Second sweep

A report written after completing the second sweep.

Jo Falck-Ytter

Created on May 8 | Last edited on May 8

The data is SN-BAS extracted with VideoMAE-V2, the same data as the first sweep. There is a very positive outlier with a mAP at almost 0.49. If I ignore that the difference is again ~0.03 from top to bottom. The weight decay scope was lowered, and it proved to be an even more influential parameter now. Perhaps it should be even lower. **Learning rate** has very little correlation, unlike the suggestion from the first sweep. But for the last sweep I will try to have the learning rate even lower, as the range in this test proved to be too small. **Droppath** was the most important parameter according to the importance metric in the first sweep. It now dropped. Because the correlation is slightly small, I will select 0.03 out of the four values because it is low but not lowest.

voting_thresh and **trunc_thresh** will be slightly altered in accordance to their correlation, but I don't expect them to have the largest impact. From inspecting a customizable graph(moving the columns, can be done on wandb.ai), trunc_thres had the second best result with value=0.9899 but most good results came with the lowest value. A lower search space for the next sweep is selected. voting_thresh seems to have a center of gravity for the good results at about 1.1. So I will increase the higher bound and lower bound, although the highest values did not improve over the second highest

All the parameters that will be ignored for the last sweep:

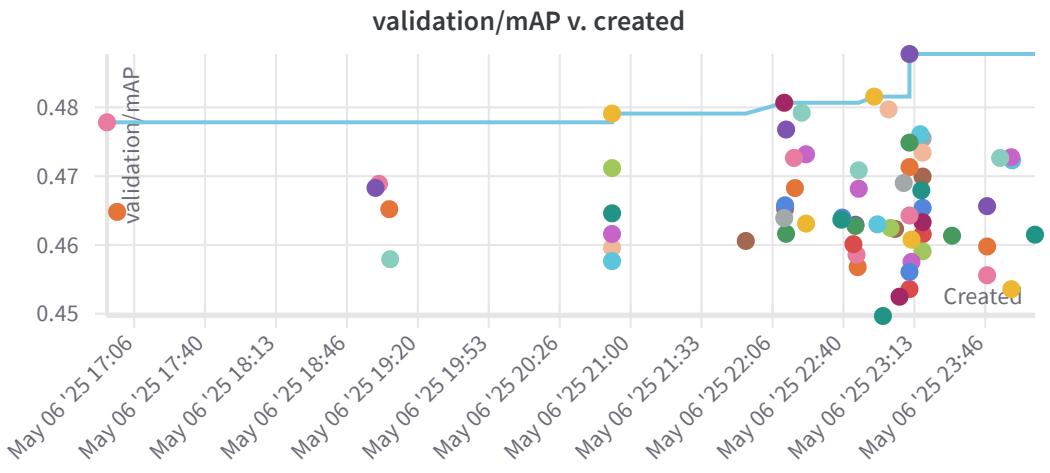
- dataset.max_seq_len
- test_cfg.pre_nms_topk
- train_cfg.droppath
- train_cfg.clip_grad_l2norm

Here is the config used for this sweep, and the diagrams are below it.

```yaml
command:
  - python
  - train_eval.py
  - configs/mamba_custom.yaml
  - --wandb
  - --wandb_project
  - video-mamba-suite
  - --wandb_entity
  - jofalck-ntnu
entity: jofalck-ntnu
method: bayes
metric:
  goal: maximize
  name: validation/mAP
parameters:
  dataset.crop_ratio:
    distribution: categorical
    values:
      - - 0.7
        - 1
      - - 0.8
        - 0.9
      - - 0.8
        - 1
      - - 0.9
        - 1
      - - 0.95
        - 1
  dataset.max_seq_len:
    distribution: int_uniform
    max: 4096
    min: 1024
  dataset.trunc_thresh:
    distribution: uniform
    max: 1
    min: 0.25
  opt.learning_rate:
    distribution: uniform
    max: 0.000125
    min: 8e-05
  opt.weight_decay:
```

```
    distribution: log_uniform_values
    max: 0.04
    min: 1e-06
  test_cfg.multiclass_nms:
    distribution: categorical
    values:
      - true
      - false
  test_cfg.pre_nms_topk:
    distribution: int_uniform
    max: 4000
    min: 1000
  test_cfg.voting_thresh:
    distribution: uniform
    max: 1.4
    min: 0.35
  train_cfg.clip_grad_l2norm:
    distribution: int_uniform
    max: 2
    min: 1
  train_cfg.droppath:
    distribution: uniform
    max: 0.2
    min: 0.05
program: train_eval.py
project: video-mamba-suite
```


Section 1



Parameter importance with respect to validation/mAP

Created with ❤️ on Weights & Biases.

https://wandb.ai/jofalck-ntnu/video-mamba-suite/reports/Second-sweep--VmlldzoxMjY3MDQzMw