

Table 1 Hopkins statistic of datasets

Dataset	SIFT1M	GIST1M	DEEP1M
Hopkins	0.496	0.528	0.479
Dataset	Text2image	MTuring	GCD/GUD
Hopkins	0.487	0.480	0.969/0.387

Table 2 Hyperparameters selection sets

Baseline	Parameters Section Sets
Vamana	$R \in \{16, 32, 64\}$, $\alpha \in \{1.0, 1.2, 1.4\}$, $efc \in \{64, 128, 256\}$
HNSW	$M \in \{16, 32, 64\}$, $efc \in \{64, 128, 256\}$
NSG	$R \in \{16, 32, 64\}$, $efc \in \{64, 128, 256\}$
HCNNG	$s \in \{3, 5, 7\}$, $T \in \{5, 10, 15\}$, $Ls \in \{750, 1000, 1250\}$

Table 3 Optimal hyperparameters in the Grid

Baseline	Best parameters for SIFT1M
Vamana	$R = 32, \alpha = 1.2, efc = 64$
HNSW	$M = 16, efc = 256$
NSG	$R = 64, efc = 256$
HCNNG	$s = 3, T = 15, Ls = 1000$
Baselines	Best parameters for GIST1M
Vamana	$R = 64, \alpha = 1.2, efc = 128$
HNSW	$M = 32, efc = 256$
NSG	$R = 64, efc = 256$
HCNNG	$s = 3, T = 15, Ls = 1250$
Baselines	Best parameters for DEEP1M
Vamana	$R = 32, \alpha = 1.2, efc = 64$
HNSW	$M = 16, efc = 128$
NSG	$R = 32, efc = 256$
HCNNG	$s = 3, T = 15, Ls = 1250$

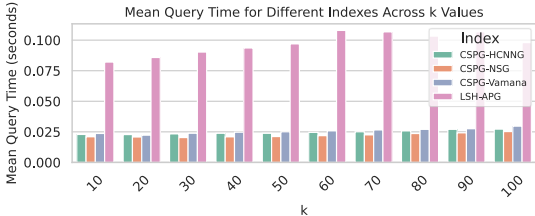


Figure 1 CSPG v.s. LSH-APG on Audio dataset

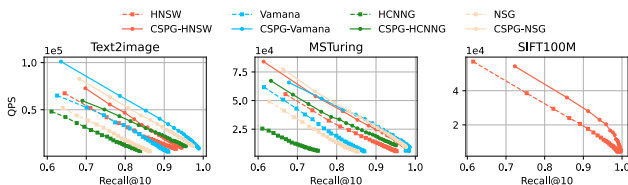


Figure 2 QPS-Recall@10 on harder datasets

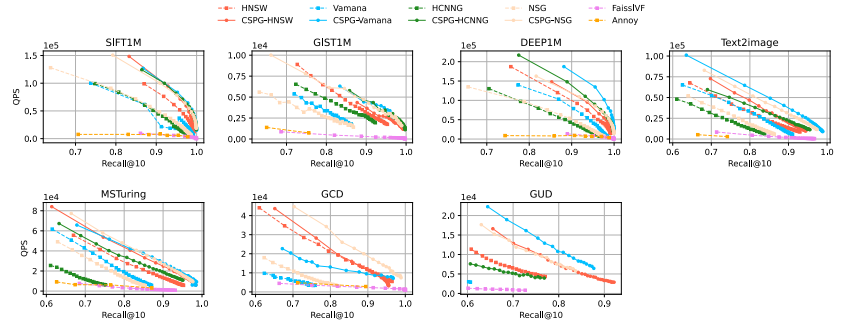


Figure 3 QPS-Recall@10 over different distribution

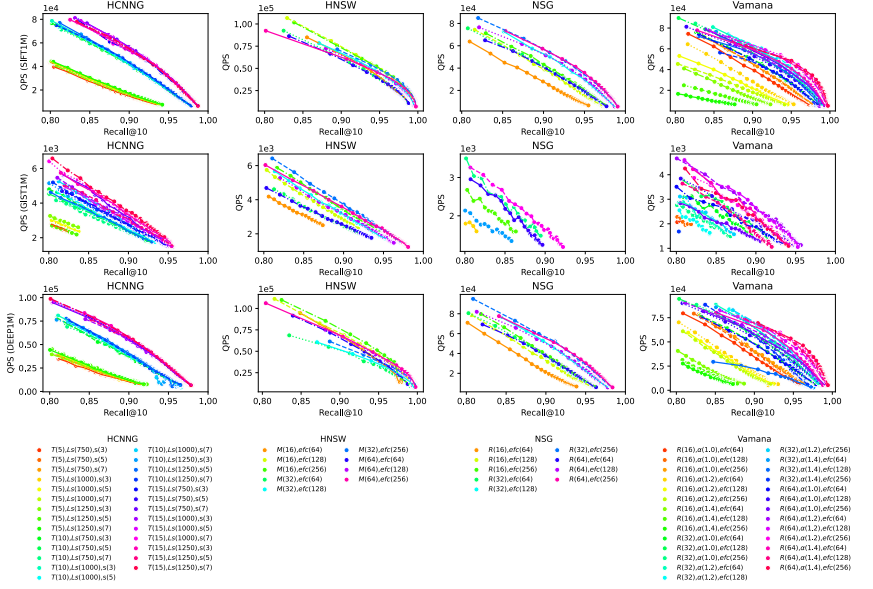


Figure 4 QPS-Recall@10 over different parameters combination

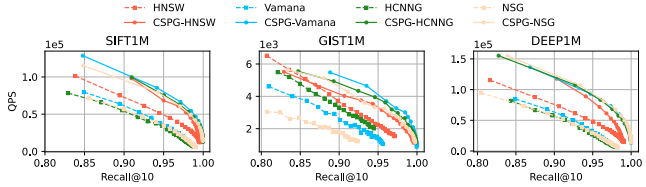


Figure 5 QPS-Recall@10 over optimal parameters

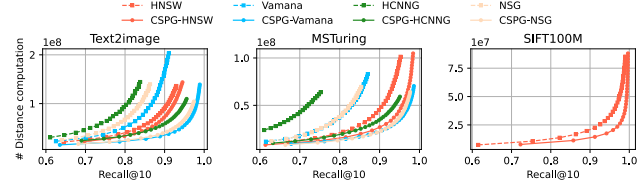


Figure 6 Comparison-Recall@10 on harder datasets