Pre-Meeting - Thursday, June 6th, 2019

Peyman Bateni

June 6, 2019

1 Updates

- I was busy finishing up and submitting a paper to CoNLL and then working on compiling the auxillary material and model saves all together should the paper get accepted. This has been since our last chat which took plenty of sleepless nights, but we finished it in time. I can provide a copy of the work if you are interested.
- I thought I was done with the literature review after finishing up the last five articles and writing the summaries for the two I thought would be useful, but a thought experiment on extending "Graph R-CNN for Scene Graph Generation" to the few-shot case resulted in me going down more papers.. hopefully will be done by end of this week.
- I've reviewed BAA and SOW, although I skimmed through parts with respect to timeline as I mainly focused on problem definitions. Second read will be needed when planning to meet the deadlines.
- I'm currently writing up the research space comparison as suggested by Professor Sigal. While I've updated the potential sections for the summaries and added details to them to make the comparison more detailed, it still needs visualization (hopefully also this week), but the summary is as follows:
  - There has been much more work on few-shot classification as oppose to detection. Also, multi-label classification has beene explored rather extensively in 2018 although there is still potential (that said multi-label deviates from the DARPA challenge).
  - Matching networks/prototype networks/siamese networks/... as many have tried various variations try to map visual feature embeddings to a class space (could be word2vec, could be other). MAML prepares model for fast adoptation. Hierchical Bayes allows for transfer or learned features from super classes to sub classes.
- Problem Definition:
  - Assumptions for the classification case:
    - $T^{train}_{source} = \{(X_{0,0}, C_0)...(X_{n,0}, C_0)...(X_{n,m}, C_m)\}$ where $X_{i,j}$ is the ith "source" training example belonging to class j, where T^{train}_{source} consist of m class with n examples each with n ¿ 10000 (? we set this last time but reviewing the LwLL description, I think this may be high).
    - $T^{train}_{target} = \{(X_{0,m+1}, C_m + 1)...(X_{k,m+1}, C_m + 1)...(X_{k,m+h}, C_m + h)\}$ where we have k target classes with h examples each with $1 \leq h \leq 10$.
    - $T^{train}_{zero-shot} = \{c_{m+k+1}...c_{m+k+d}\}$ defined the zero shot case where no examples are provided for the d zer-shot classes. While the notation allows for T^{train}_{target} and $T^{train}_{zero-shot}$ to both be part of the problem definition, they are usually dealt with separately in the few-shot learning case (with d = 0) and the zero-shot learning case (with h = 0).

* At test time separate T^{train}_{source}, T^{train}_{target} and $T^{train}_{zero-shot}$ are provided although the number of examples per class in T^{train}_{source} and T^{train}_{target} may be different from that of training although still consistently the same for each class.
    * Assuming an integer encoding of $c \in Z$ we are given a function $D : c- > \Sigma^+$ where $\sigma = \{A, B, ...\}$ with A, B, .. denoting the string labels.
    * Assuming presence of a secondary source of class embeddings such as word2vec, there exists a function $J : c- > V$ such that $V = V_1, V_2, ...V_{m+k+h}$ consisting of the vector embeddings for each class.
  - Assumptions for the detection case:
    * Problem setting is somewhat similar with the addition of boundary boxes to each example/-class topple in each of the sets.
    * $T^{train}_{source} = \{(X_{0,0}, B_{0,0}, C_0)...(X_{n,0}, B_{n,0}, C_0)...(X_{n,m}, B_{n,m}, C_m)\}$
    * $T^{train}_{target} = \{(X_{0,m+1}, B_{0,m+1}, C_m+1)...(X_{k,m+1}, B_{k,m+1}, C_m+1)...(X_{k,m+h}, B_{k,m+h}, C_m+h)\}$
    * $T^{train}_{zero-shot} = \{c_{m+k+1}...c_{m+k+d}\}$
    * Similar functions D and J are present.
  - Distinctions to make:
    * Few-shot learning -¿ only performance on T^{test}_{target} is important.
    * Zero-shot learning -¿ only performance on $T^{test}_{zero-shot}$ is important.
    * Generalized few-Shot learning -¿ performance is defined as classification/detection accuracy on both T^{test}_{source} and T^{test}_{target}.
    * Generalized zero-Shot learning -¿ performance is defined as classification/detection accuracy on both T^{test}_{source} and $T^{test}_{zero-shot}$.
    * Open-set learning -¿ in either case is when h (ie. number of few-shot cases) or d (ie. number of zero-shot cases) are not given (may be unbounded, may be only given at test time).
- Five ideas I suggested last time:
  - Visual GloVe
  - Contextual graph matching
  - Unbounded hierarchical nonparametric bayes
  - Transfer functions with mixture of gaussian boundary definitions
  - Generalizable AIR

2 Goals for this Meeting

- Going through the problem definition and discussing solutions.
- Added papers if necessary, research guidance, other material on prob modelling/bayes I should go over.