EECE 7370 Assignment 2

Review for the paper: Group Normalization

1. **Reviewer:** Sakshi Bhatia, Reviewed on: 29 September, 2024

2. **Paper Details:**
   a. Paper Title: Group Normalization
   b. Authors: Yuxin Wu, Kaiming He
   c. Published: 2018, ECCV (European Conference on Computer Vision)
   d. Citation: Wu, Y., He, K. (2018). Group Normalization. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science(), vol 11217. Springer, Cham. https://doi.org/10.1007/978-3-030-01261-8_1

3. **Paper Summary:** The paper presents a novel normalization technique called Group Normalization (GN). It addresses the limitations of other normalization techniques such as Batch Normalization (BN), Layer Normalization (LN), and Instance Normalization (IN), and improves upon those. GN divides channels into groups and computes the mean and variance for normalization within each group. Since it normalizes along groups of feature channels, its computation is independent of the batch sizes.

4. **Main Contributions:**
   a. Group Normalization technique: The concept of Group Normalization is introduced. The mathematics and code implementation of this method are well explained.
   b. Detailed Experimental analysis: The paper performs various experiments to compare the accuracy of BN with other normalization techniques over a wide range of architectures and computer vision tasks.

5. **Paper Strengths:**
   a. Well-written and explained: The paper discusses in detail the drawbacks of common normalization techniques, and a detailed comparison of the proposed method is done with the existing methods. The paper builds well upon how the existing limitations have been overcome by the Group Normalization method. The extensive experiments and analysis strengthen the authors' claims.
   b. Improved accuracy: With a batch size of 2, GN has a 10.6% lower error than BN on ResNet-50 trained in ImageNet. GN shows improved results over BN on Mask R-CNN on COCO object detection and segmentation and 3D convolutional networks for Kinetics video classification.

c. Simplicity: GN is simple and can be easily implemented in deep learning architectures by just grouping channels and applying normalization over these groups.
d. Versatility: GN is demonstrated to work well across different Computer Vision tasks such as image classification, object detection, image segmentation, and video classification. It is experimentally shown to work well on various architectures such as ResNet-50, ResNet-101, VGG-16, and Mask R-CNN. GN allows us to have good accuracy with small groups, hence making it feasible to work on 'high-capacity' models while being memory efficient. BN has a trade-off between the temporal length and batch size when spatial-temporal features are present. This limitation is also overcome by Group Normalization. Additionally, GN can easily transfer from pre-training to fine-tuning. The authors also mention the potential applicability of this method to sequential and generative models.

6. **Paper Weaknesses:**
    a. Group Hyperparameter Optimization: The paper introduces another hyperparameter, i.e., the number of groups, G. This hyperparameter would have to be chosen/optimized carefully depending on your use case. The paper does not touch upon the topic of choosing the optimal hyperparameter G.
    b. Computational Speed: The authors do not talk about the computational efficiency of this normalization technique as compared to other common normalization techniques such as Group Normalization.
    c. Performance with large batch size: When working with large batches, such as 32 images per GPU, the performance of GN is a little worse than BN.

7. **Comments on Experiments:** The paper provides sufficient experiments across various types of architectures and different computer vision applications. The authors provide comparisons against BN, LN, and IN, showing that GN performs better, particularly when batch sizes are small. The performance of the method with varying group sizes has also been demonstrated well. The experimental results are convincing for this paper.

8. **Final Recommendation: Accept**
    The paper offers a significant contribution to the field of Computer Vision by proposing an improved normalization method in cases where we require small batch sizes and may have memory constraints. It is well written and well illustrated with plenty of graphs, tables, and visuals where necessary. Given the rigorous evaluation and its potential practical applications, this paper deserves acceptance.