N & K.
- And, We fulfill all our deliverables as promised in the draft.

End!