Supplemental Material `NCQS`: Nonlinear Convex Quadrature Surrogate Hyperparameter Optimization

Sophia Abraham[1], Kehelwala Dewage Gayan Maduranga[2], Jeffery Kinnison[1], Jonathan Hauenstein[1], and Walter Scheirer[1]

[1]*University of Notre Dame, Notre Dame, IN 46556, USA*
[2]*Tennessee Technological University, Cookeville, TN 38505*

.1. Extended Experimental Setup

.1.1   EVM Open-Set Experiments

Table 2 summarizes the dataset experiments in terms of number of search trials (ST), number of known (KC) and unknown classes (UC), and number of training, validation, negatives, and testing samples. Negatives are samples from the unknown class without labels used to better inform the "1 vs. rest" classifiers when training the EVM. The number of trials for each experiment was determined and adjusted according to the time required to train a single model for the given dataset. If the dataset contained many images per class, fitting the EVM to the feature vectors for large samples required longer compute time and thus the number of search trials was reduced. Additionally the ranges for each of the hyperparameters in the open-set experiments can be found in Table 1.

Threshold, cover threshold, and distance multiplier are sampled from a uniform distribution of range $[0, 1]$. The fitting algorithm for the EVM requires that the tailsize not exceed greater than half the number of training samples. For this reason, the tailsize hyperparameter was sampled from a uniform distribution between the range $[0, 0.5]$ and used as a multiplier to determine how many negative samples were included in estimating the model parameters.

.1.2   ATR Experiments

The images in the DSIAC dataset are captured by two forward-looking, ground-based sensors: an L3 Cincinnati Electronics Night Conqueror MWIR camera using a 640×480 pixel Indium Antimonide focal plane array with a 28 micron pitch, and an Illunis visible light camera. Images were col- lected from both cameras during daytime and from the MWIR camera at night. For each target/range/time scenario, images were captured at 30 Hz for one minute, yielding 1,680 images per sequence. We use an updated DSIAC protocol provided by Domenick Poster [2] where instead of using the provided annotations, we calculate corrected bounding boxes using positional metadata available for the MWIR imagery. We only use data from the following classes: Pickup, SUV, BTR70, BRDM2, BMP2, T72, ZSU23-4, and 2S3 (dismounted civilian and towed artillery scenarios have been omitted). See Table 3 for the sizes of our training, validation, and testing subsets. We only use the data from the 1000 to 2000 dataset ranges for our experiments. The search space for the ATR experiment is also included in Table 4.

.1.3   Closed Set HPOBench Experiments

For specifications regarding the datasets and experiments from HPOBench please refer to the original paper [1].

References

[1] Katharina Eggensperger, Philipp Müller, Neeratyoy Mallik, Matthias Feurer, René Sass, Aaron Klein, Noor Awad, Marius Lindauer, and Frank Hutter. Hpobench: A collection of reproducible multi-fidelity benchmark problems for hpo. *arXiv preprint arXiv:2109.06716*, 2021. 1

[2] Domenick Poster. Defense systems information analysis center (dsiac) atr algorithm development image database. Online, 2023. 1

Parameter	Description	Domain
Threshold	Probability threshold used to determine if an input coordinate point should be classified as 'unknown' if the point falls below this probability of inclusion.	$[0, 1]$
Tailsize	Defines how many negative samples are used to estimate the model parameters.	$[0, 0.5]$
Cover threshold	The probability threshold used to eliminate extreme vectors if they are covered by other extreme vectors with that probability.	$[0, 1]$
Distance multiplier	The multiplier to compute margin distances.	$[0, 1]$
Distance function	The distance function used to compute the distance between two samples.	Cosine or Euclidean

Table 1: Hyperparameters used in training the EVM.

Table 4: Hyperparameters and their domain ranges for Tiny-YOLO in ATR experiment

Experiment	ST	KC	UC
MNIST	200	6	4
LFW	1000	34	5715
ImageNet	200	40	51

Dataset Split	Training	Validation	Testing
MNIST	28824	12000	10000
LFW	1333	2481	3309
ImageNet	10960	19428	48750

Table 2: Experiment details for EVM open-set experiments

Range	Train	Val	Test
1000	1843	461	-
1500	-	-	458
2000	1834	458	-
2500	-	-	425
3000	1728	432	-
3500	-	-	425
4000	1709	427	-
4500	-	-	413
5000	1601	400	-

Table 3: Table with Ranges and Training, Validation, and Test sets for DSIAC data.

Hyperparameter	Domain Range
lr0	[0.0, 0.1]
lrf	[1, 5]
momentum	[0.5, 0.999]
weight_decay	$[10^{-8}, 10^{-5}]$
warmup_epochs	[0, 25]
warmup_momentum	[0.5, 0.9]
warmup_bias_lr	$[10^{-4}, 10^{-1}]$
box	[0, 1]
cls	[0, 1]
cls_pw	[0, 1]
obj	[0, 1]
obj_pw	[0, 1]
iou_t	[0, 0.5]
anchor_t	[1.0, 10.0]
fl_gamma	[0.0, 0.3]
hsv_h	$[10^{-4}, 10^{-1}]$
hsv_s	[0, 1]
hsv_v	[0, 1]
degrees	[0, 180]
translate	[0, 0.3]
scale	[0, 1]
shear	[0, 45]
perspective	[0, 0.001]
flipud	[0, 1]
fliplr	[0, 1]
mosaic	[0.5, 1.0]
mixup	[0.5, 1.0]
copy_paste	[0.5, 1.0]
anchors	[1.0, 10.0]