Visual Dialog – Remake

1. Introduction & Related works 1.5 trang.
2. Biểu đồ thuật toán. 0.5 trang.
3. Một hình vẽ tổng quát model 2 trang.
4. Một table Ablative study: (2 trang)
   1. Number of attentions: 0, 1, 2, 3
   2. Share/Not sharer attention layers
   3. Train Disc/Gen/Misc
   4. Bottom-up attention features 25, 36, 50
   5. Aggregration: [Img, Hist, Ques], [Img, Ques], [Ques]
   6. Positional Embeddings for Hist, Ques and Image.
   7. Fine-tune with dense annotations.
5. Table so sánh với state-of-the-art. (2/3 trang)
6. Một table về Ensemble: (1/3 trang)
   1. Disc/Gen
   2. Disc + Gen: Avg.
   3. Disc + Gen: Reprocical
   4. 4 Disc + 4 Gen.
7. Analysis Computational cost: Table of Floppings & Parameters. (1/3 trang)
8. Qualitative Results. (2/3) trang

Thực nghiệm:

V1.0 – Legacy – share attention – [img, ques] – 30 epochs.

V2.1 – BTA – [img, ques] – 30 epochs.

V2.2 – BTA – share attention – [img, ques] – 30 epochs.