Supplementary - Black-Box Adaptation for Medical Image Segmentation

Jay N. Paranjape[1,✉], Shameema Sikder[2,3], S. Swaroop Vedula[3], and Vishal M. Patel[1]

[1] Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, USA
jparanj1@jhu.edu
[2] Wilmer Eye Institute, The Johns Hopkins University, Baltimore, USA
[3] Malone Center for Engineering in Healthcare, The Johns Hopkins University, Baltimore, USA

SPSA-GEASS: Given a point prompt, the foundation model produces a reasonable mask without any sort of optimization (zero-shot performance). Thus, the approximated gradients are small in magnitude and the system tends to get stuck at local minimas. In order to alleviate this, we propose a small modification to the SPSA-GC algorithm called SPSA-Gradient Enhancement Assistance (SPSA-GEASS). As shown in Figure 1, we define two parameters *strike* and *cooldown*. *strike* is initialized to zero during the start of the training. During training, if the magnitude of gradient $\hat{g}_i$ falls below a threshold, *strike* is incremented until it reaches a limit k_1. After this, the learning rate α and the perturbation step parameter c in SPSA-GC are multiplied by hyperparameters η_1 and η_2 respectively, for a number of iterations equal to *cooldown*. This emulates the effect of taking a large step so that the model might move out of a local minima. The parameter *strike* determines the tolerance of low-magnitude gradients during training, while the parameter *cooldown* controls the number of iterations for which large steps are taken. However, these parameters ($strike, cooldown, \eta_1, \eta_2, k_1$) require careful selection and would vary across datasets and other hyperparameters like the learning rate.

Training Progress of BAPS: To test the effectiveness of BAPS, we compare its training progress with that of Visual Prompting for two datasets with MedSAM as the foundation model. These plots are shown in Figure 2. Here, we see a clear gap between the loss values, with BAPS consistently showing a lower loss value.

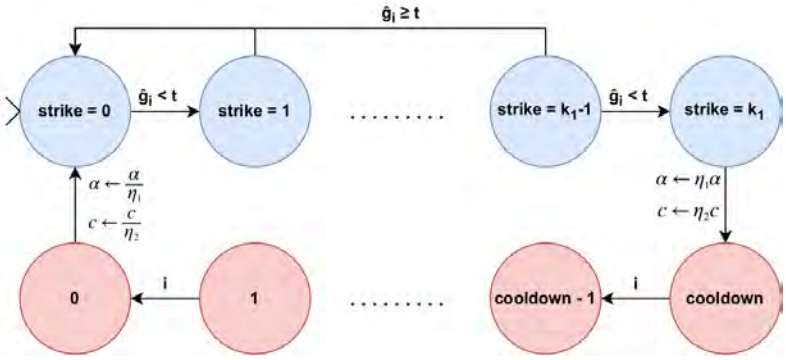
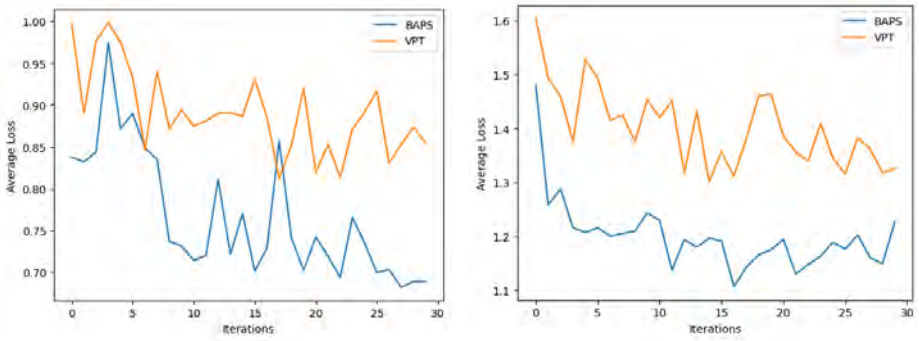


Fig. 1: Overview of SPSA-GEASS. The system starts at $strike = 0$. If the estimated gradient is greater than the threshold, *strike* increases, or else the system reverts to the original state. If *strike* reaches k_1, the learning rate and perturbation step parameter increase significantly. Then, the *cooldown* reduces every iteration until 0. After this, the system returns to its initial state.



(a) ISIC2018 (b) REFUGE

Fig. 2: The training progress of BAPS versus VPT with MedSAM as the blackbox foundation model.