A Appendix

A.1 Proof of Theorem 1

As illustrated in Sec. 3.2, it is hard to build the unlearned data x^u for the feature unlearning since adding the perturbation may influence the model accuracy seriously. Suppose the feature is successfully removed when the norm of perturbation is larger than C. We define the utility loss ℓ_1 with unlearning feature successfully:

$$\ell_1 = \min_{\|\delta_{\mathcal{F}}\| \geq C} \mathbb{E}_{(x,y)\in\mathcal{D}} \min_{\theta} \ell\big(f_\theta(x+\delta_{\mathcal{F}}), y\big) \tag{10}$$

And we define the maximum utility loss with the norm perturbation less than C as:

$$\ell_2 = \max_{\|\delta_{\mathcal{F}}\| \leq C} \mathbb{E}_{(x,y)\in\mathcal{D}} \min_{\theta} \ell\big(f_\theta(x+\delta_{\mathcal{F}}), y\big) \tag{11}$$

Assumption 3. *Assume $\ell_2 \leq \ell_1$*

Assumption 3 elucidates that the utility loss associated with a perturbation norm less than C is smaller than the utility loss when the perturbation norm is greater than C. This assumption is logical, as larger perturbations would naturally lead to greater utility loss.

Assumption 4. *Suppose the federated model achieves zero training loss.*

We have the following theorem to elucidate the relation between feature sensitivity removing via Algo. 1 and exact unlearning (see proof in Appendix).

Theorem 2. *If Assumption 3 and 4 hold, the utility loss of unlearned model obtained by Algo. 1 is less than the utility loss with unlearning successfully, i.e.,*

$$\ell_u \leq \ell_1, \tag{12}$$

where $\ell_u = \mathbb{E}_{(x,y)\in\mathcal{D}}\big(\ell(f_{\theta^u}(x), y)$

Proof. When the unlearning happens during the federated training, the unlearning clients would also optimize the training loss and feature sensitivity simultaneously. Specifically, the optimization process could be written as:

$$\theta_u = \arg\min_{\theta} \mathbb{E}_{(x,y)\in\mathcal{D}}\big(\ell(f_\theta(x), y) + \lambda \mathbb{E}_{\delta_{\mathcal{F}}} \frac{\|f_\theta(x) - f_\theta(x+\delta_{\mathcal{F}})\|_2}{\|\delta_{\mathcal{F}}\|_2}\big),$$

where $\lambda \geq \frac{1}{C}$ is one coefficient. Without loss of generality, we assume the $\ell(f_\theta(x), y) = \|f_\theta(x) - y)\|$. Denote

$$\Theta^* = \arg\min_{\theta} \mathbb{E}_{(x,y)\in\mathcal{D}} \ell(f_\theta(x), y).$$

If Assumption 4 holds, then $f_{\theta^*}(x) = y$ for any $\theta^* \in \Theta^*$. Therefore, for any $\|\delta_{\mathcal{F}}\| \geq \frac{1}{\lambda}$ such that

$$\begin{aligned}
&\mathbb{E}_{(x,y)\in\mathcal{D}}\big(\ell(f_{\theta^*}(x), y) + \lambda \mathbb{E}_{\|\delta_{\mathcal{F}}\| \geq \frac{1}{\lambda}} \frac{\|f_\theta(x) - f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2}{\|\delta_{\mathcal{F}}\|_2}\big) \\
&= \lambda \mathbb{E}_{(x,y)\in\mathcal{D}} \mathbb{E}_{\|\delta_{\mathcal{F}}\| \geq \frac{1}{\lambda}} \frac{\|y - f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2}{\|\delta_{\mathcal{F}}\|_2} \\
&\leq \mathbb{E}_{(x,y)\in\mathcal{D}} \mathbb{E}_{\|\delta_{\mathcal{F}}\| \geq \frac{1}{\lambda}} \|y - f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2.
\end{aligned} \tag{13}$$

The last inequality is due to Therefore, we further obtain:

$$\begin{aligned}\ell_u &\leq \min_{\theta\in\mathbb{R}^d} \mathbb{E}_{(x,y)\in\mathcal{D}}\big(\ell(f_\theta(x),y) + \lambda\mathbb{E}_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}} \frac{\|f_\theta(x)-f_\theta(x+\delta_{\mathcal{F}})\|_2}{\|\delta_{\mathcal{F}}\|_2}\big) \\
&\leq \min_{\theta\in\Theta^*} \mathbb{E}_{(x,y)\in\mathcal{D}}\big(\ell(f_\theta(x),y) + \lambda\mathbb{E}_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}} \frac{\|f_\theta(x)-f_\theta(x+\delta_{\mathcal{F}})\|_2}{\|\delta_{\mathcal{F}}\|_2}\big) \\
&\leq \min_{\theta\in\Theta^*} \mathbb{E}_{(x,y)\in\mathcal{D}}\mathbb{E}_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}} \|y-f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2 \\
&\leq \mathbb{E}_{(x,y)\in\mathcal{D}}\mathbb{E}_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}} \min_{\theta\in\Theta^*} \|y-f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2 \\
&= \mathbb{E}_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}}\mathbb{E}_{(x,y)\in\mathcal{D}} \min_{\theta\in\Theta^*} \|y-f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2 \\
&\leq \max_{\|\delta_{\mathcal{F}}\|\geq\frac{1}{\lambda}} \mathbb{E}_{(x,y)\in\mathcal{D}} \min_{\theta\in\mathbb{R}^d} \|y-f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2 \\
&\leq \max_{\|\delta_{\mathcal{F}}\|\leq C} \mathbb{E}_{(x,y)\in\mathcal{D}} \min_{\theta\in\mathbb{R}^d} \|y-f_{\theta^*}(x+\delta_{\mathcal{F}})\|_2 \\
&= \ell_2,\end{aligned} \tag{14}$$

where the last inequality is due to $\lambda \geq \frac{1}{C}$. According to Assumption 3, we have $\ell_u \leq \ell_1$

□

A.2 Experimental Setup

Datasets *MNIST*[90]: Both the *MNIST*[90] and *Fashion-MNIST(FMNIST)*[92] datasets contain images of handwritten digits and attire, respectively. Each dataset comprises 60,000 training examples and 10,000 test examples. In both datasets, each example is represented as a single-channel image with dimensions of 28x28 pixels, categorized into one of 10 classes. Additionally, the *Colored-MNIST(CMNIST)*[90] dataset, an extension of the original MNIST, introduces color into the digits of each example. Consequently, images in the Colored MNIST dataset are represented in three channels. *CIFAR*[93]: The *CIFAR-10*[93] dataset comprises 60,000 images, each with dimensions of 32x32 pixels and three color channels, distributed across 10 classes. This dataset includes 6,000 images per class and is partitioned into 50,000 training examples and 10,000 test examples. Similarly, the *CIFAR-100*[93] dataset shares the same image dimensions and structure as *CIFAR-10* but extends to 100 classes, with each class containing 600 images. Within each class, there are 500 training images and 100 test images. Moreover, *CIFAR-100* organizes its 100 classes into 20 superclasses, forming the *CIFAR-20 dataset*[93]. *CelebA* [85]: A face recognition dataset featuring 40 attributes such as gender and facial characteristics, comprising 162,770 training examples and 19,962 test examples. This study will focus on utilizing the *CelebA*[85] dataset primarily for gender classification tasks.

(a) MNIST (b) CMNIST (c) FMNIST (d) CIFAR-10 (e) CIFAR-20 (f) CIFAR-100 (g) CelebA

Figure 8: Visual representation of dataset samples utilized in this study.

Adult Census Income (Adult)[86] includes 48, 842 records with 14 attributes such as age, gender, education, marital status, etc. The classification task of this dataset is to predict if a person earns over $50K a year based on the census attributes. We then consider marital status as the sensitive feature that aim to unlearn in this study. *Diabetes*[87] includes 768 personal health records of females at least 21 years old with 8 attributes such as blood pressure, insulin level, age and etc. The classification task of this dataset is to predict if a person has diabetes. We then consider number of pregnancies as the sensitive feature that aim to unlearn in this study.

Baselines The baseline methods in this study:

Baseline: Original model before unlearning.

Retrain: In scenarios involving sensitive feature unlearning, the retrained model was simply trained using a dataset where Gaussian noise was applied to the unlearned feature region. This approach may lead to performance deterioration, as discussed in Sec. 3.2. For backdoor feature unlearning scenarios, the retrained model was trained using the retain dataset $\mathcal{D}_r$, also referred to as the clean dataset. In biased feature unlearning scenarios, the retrained model was trained using a combination of 50% from each of the retain dataset $\mathcal{D}_r$ (bias dataset) and the unlearn client local dataset $\mathcal{D}_u$ (unbias dataset). This ensures fairness in the model's performance across both datasets.

Fine-tune: The baseline model is fine-tuned using the retained dataset $\mathcal{D}_r$ for 5 epochs. *Class-Discriminative Pruning(FedCDP)*[66]: A FU framework that achieves class unlearning by utilizing Term Frequency-Inverse Document Frequency (TF-IDF) guided channel pruning, which selectively removes the most discriminative channels related to the target category and followed by fine-tuning without retraining from scratch.

FedRecovery[62]: A FU framework that achieves client unlearning by removing the influence of a client's data from the global model using a differentially private machine unlearning algorithm that leverages historical gradient submissions without the need for retraining.

A.3 Attention Map

A.3.1 Backdoor Feature Unlearning

Attention map analysis for backdoor samples across model iterations of baseline, retrain, and unlearn model using our proposed Ferrari method on MNIST(Fig. 9), FMNIST(Fig. 10), CIFAR-10(Fig. 11), CIFAR-20(Fig. 12) and CIFAR-100 (Fig. 13)datasets.

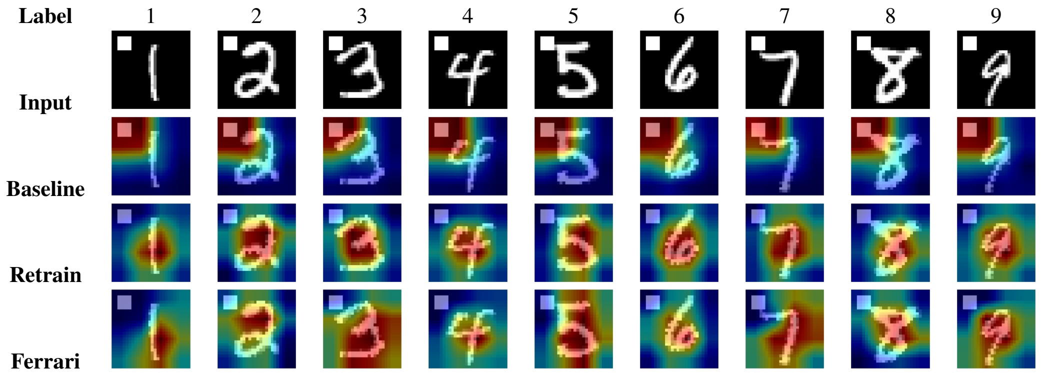


Figure 9: MNIST

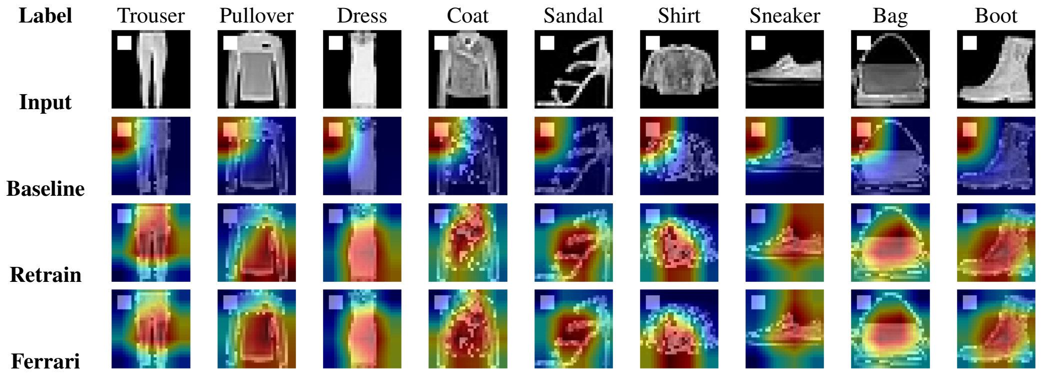


Figure 10: FMNIST

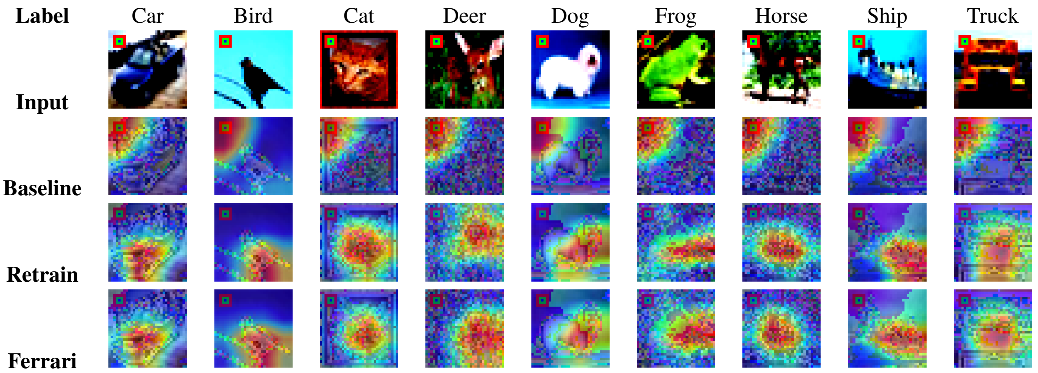


Figure 11: CIFAR-10

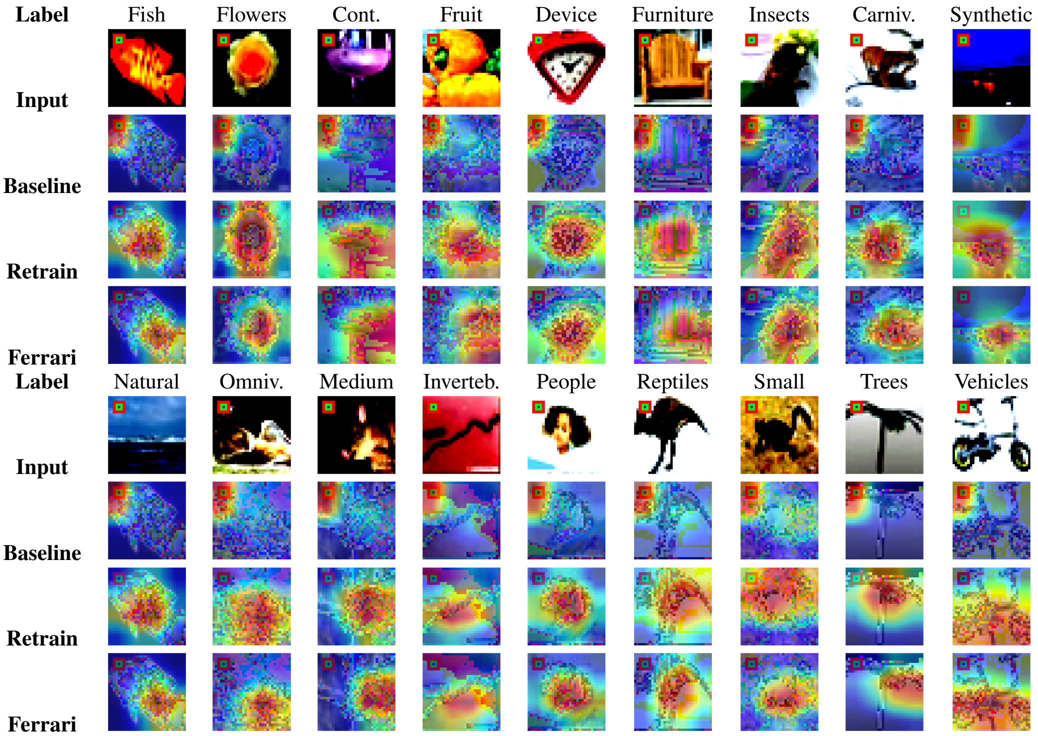
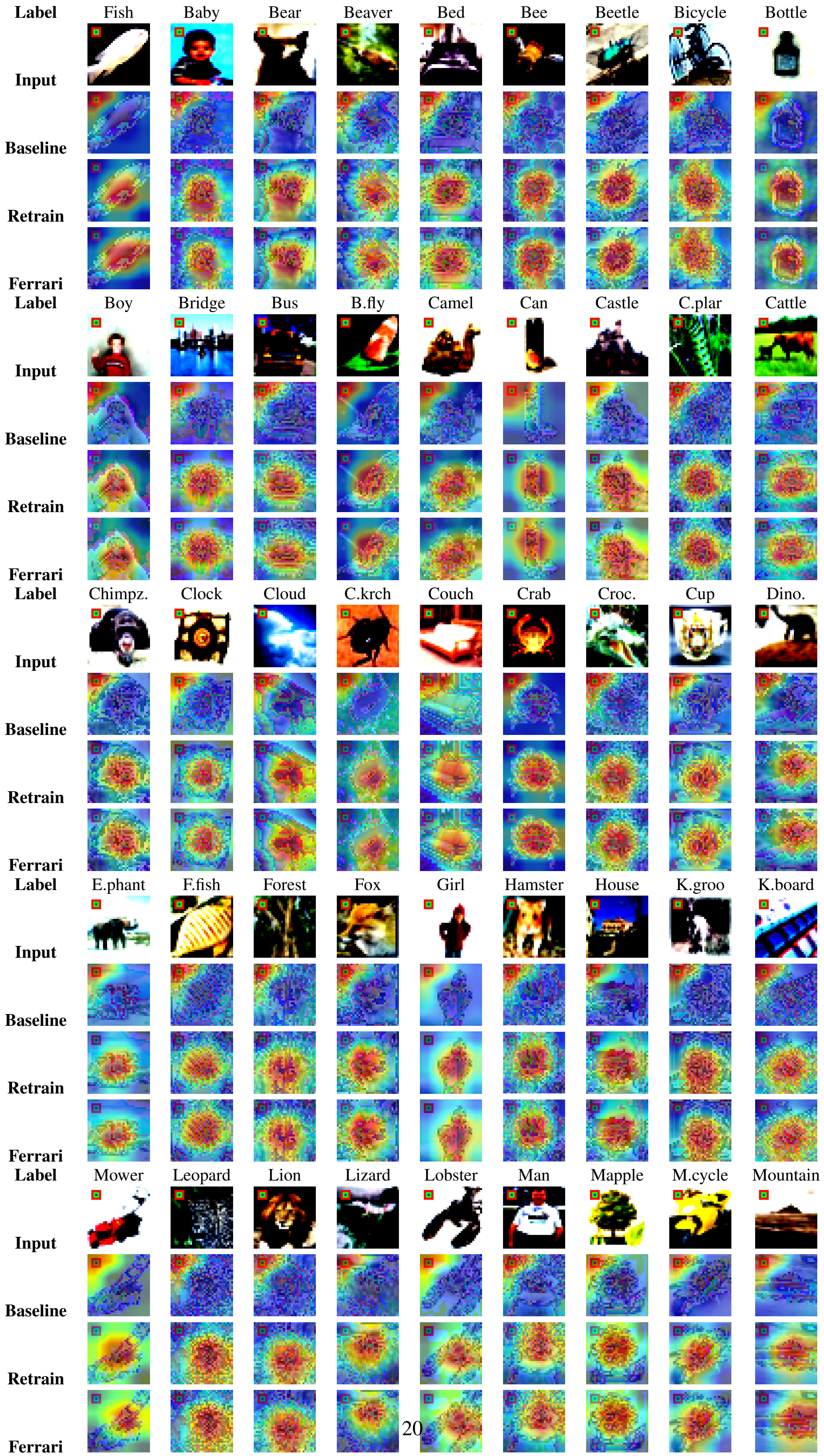


Figure 12: CIFAR-20



Continued on next page

Continued from previous page

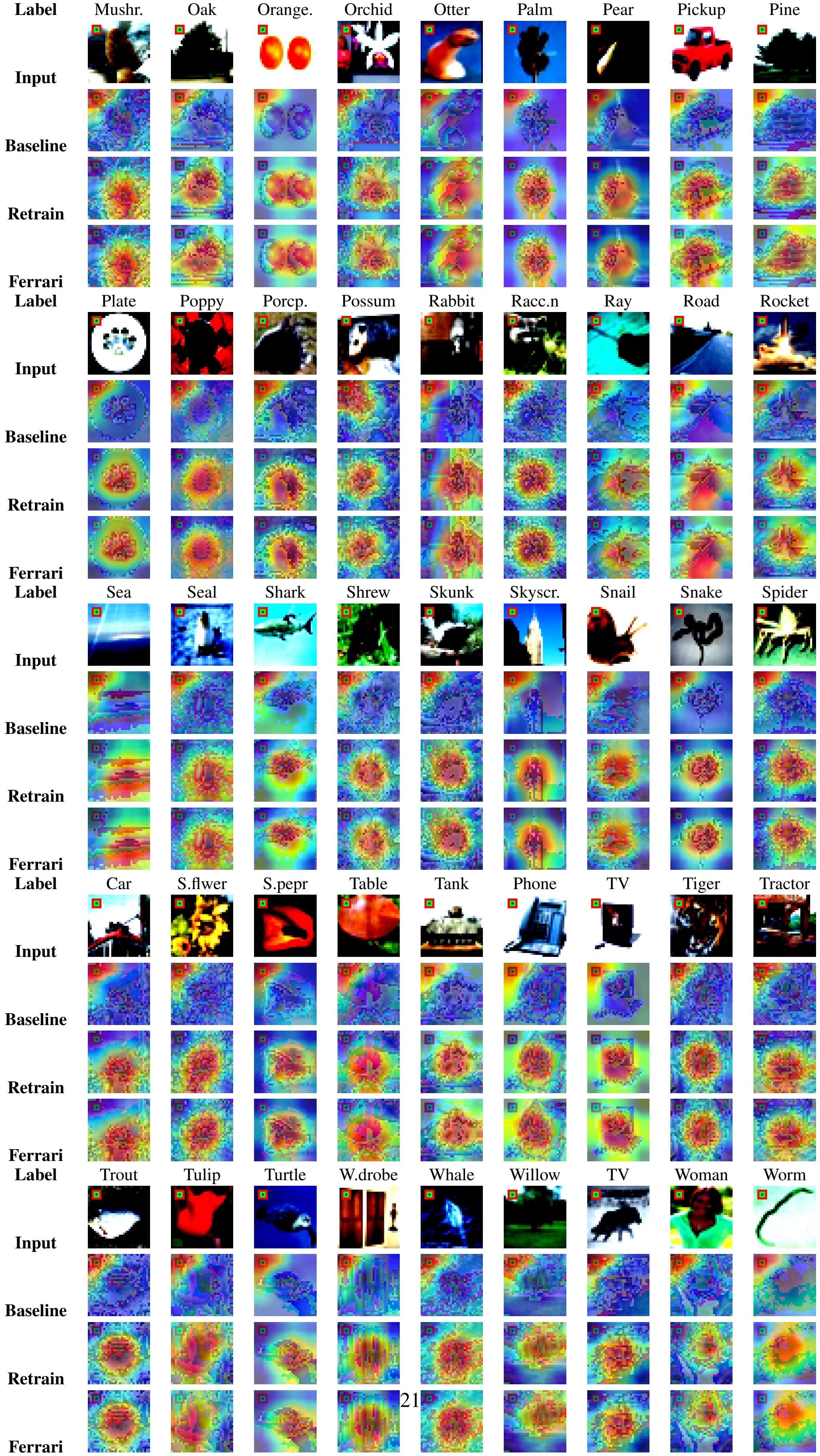
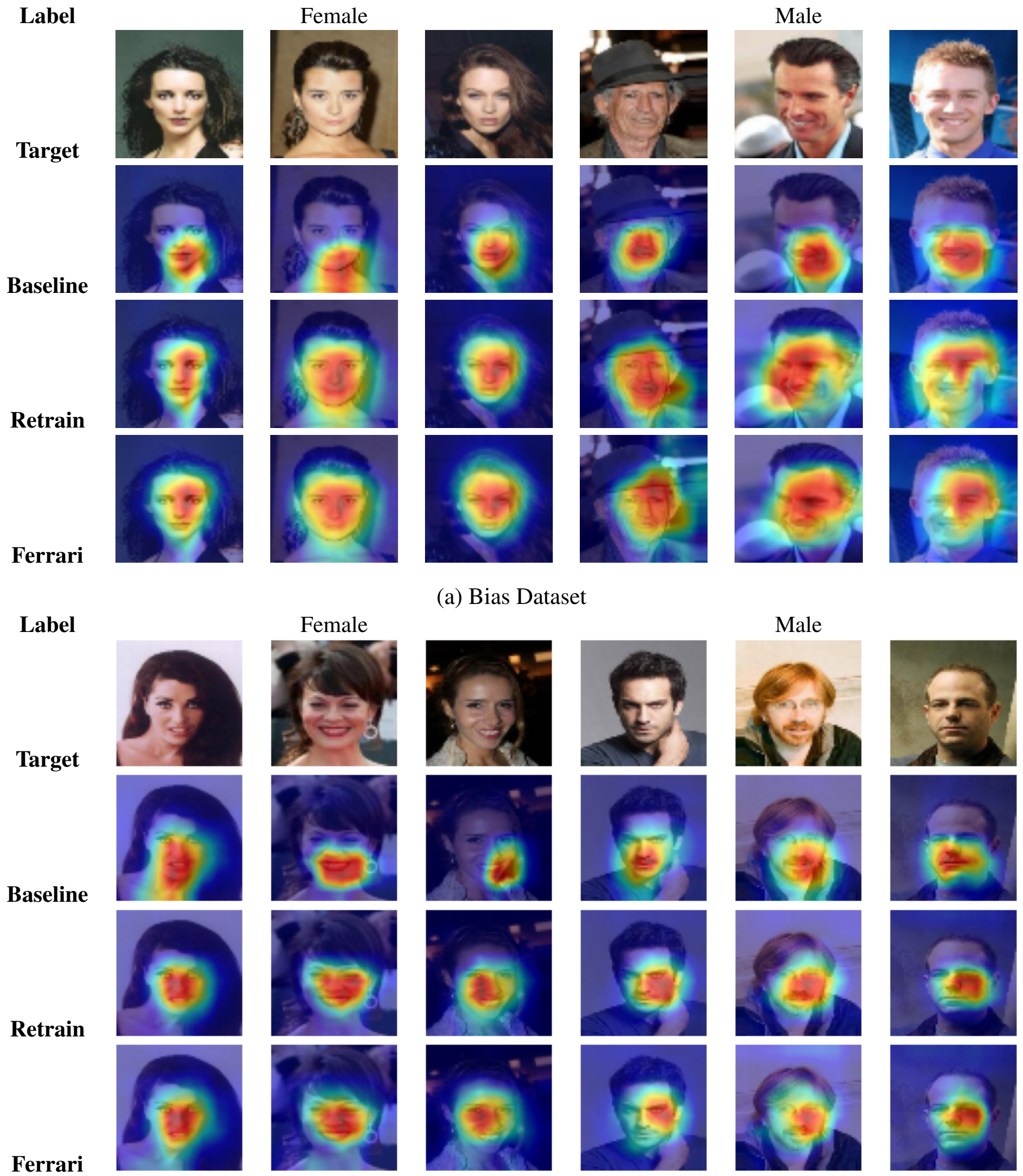


Figure 13: CIFAR-100

A.3.2 Biased Feature Unlearning



(a) Bias Dataset

(b) Unbias Dataset

Figure 14: Attention map analysis for bias and unbias samples across model iterations of baseline, retrain, and unlearn model using our proposed Ferrari to unlearn 'mouth' on CelebA dataset.

A.4 Limitation and Future Work

While our proposed approach of federated feature unlearning demonstrates effectiveness in various unlearning scenarios using only the local dataset of unlearning clients without requiring participation from other clients, thus simulating practical application, it has some inevitable limitations.

The proposed approach necessitates access to the entire dataset from the unlearning client to achieve maximal unlearning effectiveness. However, as demonstrated in Section 5.5, a partial dataset comprising at least 70% of the data yields similar performance to the full dataset. In certain cases, the unlearning client may lose a significant portion of their data, rendering our approach ineffective in such scenarios. Therefore, future work should investigate federated feature unlearning approaches that require only a small portion of the unlearning client's dataset. Additionally, the proposed approach has only been proven effective for classification models, as it was specifically designed for this purpose. Its effectiveness in other domains, such as generative models, remains to be investigated.

Therefore, future work should explore methods that require only a small portion of the client's dataset. Additionally, future research will investigate advanced perturbation techniques, support for diverse data types and models, and integration with other privacy-preserving methods to further enhance data protection in FL systems.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include theoretical results.

- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [NA]

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.

- Please see the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
   - If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.