Face to Cartoon translation

Jo MinKi

1. Introduction & Related works  
   In recent years, CNN have become powerful model for computer vision tasks. However, the computational cost is little bit expensive so there were many trial for compressing the model. These works have a similar goal, 1) reduce the model size, 2) decrease the run-time memory footprint, 3) lower the number of computing operations without losing the accuracy.
2. Related works.  
   1. Low rank Decomposition  
      Since I’m not good at the linear algebra, I don’t know what this exactly is. As far as the explanation, approximates the weight matrix in low rank matrix using techniques like Singular value decomposition. This method usually works well on fully connected layer.

Pros: Able to compress the model size to 1/3  
Cons: This does not give any notable speed acceleration, because computing operation in CNN mainly come from convolutional layer. Furthermore, according to recent study that FC layer makes the model easily prone to the overfitting, many people building there model without FC layer as possible as they can.

* 1. Weight Quantization   
     I heard that this method compress the model so the mobile device can handle it. Before the training, network weights are gashed to the binary numbers.   
       
     Pros: Through this method, we can get 35x to 49x compressed AlexNet and VGGNet.   
     Cons: This method can neither save run-time memory nor inference time, since during inference shared weights need to be restored to their original position. In addition, I heard that 이건 그냥 생 노가다라고 하던데
  2. Weight Prunning / Sparsifying 이거 왜 쓰다 말았

1. L1 norm regularization   
   <https://medium.com/mlreview/l1-norm-regularization-and-sparsity-explained-for-dummies-5b0e4be3938a>
2. Issue
   1. Data
      1. Group data by sex   
         Might not be a problem cause male character and female character looks similar..
      2. Collecting Data (프리드로우, 외모지상주의(머리카락), …)  
         There was an opinion that the quantity of the data might be the problem.
   2. Shape of the face
      1. Facial expression estimation   
         According to the other work, it can be estimated by comparing the color value and location value of the landmark point.
         1. However, it seems that this model cannot find the exact position of the landmark when it get a cartoon data. (especially the eyes.)
      2. Cartoon feature   
         Cause the target domain(cartoon) is too much different from target domain, some hard conditioning considering the Cartoon character’s feature might be needed. For example, resizing the eyes could be one of them.
         1. At the former work(UGATIT), possible issue was that learning the attention from the scratch might go wrong, so giving the attention by logic would’ve been help. Actually It was just fine without that though.
3. 섞을거
   1. Deformation
   2. 3D mm <http://gvv.mpi-inf.mpg.de/projects/MZ/Papers/arXiv2017_FA/paper.pdf>   
      Facial information estimation
   3. Python package blender  
      3D model to 2D image (막해도 괜찮을거다는 의견 둘)

Question.

1. Do we know the ? If we don’t, where are we heading for?  
   1. No. we are not knowing the exact function that we should head for. In practically, we are just adding a skip connection on the network so we get the ­­ as a result. So don’t take the explanations on the 3.1 too much seriously.