Budget-Aware Data Selection for Low-Shot Image Classification: A Controlled Empirical Benchmark

Aulia Ilmi Maghfira Ridwan*, Zidane Putra Sanjaya, Bagus Prasetiyo, and Novanto Yudistira

Master of Computer Science, Universitas Brawijaya, Malang, Indonesia
auliailmimr@student.ub.ac.id, zidanesanjaya81@student.ub.ac.id, bagusprasetiyo@student.ub.ac.id, yudistira@ub.ac.id

Abstract. This paper presents a controlled empirical benchmark of data selection for low-shot image classification under a limited annotation budget. Given a small labelled training set and an additional unlabelled candidate pool, we investigate which samples should be selected to improve downstream classification performance. We evaluate ten experimental conditions on CIFAR-10 and Fashion-MNIST using identical data splits, model architectures, training settings, budget levels, and random seeds. The evaluated conditions include random selection, pseudo-label-balanced selection, informativeness heuristics, geometry-based methods, augmentation, and an RL-inspired selector.
The results show that no single strategy performs best across both datasets. Augmentation achieves the highest aggregate area under the budget accuracy curve on CIFAR-10. On Fashion-MNIST, k-center achieves the highest value, although uniform random selection is only slightly behind. The RL-inspired selector is competitive on Fashion-MNIST and outperforms pseudo-label-balanced selection in several experimental settings. However, this improvement does not occur consistently across datasets and budgets. Analysis of the selected subsets also shows that high feature diversity alone does not guarantee better classification performance. Distributional alignment and stable class coverage remain important when the selection budget is limited. The main contribution of this work is therefore a reproducible empirical benchmark and an analysis of a hand-weighted RL-inspired selector, rather than a new reinforcement learning algorithm.

Keywords: low-shot learning · budget-aware selection · coreset selection · pseudo-label selection · RL-inspired selection · data efficiency

1 Introduction

Convolutional neural networks require large labelled datasets, yet annotation is expensive in domains such as medical imaging, industrial inspection, and rare-category recognition. With only a few labelled examples per class, models readily overfit and exhibit high variance, motivating low-shot learning.

* Corresponding author.

© The Author(s) 2026
M. Murtadha et al. (eds.), *Proceedings of the 2026 11th International Conference on Sustainable Information Engineering and Technology (SIET 2026)*, Atlantis Highlights in Computer Sciences 26,
https://doi.org/10.2991/978-94-6239-828-3_5

Additional samples can be generated or drawn from a larger pool [6,7,8], but their value is uneven. Redundant or low-quality data may weaken rather than improve the training signal [1,3,2]. Under a fixed annotation and training budget, the practical question is therefore which candidates should be selected.

Existing work studies augmentation, informativeness, diversity, and reinforcement-learning acquisition, but comparisons often vary in their splits, models, or evaluation protocols. We address this gap with four research questions: how much candidate data improves the low-shot baseline; which methods perform best across datasets, shots, and budgets; how selected-set diversity, balance, and alignment relate to performance; and whether a hand-weighted RL-inspired selector justifies its additional cost.

We compare ten experimental conditions under identical data splits, models, budgets, and seeds. The benchmark includes reference conditions, random strategies, informativeness and geometry heuristics, augmentation, and a hand-weighted RL-inspired selector. Performance is evaluated using predictive quality, data efficiency, selected-set properties, computational cost, and statistical uncertainty.

The contributions are a reproducible budget-aware benchmark, a multi-metric comparison of conventional selectors, and a transparent evaluation of a fixed RL-inspired scoring rule. The results identify where simple Random and k-center selection remain competitive and clarify that the evaluated RL-inspired method is not a trained policy.

2 Related Work

Recent augmentation methods use GANs and diffusion models to enlarge training sets, although sample quality and diversity remain important for downstream generalisation [6,7,**?**,1]. Coreset and pruning studies similarly show that importance-based selection can fail when it neglects coverage [3,9,10,11].

Active learning selects informative candidates using uncertainty, diversity, or hybrid acquisition criteria. Prior work shows that no single strategy consistently dominates across experimental settings [12]. Reinforcement learning has also been used to formulate sequential acquisition and curriculum selection [13,16,17].

Our contribution is a controlled benchmark rather than a new selection algorithm. All methods use identical data splits, candidate pools, models, budgets, and random seeds. The evaluation combines predictive metrics, budget efficiency, selected-set properties, computational cost, and statistical uncertainty. The RL-inspired selector is included as a transparent fixed-score baseline rather than a trained reinforcement-learning policy.

3 Problem Formulation

3.1 Budget-Constrained Selection

Let $D_{\text{real}} = \{(x_i, y_i)\}_{i=1}^{N}$ denote the labelled low-shot training set and let $D_{\text{pool}} = \{x'_j\}_{j=1}^{M}$ denote an initially unlabelled candidate pool. Ground-truth candidate

labels are hidden from every selection method and are revealed only after the subset has been selected, thereby simulating annotation under a fixed budget.

For a budget ratio $b \in (0, 1]$, the selector chooses an index subset $I_S \subseteq \{1, \ldots, M\}$ satisfying

$$|I_S| = \text{round}(bM). \tag{1}$$

After selection, the labels of the selected candidates are revealed and the downstream classifier is trained on $D_{\text{real}} \cup D_S$. The test set is never used for selection, checkpoint selection, or hyperparameter configuration.

The benchmark pool is constructed with equal numbers of samples per ground-truth class to provide a controlled source pool. These labels are used only during benchmark construction and post-selection evaluation; they are not exposed to the selectors. Methods requiring class assignments use pseudo-labels predicted by the low-shot base classifier.

4 Methods: Selection Baselines

We evaluate ten experimental conditions: two budget-independent references, seven conventional budget-aware baselines, and one hand-weighted RL-inspired selector. The references are No Candidate and All Candidate. The conventional baselines are Random, Pseudo-Label-Balanced Random, Confidence, Uncertainty, k-center, Feature-distance, and Augmentation. The final condition is the hand-weighted RL-inspired selector. All methods operate on the same candidate pool and feature representation extracted from a classifier trained only on the low-shot labelled data. Augmentation uses the same random budget-constrained subset as the Random baseline, but applies RandAugment, MixUp, and CutMix during downstream training; its selected-set metrics are omitted from later analysis because its distinguishing component is the training procedure rather than the subset-selection rule.

Reference baselines. No Candidate trains only on D_{real}. **All Candidate** uses the full candidate pool as a budget-independent upper reference.

Random baselines. Random samples uniformly without replacement. **Pseudo-Label-Balanced Random** assigns every candidate a pseudo-label using the low-shot base classifier and then samples approximately equal numbers from each predicted class. Ground-truth candidate labels are not used during selection. Consequently, predicted-class allocation is balanced in the pseudo-label space while the actual class distribution revealed after annotation can still vary slightly from uniformity.

Informativeness-based heuristics. Confidence selects candidates with the highest maximum-softmax probability. **Uncertainty** selects candidates with the highest predictive entropy.

Geometry-based heuristics. k-center performs greedy farthest-first selection in embedding space. **Feature-distance** selects candidates whose embeddings are closest to the centroid of their predicted class.

Strong augmentation. Augmentation applies RandAugment with MixUp and CutMix to the budgeted training set.

4.1 Hand-Weighted RL-Inspired Selector

The evaluated RL-inspired selector is not a trained reinforcement- learning policy. It contains no policy network, training episodes, policy-gradient updates, or reward-driven parameter optimisation. Instead, it applies a fixed scoring rule inspired by a contextual accept-or-reject formulation.

For candidate j, the normalised predictive entropy is

$$\widetilde{H}_j = -\frac{\sum_{c=1}^{C} p_{jc} \log(p_{jc} + \epsilon)}{\log C}. \tag{2}$$

The Euclidean distance between the candidate embedding and the centroid of its predicted class is min–max normalised to $\widetilde{d}_j$. The fixed base score is

$$q_j = 0.5\widetilde{H}_j + 0.5(1 - \widetilde{d}_j). \tag{3}$$

During greedy selection, a predicted-class balance penalty is applied:

$$a_j = q_j - 0.7P_{\text{bal}}(j). \tag{4}$$

A candidate is accepted when $a_j > 0.15$ or when its predicted class has not reached the target allocation. If the budget remains unfilled, candidates are added in descending order of q_j.

All coefficients are fixed across datasets, shot settings, budgets, and seeds. The results therefore evaluate a hand-weighted RL-inspired scoring rule rather than a learned RL policy.

5 Experimental Setup

5.1 Dataset and Split Protocol

We use the official training and test partitions of CIFAR-10 and Fashion-MNIST. For every dataset, class, and seed, the official training partition is shuffled once. The first 100 samples per class form the maximum low-shot source, the next 100 samples per class form the validation set, and the following 500 samples per class form the candidate pool.

The 10-, 50-, and 100-shot training sets are nested subsets of the same maximum low-shot source. The validation set and candidate pool remain fixed across shot settings for a matched dataset and seed. The low-shot set, validation set, candidate pool, and official test set are mutually disjoint. All split indices are saved.

Fashion-MNIST images are resized to 32×32. CIFAR-10 remains three-channel and Fashion-MNIST remains single-channel. Input values are linearly scaled to $[-1, 1]$.

5.2 Architecture and Training Configuration

The shared CNN consists of two 3×3 convolutional layers with 32 channels followed by max pooling, two 3×3 convolutional layers with 64 channels followed by a second max-pooling layer, and one 3×3 convolutional layer with 128 channels followed by adaptive average pooling. Each convolution is followed by batch normalisation and ReLU. The pooled representation is projected to a 128-dimensional embedding, followed by ReLU and a linear 10-class classification head.

All models are trained for 30 epochs using Adam with learning rate 10^{-3}, weight decay 5×10^{-4}, batch size 128, cross-entropy loss, and cosine annealing. The checkpoint with the lowest validation loss is retained. We use random seeds 42, 123, and 256 for all methods.

5.3 Evaluation Protocol

We report test accuracy, macro-F1, area under the budget–accuracy curve (AUBC), selected-set diversity, actual class-balance entropy, proxy distributional distance, and wall-clock selection and training cost. Results are aggregated across all three seeds. Ground-truth candidate labels are used only after selection for annotation and selected-set evaluation.

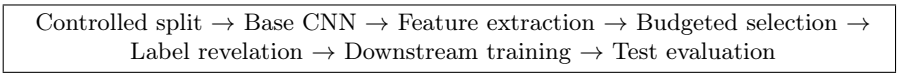


Fig. 1. Experimental workflow. Candidate labels remain hidden until selection is complete.

6 Results

6.1 Accuracy and Macro-F1

For each seed, results are first averaged across the three shot settings. We then report the mean across seeds 42, 123, and 256; statistical uncertainty is retained in Table 6. Tables 1 and 2 show that candidate data substantially improves over the No Candidate reference. Random is strongest at the 25% budget on both datasets, while Augmentation leads CIFAR-10 and k-center leads Fashion-MNIST at 50%. At full budget, non-augmentation selectors use the identical candidate set.

6.2 Budget-Accuracy Curve and Data Efficiency

To summarise efficiency across budgets we report the normalised area under the budget-accuracy curve (AUBC) in Table 3; the RL-inspired selector's per-shot

Table 1. Aggregate CIFAR-10 performance across shot settings. Each entry reports accuracy (%) / macro-F1.

Method	25%	50%	100%
No Candidate	37.37 / .3470	37.37 / .3470	37.37 / .3470
All Candidate	64.64 / .6444	64.64 / .6444	64.64 / .6444
Random	**56.09** / **.5613**	60.89 / **.6079**	64.64 / .6444
Pseudo-label bal.	55.57 / .5552	60.34 / .6035	64.64 / .6444
Confidence	51.46 / .5088	59.04 / .5893	64.64 / .6444
Uncertainty	54.05 / .5399	58.61 / .5870	64.64 / .6444
k-center	55.14 / .5518	59.90 / .5985	64.64 / .6444
Feature-distance	52.52 / .5200	58.45 / .5817	64.64 / .6444
Augmentation	55.45 / .5442	**61.34** / .6059	**67.31** / **.6684**
RL-inspired	52.45 / .5199	59.20 / .5892	64.64 / .6444

Table 2. Aggregate Fashion-MNIST performance across shot settings. Each entry reports accuracy (%) / macro-F1.

Method	25%	50%	100%
No Candidate	62.28 / .6036	62.28 / .6036	62.28 / .6036
All Candidate	88.08 / .8810	88.08 / .8810	88.08 / .8810
Random	**84.78** / **.8479**	86.70 / .8671	88.08 / **.8810**
Pseudo-label bal.	84.04 / .8409	86.23 / .8621	88.08 / .8810
Confidence	79.08 / .7897	82.97 / .8299	88.08 / .8810
Uncertainty	80.65 / .8005	84.87 / .8474	88.08 / .8810
k-center	84.45 / .8437	**86.89** / **.8690**	88.08 / .8810
Feature-distance	79.68 / .7957	84.66 / .8469	88.08 / .8810
Augmentation	81.11 / .7981	84.82 / .8418	**88.18** / .8805
RL-inspired	83.44 / .8343	86.75 / .8677	88.08 / .8810

budget-accuracy breakdown is reported later in the per-shot analysis. Among the budgeted conditions, Augmentation achieves the highest aggregate AUBC on CIFAR-10 (0.6234), followed by Random (0.6134) and Pseudo-label-balanced selection (0.6098). On Fashion-MNIST, k-center achieves the highest AUBC (0.8688), although its advantage over Random (0.8684) is very small. The RL-inspired selector reaches 0.5989 on CIFAR-10 and 0.8664 on Fashion-MNIST. It is therefore relatively weak on CIFAR-10 but competitive with the strongest selectors on Fashion-MNIST. These results indicate that no selection strategy dominates across datasets and that simple Random selection remains a strong baseline.

6.3 Selected Data Analysis

To understand *why* the methods rank as they do, we examine the composition of the selected subsets. Table 4 reports diversity, class balance, and proxy distributional distance at a representative operating point (CIFAR-10, 50 shots, 25% budget), averaged over three seeds. Augmentation is omitted from this analysis because its selected indices follow the Random baseline and its distinguishing component is the use of input-space transformations during downstream training. Confidence selection produces the highest diversity (11.85), but it also has

Table 3. Normalised AUBC, averaged across three shot settings and three random seeds. Higher values indicate greater data efficiency.

Method	CIFAR-10	F-MNIST
No Candidate	0.3737	0.6228
All Candidate	0.6464	0.8808
Random	0.6134	0.8684
Pseudo-label bal.	0.6098	0.8648
Confidence	0.5964	0.8402
Uncertainty	0.5986	0.8524
k-center (div.)	0.6069	**0.8688**
Feature-distance	0.5952	0.8497
Augmentation	**0.6234**	0.8532
RL-inspired	0.5989	0.8664

relatively weak class balance (0.954) and a large distributional distance (1.92). Thus, high pairwise diversity alone does not guarantee strong downstream performance. Random and Pseudo-label-balanced selection provide the closest distributional alignment with the real low-shot data, with distances of 0.36 and 0.26, respectively. Random obtains the highest actual-label balance (0.999), closely followed by Pseudo-label-balanced selection (0.998). This distinction matters because the latter enforces balance in pseudo-label space rather than using hidden ground-truth candidate labels. The RL-inspired selector has the lowest diversity (5.03), lowest class balance (0.937), and largest distributional distance (2.31), showing that its fixed soft penalties do not guarantee representative coverage.

Table 4. Selected-set properties at CIFAR-10, 50-shot, and 25% budget, averaged across three seeds. Augmentation is omitted because its distinguishing component is the downstream training transformation. Div.: mean pairwise feature distance; Bal.: actual-label class-balance entropy (1.0 = uniform); Dist.: proxy distance to the real low-shot data.

Method	Div.	Bal.	Dist.
Random	9.09	**0.999**	0.36
Pseudo-label bal.	9.11	0.998	**0.26**
Confidence	**11.85**	0.954	1.92
Uncertainty	5.20	0.989	1.14
k-center (div.)	10.27	0.993	1.22
Feature-distance	8.52	0.988	1.11
RL-inspired	5.03	0.937	2.31

6.4 RL-Inspired Results

The RL-inspired selector shows a dataset-dependent pattern. On CIFAR-10, it remains below the strongest baselines in the aggregate comparison. On Fashion-MNIST, it outperforms Pseudo-label-balanced selection in the 50-shot and 100-shot settings at both the 25% and 50% budgets. Its strongest relative result occurs in the Fashion-MNIST 100-shot setting at the 25% budget, where it reaches 86.19%. These findings indicate that the fixed scoring rule can be competitive when the base representation is sufficiently stable, but its advantage does not generalise across datasets.

6.5 Cost Analysis

Table 5. Average computational cost in minutes. Selection time excludes the shared base-model training and feature-extraction stages.

	CIFAR-10			F-MNIST		
Method	Sel.	Train	Total	Sel.	Train	Total
Random	0.000012	0.2225	0.2225	0.000012	0.2194	0.2194
Pseudo-label	0.000019	0.2225	0.2225	0.000019	0.2198	0.2198
Confidence	0.000009	0.2225	0.2225	0.000009	0.2199	0.2199
Uncertainty	0.000008	0.2227	0.2227	0.000008	0.2198	0.2198
k-center	0.020637	0.2230	0.2436	0.019408	0.2200	0.2394
Feature-dist.	0.000069	0.2215	0.2216	0.000065	0.2181	0.2182
Augmentation	0.000011	1.4227	1.4227	0.000012	1.2000	1.2000
RL-inspired	0.001057	0.2224	0.2235	0.001069	0.2195	0.2206

Augmentation has the largest total cost because it increases the complexity of downstream training. Among the subset-selection methods, k-center has the highest selection overhead because of its iterative farthest-point computation. The additional cost of the RL-inspired scoring rule remains small relative to classifier training.

6.6 Statistical Analysis

We report the mean, standard deviation, and 95% confidence interval across three random seeds. Paired comparisons use matched shot, budget, and seed configurations and are conducted separately for each dataset. Because the experiments use three seeds, the confidence intervals are interpreted as measures of experimental uncertainty rather than evidence of universal method superiority.

Table 6. Paired accuracy comparisons against Random selection. Tests use matched shot, budget, and seed configurations. Adjusted p-values use Holm correction.

Dataset	Comparison	n	Mean diff.	p	Adjusted p
CIFAR-10	Pseudo-label balanced vs Random	36	0.0004	0.8516	0.8516
CIFAR-10	k-center vs Random	36	−0.0048	0.1675	0.3350
CIFAR-10	RL-inspired vs Random	36	−0.0148	0.0005	0.0015
Fashion-MNIST	Pseudo-label balanced vs Random	36	−0.0050	0.0030	0.0091
Fashion-MNIST	k-center vs Random	36	−0.0016	0.3291	0.3291
Fashion-MNIST	RL-inspired vs Random	36	−0.0102	0.0804	0.1608

7 Discussion

The results do not identify a universally dominant strategy. Random selection remains a strong baseline at limited budgets. Augmentation provides the highest aggregate efficiency on CIFAR-10, but its training cost is substantially larger. K-center obtains the highest Fashion-MNIST AUBC by a very small margin over Random. The per-shot results also show that Uncertainty becomes more effective as the initial labelled set grows, indicating that acquisition quality depends on the reliability of the base classifier.

Selected-set analysis shows that diversity alone does not guarantee strong downstream performance. Confidence produces the most diverse subset but also exhibits weak class balance and poor distributional alignment. The RL-inspired selector is competitive on Fashion-MNIST but not on CIFAR-10. Its fixed weights therefore do not provide a dataset-independent selection rule.

8 Limitations and Threats to Validity

The benchmark covers two relatively small datasets, a compact CNN, and a controlled class-balanced candidate pool. Conclusions may change for larger images, deeper backbones, noisy labels, or naturally imbalanced deployment data. Three seeds support matched comparisons, although additional repetitions would improve statistical precision. The RL-inspired selector is a fixed scoring rule and cannot support claims about policy learning or reward optimisation. Future work should evaluate learned selector weights, controlled component ablations, larger datasets, deeper architectures, and naturally imbalanced candidate pools.

9 Conclusion

We presented a controlled benchmark of budget-aware data selection for low-shot image classification. Adding candidate data provides a substantial improvement over training only on the initial labelled set, but no selection strategy performs best across all conditions.

Augmentation achieves the highest aggregate AUBC on CIFAR-10, while k-center achieves the highest Fashion-MNIST AUBC by a narrow margin over Random. Diversity alone does not explain downstream performance; class coverage and distributional alignment are also important.

The RL-inspired selector is a hand-weighted scoring rule rather than a trained policy. It is competitive in several Fashion-MNIST settings, but its benefit does not generalise consistently across datasets. These findings position the work as a controlled empirical benchmark rather than a new reinforcement-learning algorithm.

References

1. A. Havrilla et al., "Surveying the Effects of Quality, Diversity, and Complexity in Synthetic Data from Large Language Models," *arXiv:2412.02980*, 2024.
2. I. Shumailov et al., "Model Collapse Revisited: From Census to Fine-Tuning," *NeurIPS*, 2024.
3. H. Zheng, R. Liu, F. Lai, and A. Prakash, "Coverage-Centric Coreset Selection for High Pruning Rates," *ICLR*, 2023.
4. J. Li et al., "A Survey of Mix-based Data Augmentation: Taxonomy, Methods, Applications, and Limitations," *arXiv:2212.10888*, 2022.
5. M. Zhu et al., "A Survey on Mixup Augmentations and Beyond," *arXiv:2409.05202*, 2024.
6. J. M. Rozanec et al., "Synthetic Data Augmentation Using GAN for Improved Automated Visual Inspection," *arXiv:2212.09317*, 2022.
7. S. Azizi et al., "Synthetic Data from Diffusion Models Improves ImageNet Classification," *CVPR*, 2023.
8. B. Trabucco et al., "Effective Data Augmentation with Diffusion Models," *arXiv:2302.07944*, 2023.
9. B. Mirzasoleiman et al., "Towards Sustainable Learning: Coresets for Data-Efficient Deep Learning," *arXiv:2306.01244*, 2023.
10. M. Maharana et al., "D2 Pruning: Message Passing for Balancing Diversity and Difficulty in Data Pruning," *arXiv:2310.07931*, 2023.
11. Z. Lin et al., "ELFS: Label-Free Coreset Selection with Proxy Training Dynamics," *arXiv:2406.04273*, 2024.
12. C. Lüth et al., "A Survey on Deep Active Learning: Recent Advances and New Frontiers," *arXiv:2405.00334*, 2024.
13. H. Sahbi et al., "Image Classification with Deep Reinforcement Active Learning," *arXiv:2412.19877*, 2024.
14. Y. Kwon and J. Zou, "Beta Shapley: A Unified and Noise-Reduced Data Valuation Framework for Machine Learning," *AISTATS*, 2022.
15. L. Watson et al., "Accelerated Shapley Value Approximation for Data Evaluation," *arXiv:2311.05346*, 2023.
16. Z. Dong et al., "Curriculum Graph Machine Learning: A Survey," *arXiv:2302.02926*, 2023.
17. J. Ren et al., "RL-Selector: Reinforcement Learning-Guided Data Selection via Redundancy Assessment," *arXiv:2506.21037*, 2025.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

