

PRIVACY-ROBUSTNESS SENSITIVITY ANALYSIS FOR SECURE DEPLOYMENT OF DIFFERENTIALLY PRIVATE DEEP LEARNING UNDER DISTRIBUTION SHIFT

B. Lavanya, S. Janani*

Department of Computer Science, University of Madras, Guindy, Chennai
India

* Corresponding Author, e-mail: lavanyaunom@gmail.com

Abstract: The widespread use of deep learning on sensitive data has positioned Differential Privacy (DP) as a central mechanism for protecting training confidentiality. Existing evaluations emphasize the privacy-utility trade-off measured on clean data, overlooking robustness under distribution shift. This study introduces Privacy-Robustness Sensitivity (PRS), a metric that explicitly quantifies the degradation in robustness induced by differential privacy by measuring the relative loss in corruption robustness with respect to clean accuracy as privacy constraints tighten. A privacy-utility-robustness evaluation framework is validated on CIFAR-10, CIFAR-10-C, CelebA, and a systematically constructed CelebA-C biometric corruption benchmark, enabling robustness-aware evaluation of privacy-preserving deep learning models for secure and resilient deployment scenarios.

Key words: Differential privacy, distribution shift robustness, three-dimensional evaluation, privacy-robustness sensitivity (PRS), privacy-preserving deep learning.

1. INTRODUCTION

In recent decades, artificial intelligence has experienced substantial progress, largely driven by advances in deep learning techniques that have enabled significant breakthroughs across a wide range of applications, with notable developments in architectures and models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs), capsule networks, graph neural networks (GNNs) and transformers [1]. However, large-scale models may memorise sensitive training samples, exposing them to privacy attacks [2]. Differential Privacy (DP) [3] provides a rigorous mathematical framework that prevents the influence of any single data record on the output of a model. In deep learning, this is typically enforced through Differentially Private Stochastic Gradient Descent (DP-SGD). While DP enhances confidentiality, it introduces a privacy-utility trade-off, often reducing predictive accuracy [4]. Existing evaluations of differentially private deep learning models primarily focus on clean accuracy, defined as performance on the original, uncorrupted data distribution, overlooking robustness under distribution shifts caused by noise, blur, and digital distortions. To address this gap, this work extends the privacy-utility perspective to a privacy-utility-robustness trade-off and proposes the

Privacy-Robustness Sensitivity (PRS) metric to analyse privacy-induced robustness variation under corruption. The main contributions of this work are as follows:

- Privacy-Utility-Robustness Framework: A structured three-dimensional evaluation methodology that jointly quantifies privacy guarantees, clean-data utility, and corruption robustness under distribution shift for deployment-oriented assessment of differentially private systems.
- Privacy-Robustness Sensitivity (PRS) Metric: A normalised sensitivity metric that measures the relative robustness degradation induced by tightening privacy budgets, enabling direct comparison across models, datasets, and privacy regimes.
- Systematic Corruption Benchmarking and Empirical Analysis: A Comprehensive validation on CIFAR-10, CIFAR-10-C, CelebA, and a systematically constructed CelebA-C benchmark, revealing non-monotonic behaviour and dataset-dependent stability patterns beyond conventional utility-centric evaluation.

Together, these contributions position robustness sensitivity as a first-class evaluation dimension for DP deep learning beyond accuracy-centric assessments. The remainder of the article is organised as follows: Section 2 provides a comprehensive review of the literature, Section 3 describes the experimental framework, Section 4 details the implementation of the proposed method, Section 5 analyses the experimental results, and Section 6 concludes the study and outlines directions for future work.

2. RELATED WORKS

Existing research on differential privacy (DP) focuses on how privacy assurances affect model performance and data usability through the privacy–utility trade-off [5]. Frameworks such as DP-UTIL [6] emphasise the conflict between formal privacy assurances and practical utility, making it difficult to choose privacy budgets for real-world deployment. Empirical analysis shows that utility degradation depends on both the perturbation mechanism and the privacy budget; for instance, output perturbation can cause significant utility loss exceeding 0.6 at lower privacy budgets ($\epsilon < 1$), while other mechanisms have lower degradation but higher privacy leakage. This demonstrates that increases in utility generally increase the vulnerability of inference attack, strengthening the accuracy-privacy trade-off.

Recent research expands this trade-off to multi-objective privacy–utility–explainability models [7] for image data processing, indicating that DP influences accuracy and model interpretability, adding decision-making restrictions. Federated Learning with Local Differential Privacy [8] examines communication overhead, privacy assurances, and model performance in distributed environments. Privacy-preserving digital twin technology for smart cities [9] shows how large-scale smart city deployments must balance privacy and utility. However, these studies largely assess utility retention and privacy guarantees without systematically examining robustness under distribution shift.

2.1. Research Gap

Despite advances in differential privacy for deep learning, existing research predominantly focuses on in-distribution settings, overlooking real-world scenarios.

Consequently, the impact of differential privacy on robustness under corruption remains insufficiently explored. In particular, it is unclear whether privacy-induced noise affects all corruption types uniformly or exhibits corruption-dependent behaviour. Moreover, current metrics rely only on absolute performance values. This limits consistent comparison across datasets and deployment conditions. To address these gaps, this work introduces a unified evaluation framework and the PRS metric to quantify the degradation of robustness under privacy constraints.

3. EXPERIMENTAL FRAMEWORK

This paper proposes a Privacy–Utility–Robustness Evaluation Framework to address the research gap. Figure 1 illustrates its workflow, which consists of four sequential stages: dataset preparation, differentially private training, evaluation under distribution shift, and PRS-based analysis.

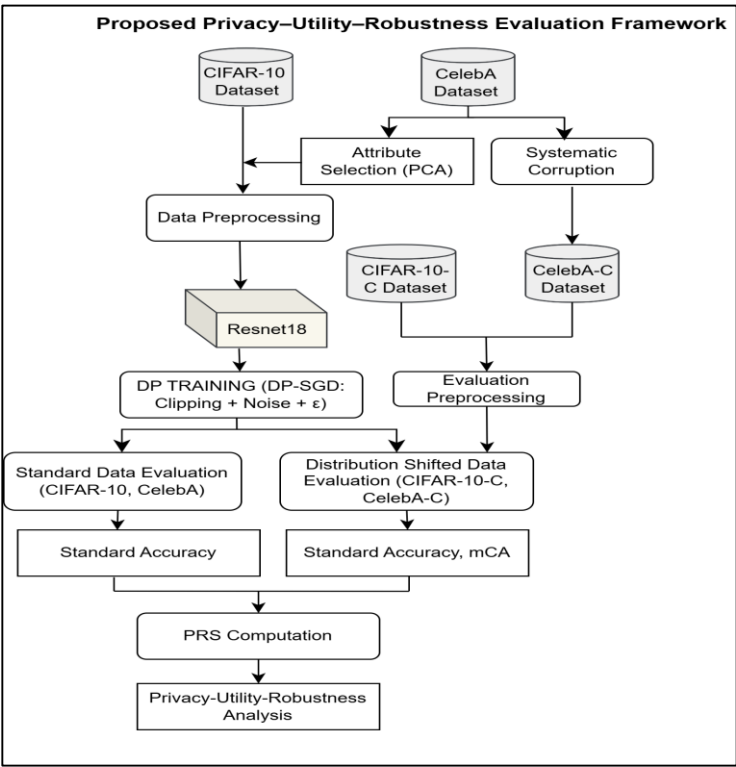


Figure 1. The proposed privacy–utility–robustness evaluation framework

3.1. Dataset Selection and Construction for Robustness Evaluation

The proposed framework is operationalised by incorporating both object-centric and privacy-sensitive datasets, as well as their corrupted counterparts, to simulate distribution shift conditions. The selection of datasets is guided by the need to evaluate differential privacy under both controlled benchmarks and privacy-sensitive real-world scenarios, while systematically analysing robustness under distribution shift. The

CIFAR-10 dataset is used as a baseline benchmark for standard image classification. To assess robustness under controlled distribution changes across various corruption types and severity levels, the corrupted form, CIFAR-10-C, is utilised, yielding a tensor of shape $(10,000 \text{ test images} \times 5 \text{ severities} \times 32 \times 32 \times 3 \text{ channels})$.

The CelebA dataset is used to represent privacy-sensitive information pertaining to facial characteristics. Prior to training, Linear PCA using singular value decomposition (SVD) is applied for attribute selection to minimise redundancy and guarantee semantic distinctness. CelebA-C is a systematically corrupted extension, created by applying 19 types of corruption to the test set at five different severity levels, forming a structured corruption tensor of shape $(19,962 \times 5 \times 218 \times 178 \times 3)$. To facilitate reproducibility and further research, the dataset is made publicly available on Zenodo under the CC BY 4.0 license and distributed as an HDF5 file with lossless gzip compression. For practical hosting and modular access, the dataset is released as four complementary subsets corresponding to noise, blur, weather, and digital corruption families [10–13]. For qualitative illustration, Figure 2 presents representative examples of the corruption types applied in CelebA-C.

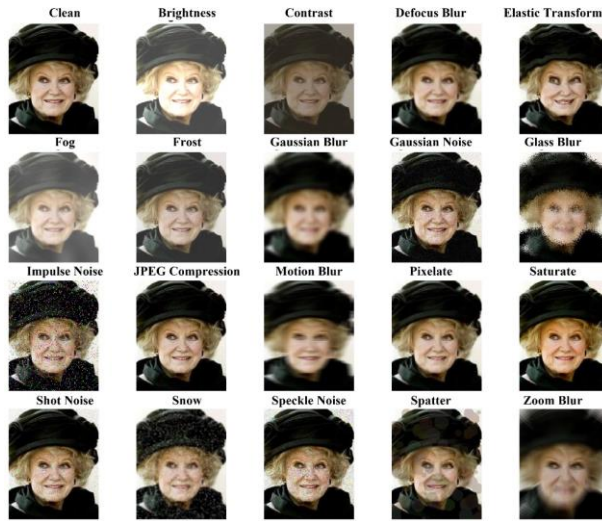


Figure 2. Visualisation of 19 corruption types from CelebA-C applied to reference CelebA test images for robustness evaluation under distribution shift.

4. IMPLEMENTATION OF PROPOSED METHOD

Differential Privacy is implemented using DP-SGD during the training phase of Resnet18 on both CIFAR-10 and CelebA datasets. In this approach, the input data remains unchanged, while privacy is enforced by perturbing gradients during optimization. The clipped gradient update is computed as:

$$\tilde{g} = \frac{1}{B} \left(\sum_{i=1}^B \text{Clip}(g_i, C) \right) + N(0, \sigma^2 C^2 I) \quad (1)$$

where g_i denotes the gradient corresponding to the i -th training sample, B denotes batch size, C is the clipping norm that limits the influence of individual samples, and

$\mathcal{N}(0, \sigma^2 C^2 I)$ represents Gaussian noise scaled by the noise multiplier σ to ensure formal privacy guarantees. The resulting noisy gradient $\tilde{g}$ is used to update model parameters during training. Figure 3 provides a schematic representation of the DP-SGD gradient perturbation workflow.

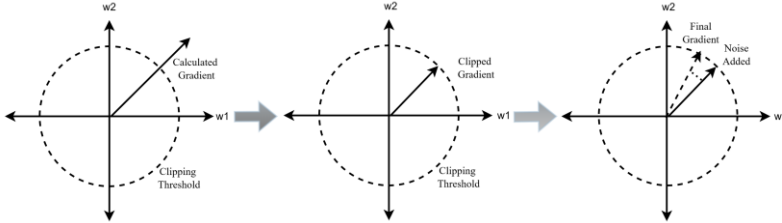


Figure 3. Gradient perturbation in DP-SGD showing per-sample gradient computation, clipping to bound sensitivity, and noise addition for privacy preservation

4.1. Non-Differentially Private (Non-DP) Training

In the non-DP training regime, both CIFAR-10 and CelebA experiments employ a ResNet-18 architecture trained for 20 epochs using the RMSprop optimiser with a fixed learning rate of 1×10^{-3} . CIFAR-10 follows a standard 83.33% -16.67% training-test split with a batch size of 512, while CelebA follows the official partition with approximately 80.34% training, 9.81% validation, and 9.85% test samples. In this work, only the CelebA training and test splits are utilised, with a batch size of 128.

The number of epochs is determined based on preliminary convergence analysis, where both training and validation performance stabilise around epochs 18–20 with negligible subsequent improvement. For CIFAR-10, the task is formulated as multi-class classification using CrossEntropy loss, whereas CelebA is treated as a multi-label classification problem using BCEWithLogitsLoss for the selected attributes. Selecting a single, well-established backbone ensures methodological consistency across experiments, allowing observed robustness variations to be attributed to privacy budgets and corruption shifts rather than architectural differences.

4.2. Differentially private (DP) training

In this work, differentially private training is implemented using DP-SGD in the PyTorch framework to train the ResNet-18 model. Under this configuration, a gradient clipping norm of $C = 1.2$ and a privacy target of $\delta = 10^{-5}$ are adopted across all experiments, while varying the noise multiplier ($\sigma = 0.5, 0.7, 0.8, 1.0$). Minor implementation-level modifications are applied to ResNet-18 by replacing in-place operations with out-of-place equivalents, thereby avoiding interference with Opacus gradient hooks. Model utility is evaluated using top-1 classification accuracy on the clean CIFAR-10 test set. Accuracy is defined as:

$$Accuracy = \frac{1}{N} \sum_{i=1}^N \mathbb{I} \left(\arg \max_j f(x_i)_j = y_i \right) \quad (2)$$

where $f(x_i)$ denotes the model output logits and y_i is the ground-truth label of N samples.

For the CelebA dataset, performance is evaluated in a multi-label facial attribute classification setting, where the model is trained to jointly predict three binary attributes using a sigmoid-based output formulation optimised with BCEWithLogitsLoss. Formally, exact-match accuracy is defined as:

$$Acc_{exact} = \frac{1}{N} \sum_{i=1}^N \mathbb{I} \left(\bigwedge_{k=1}^3 \hat{y}_{i,k} = y_{i,k} \right) \quad (3)$$

where $\hat{y}_{i,k}$ and $y_{i,k}$ denote the predicted and ground-truth labels of N samples for the attribute $k \in \{\text{Gender, Young, Smiling}\}$.

4.3. Distribution Shift and Corruption Modelling

Robustness is evaluated under distribution shift by testing the model on corrupted inputs while keeping the training distribution unchanged, thereby isolating corruption-induced effects. For a given corruption type c and severity level s , Robustness accuracy is defined as

$$Acc_{c,s} = \frac{1}{N_c} \sum_{i=1}^{N_c} \mathbb{I} \left(\arg \max_j f(x_i^{c,s})_j = y_i \right) \quad (4)$$

where $x_i^{c,s}$ denotes the corrupted input, y_i is the ground-truth label, and $\mathbb{I}(\cdot)$ is the indicator function.

4.4. The Proposed Privacy–Robustness Sensitivity (PRS) Metric

To explicitly quantify how Differential Privacy affects robustness under distribution shift, the Privacy–Robustness Sensitivity (PRS) is introduced as a sensitivity-based metric that quantifies the relative degradation in robustness induced by privacy constraints. Unlike conventional robustness metrics that report absolute accuracy under corruption, PRS captures how sensitive corruption robustness is to changes in privacy, normalised to the clean-data utility. To systematically characterise corruption robustness, mean Corruption Accuracy (mCA) is defined as the average model accuracy across all corruption types and severity levels under the same privacy budget:

$$mCA = \frac{1}{|C||S|} \sum_{c \in C} \sum_{s \in S} Acc_{c,s}(\epsilon) \quad (5)$$

where $Acc_{c,s}(\epsilon)$ denotes the classification accuracy under the corruption type c and severity level s at privacy budget ϵ . In this work, C consists of 19 corruption types and S includes five severity levels, ensuring comprehensive evaluation under diverse distribution shifts. Building upon this aggregated robustness formulation, PRS is proposed as:

$$PRS(\epsilon) = \frac{A_{clean}(\epsilon) - mCA}{A_{clean}(\epsilon)} \quad (6)$$

where $A_{clean}(\epsilon)$ denote the clean test accuracy under the privacy budget ϵ , computed according to Equations (2) and (3). $PRS(\epsilon)$ quantifies the relative loss in robustness compared to clean-data performance. Since $0 \leq mCA(\epsilon) \leq A_{clean}(\epsilon)$ it follows that:

$$0 \leq PRS(\epsilon) \leq 1 \quad (7)$$

This formulation expresses the fractional loss in robustness relative to clean performance, enabling direct comparison across models, datasets, and privacy budget settings. While prior works primarily rely on standard or task-specific performance indicators, the proposed metric serves as an enabling component of a broader privacy-

utility-robustness evaluation framework. PRS = 0 indicates no robustness degradation, while PRS = 1 indicates severe robustness loss under privacy.

5. RESULTS AND ANALYSIS

This section presents a comprehensive analysis of the experimental results, examining the impact of Differential Privacy on model utility and robustness under distribution shift.

5.1. Privacy-Utility Analysis

An initial analysis examines the privacy–utility trade-off on clean test data in both non-DP and DP settings. Notably, performance is stronger on the CelebA, as increased data scale enables tighter privacy guarantees, whereas CIFAR-10 requires higher ϵ under the same noise settings, underscoring the role of dataset size in mitigating privacy-induced utility loss. Figure 4 depicts the privacy–utility trade-off for CIFAR-10 and CelebA, respectively.

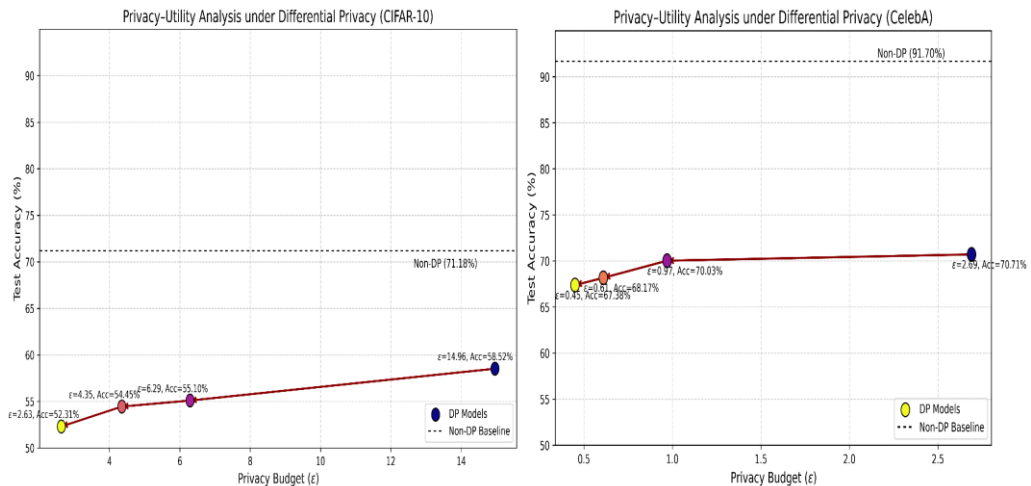


Figure 4. Privacy–Utility trade-off on CIFAR-10 (left) and CelebA (right). Test accuracy versus privacy budget ϵ for varying noise levels σ ; dashed lines indicate non-DP baselines

5.2. Utility-Robustness Analysis

Utility robustness is quantified using mCA. Figure 5 illustrates the robustness of the non-DP model to 19 corruption types on CIFAR-10-C and CelebA-C. The clean accuracy serves as an upper-performance baseline. In both datasets, accuracy consistently decreases with increasing severity, indicating sensitivity to stronger perturbations. Among the corruptions, noise and blur-based distortions cause the most significant degradation, whereas brightness and contrast have comparatively lower impact. These observations highlight the varying sensitivity of models to different corruption families and establish a baseline for analysing the additional effect of differential privacy on robustness.

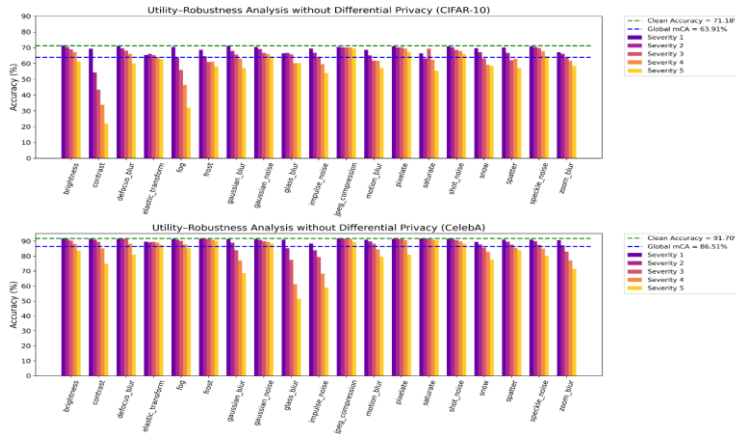


Figure 5. Utility-robustness analysis of the non-DP model on CIFAR-10-C (Top) and CelebA-C (Bottom). Bars represent accuracy across severity levels (1–5), for each corruption type, dashed lines indicating clean accuracy and global mean corruption accuracy (mCA)

5.3. Privacy-Utility-Robustness Analysis

Experiments on CIFAR-10-C and CelebA-C demonstrate a clear privacy-utility-robustness trade-off under Differential Privacy: as the privacy budget ϵ decreases, both clean accuracy and mCA decline, indicating reduced utility and robustness as shown in Figures 6 and 7. The impact of differential privacy is not uniformly detrimental across all corruption types. While most distortions exhibit reduced accuracy under stronger privacy constraints, certain corruption categories, particularly photometric transformations such as brightness, contrast, and saturation, show relatively stable performance due to the implicit regularisation effect introduced by DP-SGD, where gradient noise reduces overfitting to fine-grained pixel patterns and encourages learning of more robust, high-level features.

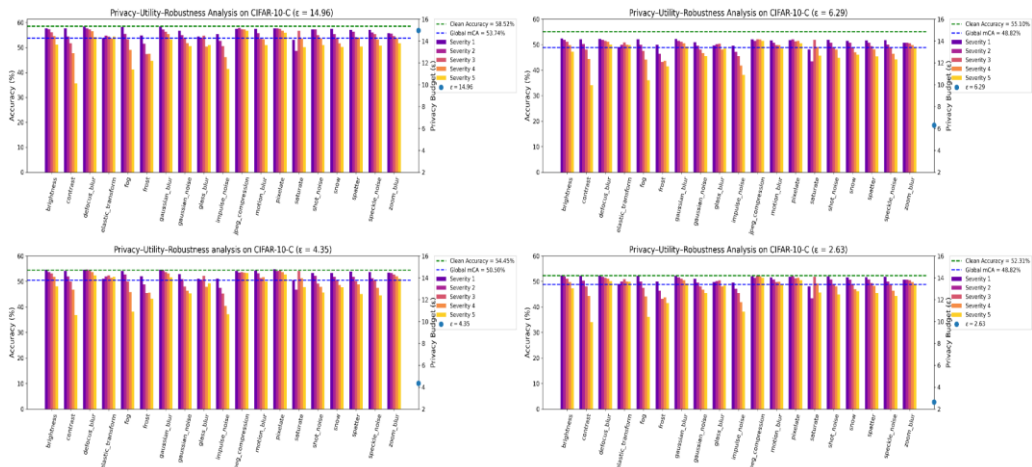


Figure 6. Privacy-utility-robustness analysis of CIFAR-10-C under Differential Privacy across multiple privacy budgets: (i) $\epsilon = 14.96$, (ii) $\epsilon = 6.29$, (iii) $\epsilon = 4.35$, and (iv) $\epsilon = 2.63$

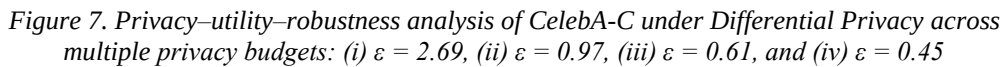


Table 1 summarises the Privacy–Robustness Sensitivity (PRS) values obtained for CIFAR-10 and CelebA under varying privacy budgets.

<i>Dataset</i>	ϵ (Privacy Budget)	Clean Accuracy (%)	mCA (%)	PRS
CIFAR-10	∞ (Non-DP)	71.18	63.91	0.1022
	14.96 ($\sigma = 1.0$)	58.52	53.74	0.0818
	6.29 ($\sigma = 0.8$)	55.10	48.82	0.0730
	4.35 ($\sigma = 0.7$)	54.45	50.50	0.0725
	2.63 ($\sigma = 0.5$)	52.31	48.82	0.0668
Celeba	∞ (Non-DP)	91.70	86.51	0.0566
	2.69 ($\sigma = 1.0$)	70.71	65.33	0.0761
	0.97 ($\sigma = 0.8$)	70.03	64.41	0.0802
	0.61 ($\sigma = 0.7$)	68.17	63.65	0.0663
	0.45 ($\sigma = 0.5$)	67.38	62.86	0.0671

The PRS analysis exhibits a non-monotonic trend with respect to the privacy budget ϵ across both datasets. CIFAR-10 consistently yields higher PRS values than CelebA, reflecting the greater susceptibility of object-centric visual features to common corruptions compared to facial attribute representation. These observations demonstrate that robustness under Differential Privacy cannot be inferred from clean accuracy alone but instead depends on the stability and invariance of learned representations shaped by dataset structure and privacy-induced regularisation.

5.5. Comparative Analysis with Existing DP Frameworks

To position this study within differential privacy research, Table 2 provides a comparative overview of key methodological differences between prior work and the proposed approach.

Table 2. Comparative analysis of prior differential privacy studies and the proposed work

Authors	Dataset	(ϵ , δ)	RDP	3D Analysis	Novel Metric	Corrupted Dataset Construction	Multi-Corruption Robustness
[10]	MNIST, Fashion-MNIST	✓	✓	✗	✓	✗	✗
[8]	CIFAR-10, COVID-19, and loan dataset.	✓	✗	✗	✓	✗	✗
[5]	MNIST, Adult, income and Breast Cancer Dataset.	✓	✗	✗	✗	✗	✗
[9]	MNIST, FER, CIFAR-10.	✓	✓	✓ (Privacy-Utility-Explainability)	✓	✗	✗
The proposed work	CIFAR-10, CIFAR-10-C, CelebA, CelebA-C.	✓	✓	✓ (Privacy-Utility-Robustness)	✓	✓	✓

5.6. Statistical Analysis

To evaluate the effect of Differential Privacy on robustness under distribution shift, paired Wilcoxon signed-rank tests were conducted on corruption-level accuracies for CIFAR-10-C and CelebA-C (Table 3). Given the bounded and non-Gaussian distribution of robustness scores, a non-parametric approach was adopted. Across all privacy budgets, DP models exhibit significantly lower robustness than the non-DP baseline. These results confirm that robustness degradation is systematic and establishes a consistent privacy–robustness trade-off across datasets and budgets.

Table 3. Wilcoxon signed-rank test results for pairwise comparisons between Non-DP and DP models under corrupted conditions

Dataset	Comparison	Median		Median Difference	% Drop	p-value
		Non-DP	DP			
CIFAR-10	Non-DP vs $\epsilon = 14.96$	63.40	54.1	9.3	14.7	$p < 0.001$
	Non-DP vs $\epsilon = 6.29$	63.40	50.0	13.4	21.1	$p < 0.001$
	Non-DP vs $\epsilon = 4.35$	63.40	50.9	12.5	19.7	$p < 0.001$
CIFAR-10-C	Non-DP vs $\epsilon = 2.63$	63.40	48.5	14.9	23.5	$p < 0.001$
CelebA	Non-DP vs $\epsilon = 2.69$	88.71	67.27	20.64	23.27	$p < 0.001$
	Non-DP vs $\epsilon = 0.97$	88.71	66.40	21.76	24.5	$p < 0.001$
CelebA-C	Non-DP vs $\epsilon = 0.61$	88.71	65.01	23.07	26.01	$p < 0.001$
	Non-DP vs $\epsilon = 0.45$	88.71	64.37	23.76	26.78	$p < 0.001$

6. CONCLUSIONS

This work extends privacy-utility evaluation by including robustness to distribution shifts utilising CIFAR-10-C and CelebA-C datasets. The proposed PRS metric measures the deterioration in robustness relative to clean accuracy and facilitates comprehensive investigation across privacy budgets. Results indicate that a reduction in ϵ diminishes both clean accuracy and mCA, thereby affirming a trade-off among privacy, utility, and robustness. The investigation also demonstrates dataset-dependent behaviour, with object-centric features exhibiting greater sensitivity to corruption under privacy conditions. Furthermore, PRS exhibits non-monotonic trends across ϵ , while specific photometric distortions maintain relative stability, indicating that the decline in robustness is not consistent across different types of distortions.

Future work may extend this evaluation framework to additional modalities such as medical imaging and natural language processing, where distribution shifts exhibit distinct structural characteristics. Investigation of adaptive privacy strategies that dynamically balance robustness and privacy constraints may further strengthen the analytical and practical foundations of robustness-aware differential privacy.

ACKNOWLEDGEMENT

This work is supported by the University Grants Commission (UGC, India) under the Junior Research Fellowship scheme (JRF).

REFERENCES

[1] Mienye, I. D., Swart, T. G. A comprehensive review of deep learning: Architectures, recent advances, and applications. *Information*, vol.15, no.12, 2024, 755. DOI: 10.3390/info15120755.

[2] Shafee, A., Awaad, T. A. Privacy attacks against deep learning models and their countermeasures. *Journal of Systems Architecture*, vol.114, 2021, 101940. DOI: 10.1016/j.sysarc.2020.101940.

- [3] Xiao, H., Xiang, Z., Wang, D., Devadas, S. A theory to instruct differentially-private learning via clipping bias reduction. *2023 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, May 2023, pp. 2170–2189. DOI: 10.1109/SP46215.2023.10179409.
- [4] Bari, N. M. S., Imtiaz, H. Empirical analysis of the privacy-utility trade-off of classification algorithms under differential privacy. *2023 26th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, December 2023, pp. 1–6. DOI: 10.1109/ICCIT60459.2023.10441153.
- [5] Seeman, J., Susser, D. Between Privacy and utility: On differential privacy in theory and practice. *ACM Journal on Responsible Computing*, vol.1, no.1, 2024, Article 3, 18 pages. DOI: 10.1145/3626494.
- [6] Jarin, I., Eshete, B. DP-UTIL: Comprehensive utility analysis of differential privacy in machine learning. *Twelfth ACM Conference on Data and Application Security and Privacy (CODASPY '22)*, Baltimore, MD, USA, April 2022, pp. 41–52. DOI: 10.1145/3508398.3511513.
- [7] Abbasi, W., Mori, P., Saracino, A. Trading-off privacy, utility, and explainability in deep learning-based image data analysis. *IEEE Transactions on Dependable and Secure Computing*, vol.22, no.1, 2025, pp. 388–405. DOI: 10.1109/TDSC.2024.3400608.
- [8] Kim, M., Günlü, O., Schaefer, R. F. Federated learning with local differential privacy: Trade-offs between privacy, utility, and communication. *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, Canada, June 6–11, 2021, pp. 2650–2654. DOI: 10.1109/ICASSP39728.2021.9413764.
- [9] Mohamed, M., Mohamed, D., Maatouk, M. M. H. Privacy-preserving digital twin technologies for smart cities: Balancing utility and data protection. *International Journal on Information Technologies and Security*, vol.17, no.4, 2025, pp. 57–68. DOI: 10.59035/CGYJ7627.
- [10] Lavanya, B., Janani, S. CelebA-C: Noise corruptions (Version 1.0) [Dataset]. Zenodo, 2026. DOI: 10.5281/zenodo.18502914.
- [11] Lavanya, B., Janani, S. CelebA-C: Blur corruptions (Version 1.0) [Dataset]. Zenodo, 2026. DOI: 10.5281/zenodo.18502568.
- [12] Lavanya, B., Janani, S. CelebA-C: Weather corruptions (Version 1.0) [Dataset]. Zenodo, 2026. DOI: 10.5281/zenodo.18503116.
- [13] Lavanya, B., Janani, S. CelebA-C: Digital corruptions (Version 1.0) [Dataset]. Zenodo, 2026. DOI: 10.5281/zenodo.18503029.

Information about the authors:

Dr B. Lavanya is an Associate Professor at the Department of Computer Science at the University of Madras, located in Guindy, Chennai, India. Her areas of research interest include data mining, bioinformatics, big data analytics, data security, and natural language processing.

S. Janani is pursuing her PhD with a UGC JRF fellowship at the Department of Computer Science at the University of Madras, Chennai, India. She is an enthusiastic researcher with research interests in machine learning, data privacy, and secure data-driven systems.

Manuscript received on 06 March 2026