Differentially Private 2D Human Pose Estimation

Supplementary Material

Symbol	Definition
Data and Features	
$\mathcal{S}_{\text{data}}$	Full dataset of size n
$\mathcal{S}_{\text{pub}}$	Public subset for subspace estimation, size m
$\mathcal{S}_{\text{priv}}$	Private subset for training, size $n-m$
x	Data sample with private features
$\psi(x)$	Public feature map (non-sensitive transformation)
B_{priv}	Private mini-batch
B_{psi}	Public mini-batch for Feature-DP
Model and Loss	
$w \in \mathbb{R}^p$	Total model parameters
$\ell(w,x)$	Total loss function
$\ell_{\text{priv}}(w,x)$	Private loss component (depends on private features)
$\ell_{\text{pub}}(w,\psi(x))$	Public loss component (depends only on $\psi(x)$)
$\hat{L}_n(w)$	Empirical risk: $\frac{1}{n}\sum_{i=1}^{n}\ell(w,x_i)$
ρ	Lipschitz constant of the gradient (smoothness parameter of loss)
Gradients and Sensitivity	
$\nabla\ell(w,x)$	Full gradient
$\nabla\ell_{\text{priv}}(w,x)$	Private gradient component
$\nabla\ell_{\text{pub}}(w,\psi(x))$	Public gradient component
G	Full gradient sensitivity: $\|\nabla\ell(w,x)\|_2 \le G$
C	Private gradient sensitivity: $\|\nabla l_{\text{priv}}(w,x)\|_2 \le C$
g^t	Gradient estimate at iteration t
$\tilde{g}$	Clipped gradient
Subspace and Projection	
k	Subspace dimension ($k \ll p$)
M_t	Gradient covariance matrix at iteration t
$\hat{V}_t \in \mathbb{R}^{p\times k}$	Top-k eigenvectors of M_t (subspace basis)
$\lambda_i(M_t)$	i-th eigenvalue of M_t (ordered: $\lambda_1 \ge \lambda_2 \ge \ldots$)
α_t	Eigengap: $\lambda_k(M_t) - \lambda_{k+1}(M_t)$
Λ	Average inverse squared eigengap: $\frac{1}{T}\sum_{t=1}^{T} 1/\alpha_t^2$
$\gamma_2(\mathcal{W}, d_w)$	Complexity measure for set of all possible model iterates $\mathcal{W}$
Privacy and Optimization	
(ε,δ)	Differential privacy parameters
σ	Gaussian noise standard deviation
T	Total number of iterations
η_t	Learning rate at iteration t

Table 1. Summary of Notation

1. Implementation Details

Our models are pretrained on the COCO *train2017* set, which consists of approximately 118k images with around 140k annotated human instances, each with 17 joint annotations. The *val2017* set consisting of around 5k images is used for validation. For evaluating the trade-off between utility and performance under various DP-SGD techniques we employ the MPII Human Pose Dataset consisting of 40k human instances, each labeled with 16 joint annotations. When transferring the model from COCO to MPII, we adjust for the keypoint discrepancy between datasets. We employ the Percentage of Correct Keypoints normalized by head (PCKh) [1] as an evaluation metric.

To assess model generalization under significant domain shifts, we utilize the Human-Art dataset [3], a large-scale benchmark designed to bridge natural and artificial visual domains . The dataset comprises 50,000 images with over 123,000 instances across 20 diverse scenarios, encompassing both natural scenes (e.g., dance, drama) and artistic representations (e.g., oil paintings, digital art) . Human-Art presents unique challenges absent in conventional datasets like MPII, including abstract depictions, distorted body proportions, and unconventional poses . We adhere to the standard MS COCO evaluation protocol, reporting Average Precision (AP) as the primary metric .

All models are trained under differential privacy constraints using DP-SGD with various clipping norms and privacy budgets (ϵ). Each model undergoes training for a total of 25 epochs, as we empirically observed no significant performance improvements when extending training beyond this duration under DP constraints. Throughout all experiments, we maintain a fixed input resolution of 256×192 pixels to ensure consistency across experiments and enable comparison with prior work.

For the privacy parameter settings, we use three gradient clipping norms $C = \{1.0, 0.1, 0.01\}$, with target privacy budgets of $\epsilon = \{0.2, 0.4, 0.6, 0.8\}$. We adopt Renyi Differential Privacy (RDP) [5] for privacy accounting with the privacy parameter $\delta = 4e-5$.

For the projection method, we randomly select 100 samples from the training dataset of MPII as S_{pub} (ensuring no image overlap with the private data) with the remaining data forming the private training set S_{priv}. The default projection dimension K is set to 50 for all experiments.

To generate the public feature map, we employ Gaussian blur as ψ with a kernel size of $(25, 25)$ and standard deviation $\sigma_X = 10$, which effectively suppresses facial and body structure details. We deliberately blur the entire image rather than selectively masking human regions, as this approach provides comprehensive privacy protection by obscuring not only human identities but also contextual environmental details as depicted in Figure 1(e).

1.1. Convergence Analysis of Feature-Projective DP

To contextualize our contribution, we first present the privacy error bounds for the baseline methods. The convergence of standard DP-SGD for non-convex objectives is limited by a privacy error term that scales with the ambient dimension p and the full gradient sensitivity G, specifically $\tilde{\mathcal{O}}(\frac{p \cdot G^2}{n\varepsilon})$ [2, 6]. The two frameworks we synthesize target distinct factors of the privacy error:

Gradient Projection (PDP-SGD) [7]: reduces the *dimensional* dependence by projecting the noise onto a lower-dimensional subspace, replacing p with the subspace dimension k. This yields a privacy error of $\tilde{\mathcal{O}}(\frac{k \cdot G^2}{n\varepsilon})$.

Feature-DP (FDP) [4]: reduces the *magnitude* dependence by privatizing only the sensitive private loss component, replacing the full sensitivity G with the private sensitivity C. This yields a privacy error of $\tilde{\mathcal{O}}(\frac{p \cdot C^2}{n\varepsilon})$.

The convergence analysis of our method formally establishes the utility gain as observed from our empirical results and is a direct corollary of the separate analyses from [4, 7].

Let the empirical risk be $\hat{L}_n(w) = \frac{1}{n}\sum_{i=1}^{n} l(w, x_i)$ on a private dataset S_{priv} of size n.

Assumption 1. The loss $l(w, x)$ can be decomposed into public and private components given as $\ell(w, x) = \ell_{priv}(w, x) + \ell_{pub}(w, \psi(x))$.

Assumption 2. The full loss $\hat{L}_n(w)$ is ρ-smooth, the full gradient $||\nabla \ell(w, x)||_2 \leq G$ is bounded where G defines the sensitivity for subspace reconstruction error and the private gradient is bounded by the threshold C as $||\nabla \ell_{priv}(w, x)||_2 \leq C$, where $C \leq G$.

Assumption 3. We have access to a separate public dataset S_{pub} of size m and $\hat{V}_t \in \mathbb{R}^{p \times k}$ is the k-dimensional projection matrix computed from top-k eigenspace as $M(w) = \frac{1}{m}\sum_{i=1}^{m} \nabla l(w, \tilde{z}_i) \nabla l(w, \tilde{z}_i)^T$ on S_{pub} (Eq.2 of main paper) at iteration w_{t-1}.

Assumption 4. Assuming the principal component of the gradient dominance condition is satisfied and under this, we denote the eigengap at iteration t as α_t and $\Lambda = \frac{1}{T}\sum_{t=1}^{T} 1/\alpha_t^2$ be average inverse squared eigengap and refer to $\gamma_2(\mathcal{W}, d_w)$ as the associated complexity measure (where $\mathcal{W}$ iterate set of the weights and d_w is distance between them), as defined in [8].

Under these assumptions, setting the total iterations $T = \mathcal{O}(n^2\varepsilon^2)$, the average expected gradient norm of feature-projective DP is bounded by:

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}||\nabla \hat{L}_n(w_t)||_2^2 \leq \underbrace{\tilde{\mathcal{O}}\left(\frac{k \cdot \rho \cdot C^2}{n\varepsilon}\right)}_{\text{Privacy Error}} + \underbrace{\mathcal{O}\left(\frac{\Lambda G^4 \rho^2 \gamma_2^2(\mathcal{W}, d_w) \ln p}{m}\right)}_{\text{Reconstruction Error}} \tag{1}$$

The convergence is bound by two terms: a reconstruction error inherited from use of public dataset S_{pub} and privacy error from the gaussian noise. By combining both the approaches, the privacy error scales with both the reduced dimension k and reduced gradient norm C which can be understood from the error bound changing from $\tilde{\mathcal{O}}(p \cdot G^2) \rightarrow \tilde{\mathcal{O}}(k \cdot C^2)$ which explains the feature-projective DP's higher utility for the same (ε, δ)-FDP guarantee.

2. Per Joint PCKh@0.5 on MPII and AP on HumanART

We provide an extensive evaluation of our proposed training strategies on the MPII dataset (Tables 2 - 5), highlighting the impact of gradient clipping norm C, privacy budget ε, and initialization strategies across four setups: vanilla DP-SGD, DP-SGD with projection, Feature DP-SGD, and our proposed Feature Projective DP. In the baseline DP-SGD (Table 2), fine-tuning from a pretrained model with $C = 0.01$ and $\epsilon = 0.8$ achieves a mean PCKh of 78.17%, while the same configuration trained from scratch drops to 12.74%, emphasizing the importance of pretrained representations under privacy constraints. Incorporating gradient projection (Table 3) significantly improves results across all settings, with fine-tuning at $C = 0.01, \epsilon = 0.8$ reaching 80.63% mean PCKh, and even training from scratch improving to 13.96%. Feature DP (Table 4), which perturbs only the private gradient component, also shows substantial gains over vanilla DP-SGD, particularly for pretrained models (80.41% at $C = 0.01, \epsilon = 0.8$). However, the most consistent and robust performance is achieved with our Feature Projective DP method (Table 5), which combines both projection and selective privatization. It reaches a peak mean PCKh of 82.50% when fine-tuning with $C = 0.01$ and $\epsilon = 0.2$, and maintains performance above 81% across all ϵ values, which shows strong utility even under tighter privacy budgets. Notably, feature projective DP is the only setup where fine-tuning from scratch reaches competitive accuracy (79.95% at $\epsilon = 0.8$), significantly outperforming all baselines and confirming its ability to generalize under both high privacy constraints and limited initialization. Additionally, we provide the qualitative results in Figure 1. These results collectively validate that our combined approach effectively mitigates utility loss inherent in private learning for structured prediction tasks such as human pose estimation.

On the HumanART dataset, we evaluate our differentially private methods under the same four configurations as MPII: DP-SGD, DP-SGD with projection, Feature DP-SGD, and Feature Projective DP-SGD, reporting COCO-style average precision (AP) and recall (AR) metrics. Unlike MPII, HumanART introduces substantial domain shift with stylized, abstract, and artistically distorted human figures, making it a much more challenging setting for generalization under differential privacy. As such, several configurations, especially those involving large clipping norms ($C = 1.0$) or training from scratch without pretraining result in negligible utility, often with AP scores below 1%, and are omitted from the main tables due to lack of interpretability. In the baseline DP-SGD setup (Table 6), only the most favorable setting ($C = 0.01, \epsilon = 0.4$) yields moderate performance, achieving 40.7 AP. However, once pro-

Table 2. MPII Results: DP-SGD.

Privacy Parameter(ϵ)	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	Mean	Mean@0.1
Finetuning									
C = 1.0									
$\epsilon = 0.2$	0.00	5.15	13.31	6.00	10.20	2.18	0.21	5.94	0.29
$\epsilon = 0.4$	0.99	6.27	14.83	7.68	12.97	2.38	3.26	8.36	0.34
$\epsilon = 0.6$	0.31	13.06	21.56	8.43	19.16	3.00	3.59	12.19	0.58
$\epsilon = 0.8$	2.97	10.43	19.16	7.06	22.31	3.75	8.12	12.53	0.61
C = 0.1									
$\epsilon = 0.2$	30.73	35.51	31.19	13.43	37.96	12.59	13.30	28.46	1.61
$\epsilon = 0.4$	43.01	49.92	47.98	20.20	50.67	18.22	21.56	39.78	2.43
$\epsilon = 0.6$	53.10	59.51	52.00	24.55	58.54	22.65	22.65	45.18	2.86
$\epsilon = 0.8$	64.56	64.44	53.09	28.29	60.05	30.29	28.74	49.93	3.42
C = 0.01									
$\epsilon = 0.2$	78.14	83.36	65.21	47.49	69.98	47.35	39.02	63.85	5.68
$\epsilon = 0.4$	83.83	88.88	77.02	63.83	74.87	62.50	52.12	73.77	9.30
$\epsilon = 0.6$	87.11	90.05	78.71	70.05	75.37	66.61	59.42	76.85	11.05
$\epsilon = 0.8$	87.79	90.32	78.95	71.99	77.91	69.07	61.55	78.17	11.91
Finetuning from scratch									
C = 1.0									
$\epsilon = 0.2$	4.09	2.96	0.82	1.51	0.71	0.75	0.59	1.39	0.06
$\epsilon = 0.4$	6.51	7.51	3.44	3.67	1.66	14.04	1.06	5.95	0.22
$\epsilon = 0.6$	16.17	11.79	7.41	3.92	9.23	9.27	1.82	8.67	0.37
$\epsilon = 0.8$	8.94	17.05	8.40	3.79	10.04	8.38	2.95	9.34	0.39
C = 0.1									
$\epsilon = 0.2$	12.48	20.87	15.70	10.91	23.39	13.96	8.36	16.11	0.68
$\epsilon = 0.4$	16.68	22.69	21.15	11.10	24.10	12.98	9.38	18.64	0.80
$\epsilon = 0.6$	23.26	28.07	21.70	13.06	26.92	14.57	9.73	21.73	1.07
$\epsilon = 0.8$	28.58	29.14	21.36	13.18	27.40	14.31	9.09	22.74	1.10
C = 0.01									
$\epsilon = 0.2$	28.89	31.52	22.43	12.11	27.37	16.72	10.39	24.05	1.16
$\epsilon = 0.4$	49.25	43.19	26.28	15.01	33.58	18.05	15.28	30.94	1.86
$\epsilon = 0.6$	60.20	50.68	30.99	16.99	37.84	20.15	20.34	35.77	2.25
$\epsilon = 0.8$	62.28	54.33	36.48	21.66	43.36	22.43	21.42	39.86	2.86
Training from scratch									
C = 1.0									
$\epsilon = 0.2$	0.14	0.05	0.37	0.02	0.64	0.26	0.35	0.30	0.02
$\epsilon = 0.4$	1.71	0.07	0.07	0.15	0.28	3.28	0.02	0.68	0.03
$\epsilon = 0.6$	0.03	4.64	0.00	1.25	0.90	3.10	0.05	1.45	0.07
$\epsilon = 0.8$	0.31	0.00	1.76	0.26	0.00	0.00	0.90	0.44	0.02
C = 0.1									
$\epsilon = 0.2$	1.09	0.03	2.97	7.01	1.28	1.23	0.50	2.33	0.09
$\epsilon = 0.4$	0.14	5.04	6.49	4.69	18.87	0.67	0.24	5.84	0.23
$\epsilon = 0.6$	0.10	9.51	8.01	9.58	22.90	2.08	1.23	8.12	0.33
$\epsilon = 0.8$	0.24	8.07	5.25	9.46	17.03	2.04	2.39	6.85	0.28
C = 0.01									
$\epsilon = 0.2$	0.31	10.31	10.07	4.90	13.33	3.41	1.77	8.17	0.36
$\epsilon = 0.4$	1.33	15.61	9.17	8.96	19.92	3.87	1.32	10.13	0.48
$\epsilon = 0.6$	6.79	17.24	13.50	9.60	21.53	5.74	2.13	12.68	0.56
$\epsilon = 0.8$	13.27	16.76	13.38	9.68	19.87	5.48	2.17	12.74	0.54

jection is introduced (Table 7), we observe significant gains particularly, for fine-tuned models with $C = 0.01, \epsilon = 0.8$ achieving 38.7 AP, and improvements also appearing at higher clipping norms. Feature DP (Table 8) follows a similar trend: without projection, utility is lower overall, peaking at 40.5 AP for $C = 0.01, \epsilon = 0.8$. The most notable performance is achieved by our feature projective DP method (Table 9), which achieves strong and stable results across all privacy levels. Specifically, fine-tuning with $C = 0.01, \epsilon = 0.4$ and $\epsilon = 0.8$ achieves 51.6 A which is the highest across all methods and settings, while maintaining over 50 AP even at $\epsilon = 0.2$, demonstrating robustness under strict privacy. Impressively, even models fine-tuned from scratch perform well under our proposed method, reaching 46.0 AP at $\epsilon = 0.8$, which is a significant jump from the near-zero utility of all other methods trained from scratch. Across the board, feature projective DP demonstrates the most reliable and consistent performance, outperforming both vanilla DP-SGD and projection or FDP variants.

Table 3. MPII Results: DP-SGD with Projection.

Privacy Parameter(ϵ)	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	Mean	Mean@0.1
Finetuning									
C = 1.0									
$\epsilon = 0.2$	3.48	14.79	6.85	8.38	17.36	8.85	8.62	10.34	0.42
$\epsilon = 0.4$	54.67	55.45	32.54	23.70	37.23	21.44	13.30	36.99	2.33
$\epsilon = 0.6$	55.97	49.10	31.99	28.22	39.67	24.10	14.34	37.90	2.31
$\epsilon = 0.8$	71.18	76.19	55.51	46.60	53.97	39.13	32.97	56.26	4.19
C = 0.1									
$\epsilon = 0.2$	88.85	89.44	75.78	68.46	61.21	63.83	54.68	73.13	10.56
$\epsilon = 0.4$	88.44	89.84	78.92	72.62	70.21	69.45	61.08	77.17	12.55
$\epsilon = 0.6$	90.31	90.20	79.27	71.43	66.02	68.75	58.01	76.11	12.21
$\epsilon = 0.8$	91.51	90.39	79.51	72.84	71.23	67.86	59.78	77.41	12.88
C = 0.01									
$\epsilon = 0.2$	92.02	90.78	79.10	72.47	72.72	70.74	64.29	78.48	13.67
$\epsilon = 0.4$	91.81	90.74	79.92	72.04	75.42	71.79	65.78	79.23	13.49
$\epsilon = 0.6$	92.29	91.78	80.48	73.75	74.16	72.29	67.88	79.89	14.28
$\epsilon = 0.8$	92.29	91.49	80.86	74.52	75.32	73.91	69.77	80.63	14.61
Finetuning from scratch									
C = 1.0									
$\epsilon = 0.2$	0.48	8.93	8.86	8.24	20.72	3.77	1.75	8.64	0.31
$\epsilon = 0.4$	4.13	14.88	14.32	6.99	3.41	5.80	3.73	8.74	0.38
$\epsilon = 0.6$	3.48	16.34	9.90	12.35	13.21	13.48	10.23	11.85	0.56
$\epsilon = 0.8$	2.69	15.73	11.20	11.79	19.65	6.83	1.94	11.22	0.51
C = 0.1									
$\epsilon = 0.2$	4.40	18.99	16.99	9.42	18.07	10.70	6.78	13.58	0.63
$\epsilon = 0.4$	12.45	15.30	17.25	10.90	21.50	9.61	7.01	14.36	0.61
$\epsilon = 0.6$	15.59	15.64	16.70	10.50	19.49	14.19	7.98	15.43	0.66
$\epsilon = 0.8$	13.34	23.30	15.51	10.08	20.06	8.68	9.05	15.92	0.66
C = 0.01									
$\epsilon = 0.2$	82.44	69.58	49.75	43.23	43.31	39.11	36.56	53.54	5.87
$\epsilon = 0.4$	83.77	75.70	55.19	50.44	52.12	46.12	45.35	59.82	7.87
$\epsilon = 0.6$	86.02	74.25	61.31	52.96	51.39	46.75	45.42	61.09	8.61
$\epsilon = 0.8$	87.14	77.77	63.32	56.18	57.14	50.92	48.89	64.28	9.68
Training from scratch									
C = 1.0									
$\epsilon = 0.2$	0.07	1.77	10.07	11.07	12.74	1.37	0.02	5.65	0.26
$\epsilon = 0.4$	1.36	6.98	17.69	9.77	5.24	0.95	0.07	6.76	0.28
$\epsilon = 0.6$	0.07	1.00	14.47	12.44	6.42	7.19	1.87	6.57	0.31
$\epsilon = 0.8$	0.17	7.24	11.71	5.91	2.68	4.27	1.23	5.89	0.24
C = 0.1									
$\epsilon = 0.2$	1.19	18.05	12.78	9.53	22.66	7.64	5.95	13.05	0.56
$\epsilon = 0.4$	0.75	13.08	6.17	5.55	21.91	3.36	5.50	9.80	0.41
$\epsilon = 0.6$	1.98	21.28	8.16	11.05	19.70	4.11	4.72	12.24	0.56
$\epsilon = 0.8$	1.30	13.20	4.48	8.31	22.62	6.73	4.27	10.57	0.43
C = 0.01									
$\epsilon = 0.2$	5.15	15.10	15.07	12.39	19.61	10.28	8.83	14.26	0.65
$\epsilon = 0.4$	9.21	20.60	15.78	10.67	22.33	13.06	7.01	15.54	0.70
$\epsilon = 0.6$	17.09	19.70	6.89	10.50	22.36	15.88	10.51	15.15	0.62
$\epsilon = 0.8$	9.48	19.55	16.12	10.91	22.45	6.61	3.57	13.96	0.60

Table 4. MPII Results: Feature DP.

Privacy Parameter(ϵ)	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	Mean	Mean@0.1
Finetuning									
C = 0.01									
$\epsilon = 0.2$	87.04	89.28	75.78	66.54	75.89	65.04	58.48	75.47	10.64
$\epsilon = 0.4$	90.76	91.30	78.51	69.95	78.55	68.85	63.86	78.60	12.71
$\epsilon = 0.6$	91.47	91.95	79.63	71.99	79.87	71.00	65.45	79.90	13.62
$\epsilon = 0.8$	92.16	92.15	79.85	72.81	80.41	71.55	66.30	80.41	14.22
C = 0.1									
$\epsilon = 0.2$	52.42	65.78	57.32	27.43	53.31	36.05	26.03	48.99	3.53
$\epsilon = 0.4$	78.04	81.62	64.53	40.43	67.70	45.14	39.58	61.73	5.44
$\epsilon = 0.6$	78.48	84.80	68.48	45.60	69.88	49.89	45.11	65.32	6.47
$\epsilon = 0.8$	82.74	85.21	70.29	48.50	71.39	54.50	47.85	67.60	7.03
C = 1.0									
$\epsilon = 0.2$	1.60	9.90	7.86	11.57	14.25	3.00	7.91	9.32	0.33
$\epsilon = 0.4$	11.49	17.05	8.30	11.08	23.89	8.20	9.57	15.01	0.59
$\epsilon = 0.6$	14.84	25.00	12.36	13.62	27.07	10.84	11.29	18.37	0.81
$\epsilon = 0.8$	21.69	29.28	16.16	14.56	26.55	15.66	15.45	22.05	1.01
Finetuning from scratch									
C = 0.01									
$\epsilon = 0.2$	71.15	57.17	38.71	22.32	47.24	28.39	29.12	43.89	3.48
$\epsilon = 0.4$	78.75	69.40	47.15	33.34	54.86	38.87	37.72	53.00	5.10
$\epsilon = 0.6$	82.44	72.69	52.00	38.19	59.22	41.69	41.14	56.78	6.07
$\epsilon = 0.8$	83.80	74.81	54.88	41.29	60.93	44.35	42.80	58.98	6.70
C = 0.1									
$\epsilon = 0.2$	33.94	30.42	19.36	12.83	27.19	15.03	10.82	23.15	1.14
$\epsilon = 0.4$	40.86	37.75	22.57	15.08	31.71	18.09	15.94	28.46	1.53
$\epsilon = 0.6$	47.68	43.27	25.09	16.31	35.35	19.56	17.78	31.90	1.92
$\epsilon = 0.8$	52.25	45.31	27.27	16.57	35.90	19.28	18.47	33.27	2.21
C = 1.0									
$\epsilon = 0.2$	12.14	8.85	8.76	10.54	11.72	9.47	4.04	10.03	0.39
$\epsilon = 0.4$	10.06	13.33	10.74	9.92	20.84	12.17	5.48	12.57	0.51
$\epsilon = 0.6$	7.03	12.65	17.11	11.29	21.67	15.35	3.19	12.84	0.52
$\epsilon = 0.8$	13.47	19.74	15.95	9.58	22.95	12.65	8.90	15.69	0.60
Training from scratch									
C = 0.01									
$\epsilon = 0.2$	6.92	18.29	17.40	10.96	22.62	14.89	3.71	15.17	0.60
$\epsilon = 0.4$	10.57	22.61	18.41	12.15	23.09	14.33	6.83	17.14	0.71
$\epsilon = 0.6$	11.39	22.18	17.16	11.91	24.23	15.90	8.86	17.57	0.82
$\epsilon = 0.8$	14.09	21.31	19.52	11.31	24.32	16.44	7.35	17.74	0.80
C = 0.1									
$\epsilon = 0.2$	4.23	0.56	7.82	8.64	19.44	1.91	1.96	7.13	0.26
$\epsilon = 0.4$	0.48	12.62	10.69	8.12	21.15	2.36	0.33	8.87	0.37
$\epsilon = 0.6$	0.48	15.08	10.64	9.00	19.72	4.33	1.20	9.60	0.41
$\epsilon = 0.8$	0.61	13.09	12.94	10.74	22.73	4.94	1.37	11.22	0.48
C = 1.0									
$\epsilon = 0.2$	0.31	0.32	0.14	0.00	0.03	0.06	0.12	0.35	0.02
$\epsilon = 0.4$	1.84	3.63	0.02	0.43	0.00	0.00	0.09	0.78	0.02
$\epsilon = 0.6$	0.00	2.62	0.05	4.61	8.34	0.12	0.00	2.39	0.10
$\epsilon = 0.8$	0.14	0.02	0.07	0.00	15.22	0.02	0.26	2.57	0.12

Table 5. MPII Results: Feature Projective DP.

Privacy Parameter(ϵ)	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	Mean	Mean@0.1
Finetuning									
C = 0.01									
$\epsilon = 0.2$	94.17	92.70	82.17	73.29	80.79	76.45	72.60	82.50	19.65
$\epsilon = 0.4$	94.13	92.70	81.46	73.10	80.51	75.44	71.00	82.01	19.23
$\epsilon = 0.6$	94.10	92.70	81.47	72.79	80.87	75.32	71.19	82.01	19.48
$\epsilon = 0.8$	94.27	92.63	81.58	72.50	80.37	75.56	70.88	81.91	19.22
C = 0.1									
$\epsilon = 0.2$	93.79	91.85	79.07	71.56	77.43	73.32	68.92	80.24	15.20
$\epsilon = 0.4$	94.03	92.56	80.72	73.19	79.85	74.95	69.96	81.60	17.03
$\epsilon = 0.6$	93.86	92.82	81.17	74.70	79.78	75.68	70.78	82.09	17.74
$\epsilon = 0.8$	94.78	93.21	82.31	74.51	80.63	76.04	71.30	82.62	18.64
C = 1.0									
$\epsilon = 0.2$	73.36	65.88	53.72	40.79	56.14	43.44	34.10	54.75	4.05
$\epsilon = 0.4$	86.53	82.17	63.52	54.50	57.68	49.20	43.01	64.02	6.09
$\epsilon = 0.6$	86.66	86.06	66.70	51.94	67.47	51.40	47.02	67.02	6.73
$\epsilon = 0.8$	91.13	89.06	73.91	63.39	62.92	59.60	51.75	71.66	9.37
Finetuning from scratch									
C = 0.01									
$\epsilon = 0.2$	92.91	88.94	76.55	67.33	77.81	71.11	65.99	78.11	16.82
$\epsilon = 0.4$	92.74	88.55	75.80	65.63	77.48	70.02	65.37	77.41	16.23
$\epsilon = 0.6$	93.21	88.62	75.61	65.15	77.79	69.33	64.52	77.23	16.37
$\epsilon = 0.8$	92.29	87.11	74.11	63.13	76.53	67.46	62.78	75.74	15.83
C = 0.1									
$\epsilon = 0.2$	91.95	85.61	71.11	59.78	73.84	64.58	61.36	73.51	14.33
$\epsilon = 0.4$	93.76	88.65	76.97	67.07	77.27	70.30	66.23	78.01	16.93
$\epsilon = 0.6$	93.49	89.08	77.09	67.81	79.09	72.23	67.48	78.86	16.93
$\epsilon = 0.8$	94.03	90.46	78.56	68.58	80.68	72.54	69.08	79.95	17.83
C = 1.0									
$\epsilon = 0.2$	4.20	8.02	13.07	9.20	22.99	7.64	7.02	10.98	0.46
$\epsilon = 0.4$	10.54	17.53	13.19	11.50	20.60	7.41	5.12	13.49	0.53
$\epsilon = 0.6$	15.52	20.67	14.91	12.11	21.64	14.99	8.10	16.29	0.78
$\epsilon = 0.8$	20.36	20.92	12.63	10.95	24.98	10.07	8.46	16.29	0.73
Training from scratch									
C = 0.01									
$\epsilon = 0.2$	16.75	19.53	17.62	12.41	24.84	14.35	9.99	17.34	0.71
$\epsilon = 0.4$	62.65	52.11	34.07	19.03	41.09	27.30	23.74	38.90	3.43
$\epsilon = 0.6$	71.66	60.36	38.86	22.51	46.65	31.51	26.26	44.28	4.39
$\epsilon = 0.8$	67.84	58.95	38.06	21.86	44.95	31.99	27.66	43.39	4.12
C = 0.1									
$\epsilon = 0.2$	14.02	17.56	14.69	10.86	20.65	15.84	7.96	15.50	0.72
$\epsilon = 0.4$	11.02	19.51	16.62	11.02	23.61	14.16	10.16	16.42	0.65
$\epsilon = 0.6$	18.21	24.56	19.24	12.34	26.62	15.68	12.49	19.78	0.93
$\epsilon = 0.8$	53.27	46.01	29.37	17.27	35.47	21.20	17.17	33.49	2.33
C = 1.0									
$\epsilon = 0.2$	1.13	14.74	9.56	8.19	22.14	14.47	3.54	12.39	0.54
$\epsilon = 0.4$	17.77	14.81	15.85	9.46	22.16	14.83	4.06	15.23	0.70
$\epsilon = 0.6$	18.49	21.08	15.39	11.12	22.54	5.14	5.48	14.67	0.66
$\epsilon = 0.8$	5.83	14.96	14.88	10.66	22.21	13.04	3.83	12.97	0.55

Table 6. HumanART Results: DP-SGD.

Privacy Parameter(ϵ)	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR	AR^{50}	AR^{75}	AR^{M}	AR^{L}
Finetuning										
C = 0.01										
$\epsilon = 0.2$	29.1	67.6	20.0	14.0	30.9	34.0	71.6	28.6	22.0	35.5
$\epsilon = 0.4$	40.7	74.7	39.3	23.8	42.7	45.8	77.1	46.7	32.9	47.9
$\epsilon = 0.6$	37.5	74.7	33.6	20.9	39.4	42.1	77.6	41.4	29.1	43.8
$\epsilon = 0.8$	39.0	75.9	36.0	22.2	40.9	43.5	78.3	43.6	30.1	45.3
C = 0.1										
$\epsilon = 0.2$	3.5	19.0	0.0	0.9	4.0	8.7	35.6	0.8	5.0	9.2
$\epsilon = 0.4$	7.7	32.4	0.5	2.3	8.5	14.6	47.7	4.0	8.7	15.3
$\epsilon = 0.6$	10.4	39.3	1.5	3.4	11.3	17.1	51.9	6.2	10.5	17.9
$\epsilon = 0.8$	12.0	43.3	2.4	3.9	13.0	18.4	54.1	7.3	11.4	19.2
Finetuning from scratch										
C = 0.01										
$\epsilon = 0.2$	0.9	5.8	0.0	0.2	1.0	3.8	19.7	0.1	2.6	3.9
$\epsilon = 0.4$	1.3	8.0	0.0	0.4	1.4	5.0	23.7	0.3	3.5	5.2
$\epsilon = 0.6$	1.7	10.7	0.0	0.5	1.9	6.1	27.5	0.6	4.2	6.4
$\epsilon = 0.8$	2.1	12.9	0.0	0.8	2.3	6.8	29.6	0.7	5.1	7.0
C = 0.1										
$\epsilon = 0.2$	0.0	0.4	0.0	0.0	0.1	0.8	5.7	0.0	0.5	0.8
$\epsilon = 0.4$	0.2	1.9	0.0	0.1	0.3	2.3	13.4	0.1	1.4	2.4
$\epsilon = 0.6$	0.5	3.6	0.0	0.1	0.6	3.3	18.6	0.0	2.2	3.4
$\epsilon = 0.8$	0.6	4.3	0.0	0.1	0.7	3.5	20.0	0.0	2.4	3.7

Table 7. HumanART Results: DP-SGD with projection.

Privacy Parameter(ϵ)	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR	AR^{50}	AR^{75}	AR^{M}	AR^{L}
Finetuning										
C = 0.01										
$\epsilon = 0.2$	38.3	73.9	35.6	21.7	40.4	43.6	76.9	43.6	30.8	45.8
$\epsilon = 0.4$	34.7	73.4	29.1	18.3	36.6	39.7	76.0	37.9	26.9	41.4
$\epsilon = 0.6$	40.3	75.8	38.4	22.5	42.6	46.1	78.5	47.0	33.1	47.8
$\epsilon = 0.8$	38.7	74.3	36.3	21.1	40.8	44.3	77.3	44.5	32.1	45.9
C = 0.1										
$\epsilon = 0.2$	31.6	69.1	24.7	17.0	33.3	36.8	72.7	33.1	26.0	38.2
$\epsilon = 0.4$	32.9	70.4	27.0	18.0	34.6	38.1	73.2	35.2	27.2	39.5
$\epsilon = 0.6$	33.6	71.3	26.9	18.4	35.4	38.6	74.3	35.2	28.2	40.0
$\epsilon = 0.8$	33.5	69.4	28.4	18.4	35.3	38.7	72.6	36.8	27.1	40.2
C = 1.0										
$\epsilon = 0.2$	1.1	6.7	0.0	0.3	1.3	4.8	21.9	0.2	3.2	5.0
$\epsilon = 0.4$	11.7	41.7	2.0	3.8	12.8	18.1	52.4	7.8	11.1	19.0
$\epsilon = 0.6$	10.7	39.4	1.9	2.9	11.7	17.1	51.2	6.8	9.7	18.0
$\epsilon = 0.8$	23.5	60.4	14.2	11.1	25.2	30.2	66.2	24.2	20.2	31.5
Finetuning from scratch										
C = 0.01										
$\epsilon = 0.2$	1.7	8.5	0.4	0.4	1.9	4.9	19.5	1.1	2.9	5.2
$\epsilon = 0.4$	3.5	15.2	0.3	1.1	3.9	8.4	29.0	2.6	5.8	8.7
$\epsilon = 0.6$	2.1	10.2	0.1	0.7	2.4	5.9	22.2	1.4	3.4	6.2
$\epsilon = 0.8$	1.4	7.3	0.1	0.5	1.6	4.3	18.2	0.6	2.4	4.5
C = 0.1										
$\epsilon = 0.2$	0.5	2.8	0.1	0.2	0.6	2.2	10.9	0.1	1.5	2.3
$\epsilon = 0.4$	0.6	4.0	0.0	0.3	0.7	2.5	12.9	0.1	1.7	2.6
$\epsilon = 0.6$	1.2	5.7	0.5	0.4	1.5	3.4	15.3	0.6	1.7	3.7
$\epsilon = 0.8$	1.1	4.2	1.0	0.3	1.2	3.1	14.0	0.5	1.8	3.2
C = 1.0										
$\epsilon = 0.2$	0.1	0.9	0.0	0.0	0.2	0.9	5.6	0.0	0.6	0.9
$\epsilon = 0.4$	0.3	2.0	0.1	0.0	0.4	1.9	11.1	0.1	1.0	2.0
$\epsilon = 0.6$	0.8	4.6	0.0	0.1	0.9	3.1	16.1	0.1	1.7	3.3
$\epsilon = 0.8$	0.2	1.3	0.1	0.0	0.3	1.3	7.5	0.1	0.9	1.4
Training from scratch										
C = 0.01										
$\epsilon = 0.2$	0.12	0.68	0.06	0.0	0.17	1.06	5.7	0.1	0.5	1.1
$\epsilon = 0.4$	0.0	0.09	0.0	0.0	0.0	0.09	0.71	0.0	0.10	0.09
$\epsilon = 0.6$	0.31	2.1	0.0	0.07	0.4	1.8	11.7	0.01	1.18	1.93
$\epsilon = 0.8$	0.11	0.82	0.0	0.04	0.2	1.2	7.0	0.02	1.15	1.20
C = 0.1										
$\epsilon = 0.2$	0.05	0.29	0.0	0.06	0.38	2.5	0.0	0.3	0.4	1.24
$\epsilon = 0.4$	0.12	0.91	0.0	0.05	0.16	1.35	7.58	0.04	1.3	1.37
$\epsilon = 0.6$	0.19	1.27	0.0	0.04	0.23	1.6	9.06	0.03	1.02	1.6
$\epsilon = 0.8$	0.02	0.11	0.0	0.0	0.03	2.09	0.0	0.25	0.32	0.31

Table 8. HumanART Results: Feature Projection DP-SGD.

Privacy Parameter(ϵ)	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR	AR^{50}	AR^{75}	AR^{M}	AR^{L}
Finetuning										
C = 0.01										
$\epsilon = 0.2$	32.6	71.8	26.4	17.1	34.6	38.1	75.2	35.1	26.6	39.6
$\epsilon = 0.4$	37.3	74.6	34.2	20.7	39.5	42.6	77.7	42.4	30.4	44.2
$\epsilon = 0.6$	39.3	75.9	37.6	22.2	41.7	44.5	78.8	45.4	31.4	46.2
$\epsilon = 0.8$	40.5	76.9	39.4	23.1	42.7	45.6	79.4	47.0	32.5	47.3
C = 0.1										
$\epsilon = 0.2$	8.7	37.0	1.1	3.1	9.4	14.1	48.7	2.9	9.6	14.7
$\epsilon = 0.4$	13.5	47.2	2.1	4.8	14.6	18.8	55.9	6.8	12.3	19.6
$\epsilon = 0.6$	15.8	51.9	3.5	5.9	17.0	21.1	59.1	9.3	14.1	22.0
$\epsilon = 0.8$	17.1	53.5	4.6	6.6	18.4	22.4	60.6	10.8	14.9	23.3
Finetuning from scratch										
C = 0.01										
$\epsilon = 0.2$	4.8	23.0	0.2	0.9	5.5	10.4	38.4	2.4	6.3	10.9
$\epsilon = 0.4$	7.7	30.0	1.7	1.7	8.5	13.6	43.1	5.0	8.5	14.2
$\epsilon = 0.6$	8.8	32.8	2.3	2.0	9.7	14.8	45.2	6.3	9.3	15.4
$\epsilon = 0.8$	9.5	34.5	2.6	2.2	10.6	15.7	46.9	7.1	10.1	16.3
C = 0.1										
$\epsilon = 0.2$	0.2	1.9	0.0	0.0	0.3	2.3	13.6	0.0	1.4	2.5
$\epsilon = 0.4$	0.7	4.9	0.0	0.0	0.8	3.2	18.3	0.09	1.7	3.3
$\epsilon = 0.6$	1.0	7.0	0.0	0.2	1.1	3.9	21.0	0.1	2.7	4.0
$\epsilon = 0.8$	1.3	9.3	0.0	0.2	1.5	4.4	23.1	0.1	2.8	4.6

Table 9. HumanART Results: Feature Projection DP-SGD plus projection.

Privacy Parameter(ϵ)	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}	AR	AR^{50}	AR^{75}	AR^{M}	AR^{L}
Finetuning										
C = 0.01										
$\epsilon = 0.2$	50.6	80.2	53.4	31.3	52.8	55.3	82.4	59.1	41.3	57.2
$\epsilon = 0.4$	51.6	80.4	55.0	33.5	53.8	56.2	82.7	60.3	42.2	58.1
$\epsilon = 0.6$	50.9	80.2	53.6	32.5	53.2	55.9	82.7	59.9	42.1	57.8
$\epsilon = 0.8$	51.6	80.2	54.7	33.6	53.7	56.2	82.5	60.2	42.3	58.1
C = 0.1										
$\epsilon = 0.2$	40.4	75.5	38.9	24.4	42.3	45.4	78.0	46.0	33.0	47.1
$\epsilon = 0.4$	41.1	76.8	40.2	24.6	43.3	46.8	79.5	48.2	34.4	48.4
$\epsilon = 0.6$	42.2	75.7	42.2	25.2	44.3	47.5	78.6	49.3	34.4	49.2
$\epsilon = 0.8$	43.0	77.0	42.9	25.7	45.1	48.4	79.9	50.3	35.7	50.1
C = 1.0										
$\epsilon = 0.2$	26.8	62.1	19.0	13.5	28.4	32.9	67.0	28.6	23.2	34.2
$\epsilon = 0.4$	32.4	70.5	25.4	19.0	34.2	39.0	74.7	36.6	29.9	40.3
$\epsilon = 0.6$	33.9	70.6	28.7	18.3	35.8	40.1	74.4	38.5	28.8	41.6
$\epsilon = 0.8$	35.4	72.1	30.7	20.2	37.2	40.7	75.1	38.6	29.9	42.1
Finetuning from scratch										
C = 0.01										
$\epsilon = 0.2$	45.3	76.9	46.4	29.5	47.3	50.4	79.7	53.0	38.0	52.1
$\epsilon = 0.4$	45.1	76.6	46.5	28.0	47.2	50.5	79.1	53.5	36.8	52.3
$\epsilon = 0.6$	43.8	75.2	44.5	28.5	45.6	49.6	78.0	52.2	38.3	51.1
$\epsilon = 0.8$	46.0	76.7	47.5	29.3	48.1	51.4	79.7	54.5	38.0	53.2
C = 0.1										
$\epsilon = 0.2$	6.8	24.4	2.3	2.8	7.4	11.6	35.9	5.2	8.8	12.0
$\epsilon = 0.4$	22.0	53.7	14.5	12.0	23.3	27.7	60.3	22.1	20.6	28.7
$\epsilon = 0.6$	29.1	62.4	24.0	17.7	30.8	34.9	67.8	32.0	26.2	36.1
$\epsilon = 0.8$	33.0	68.0	27.9	21.5	34.4	38.6	72.0	36.5	30.0	39.8

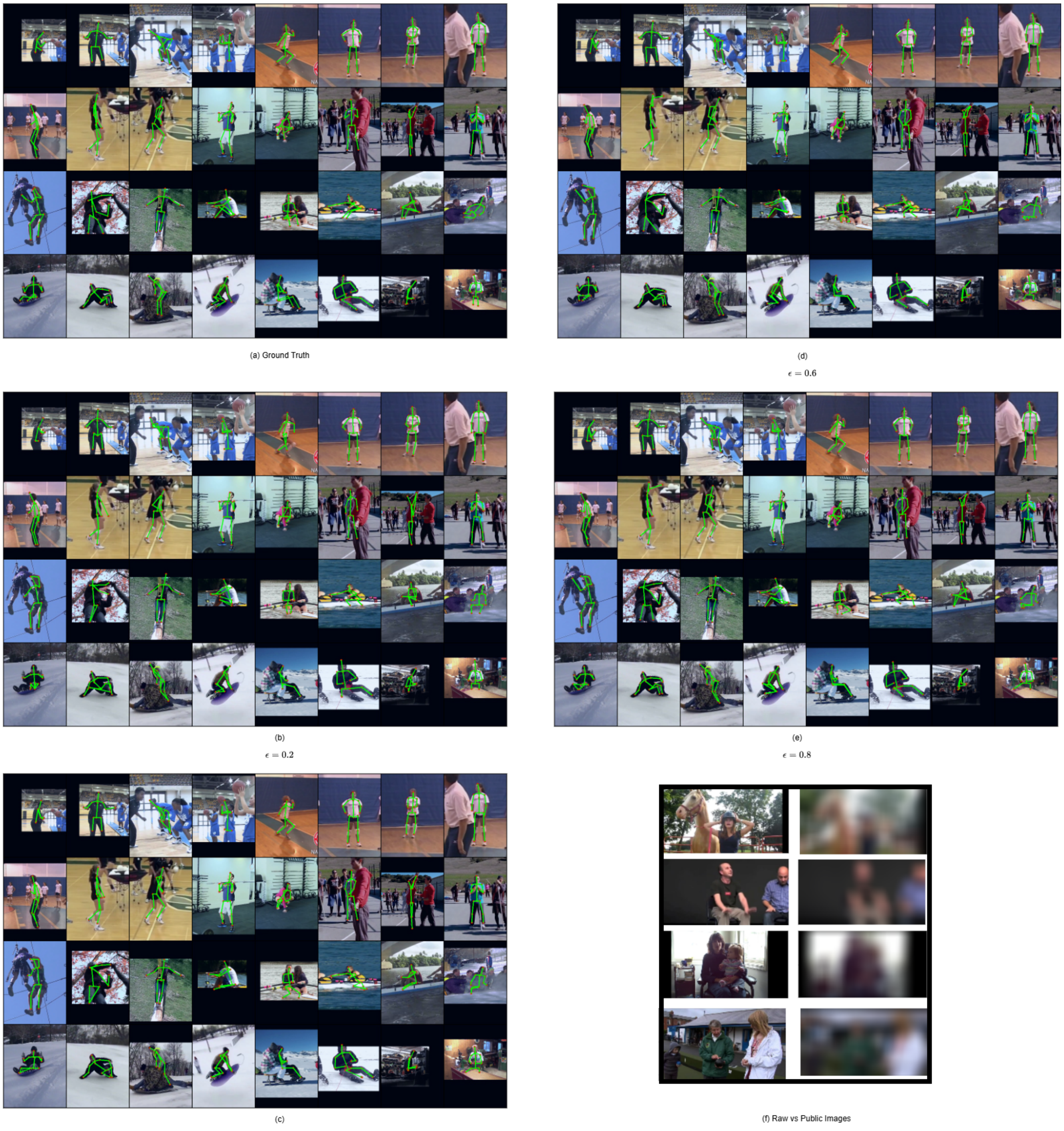

Figure 1. Figures (a-e)Depiction of qualitative results on DP-SGD, Projection DP-SGD and Feature Projection DP-SGD. We specifically show results on Finetuning with $C = 0.1$ at various privacy budgets. (f) Representation of Raw (Private) image compared to public feature (gaussian blurred).

References

[1] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on computer Vision and Pattern Recognition*, pages 3686–3693, 2014. 1

[2] Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *2014 IEEE 55th annual symposium on foundations of computer science*, pages 464–473. IEEE, 2014. 2

[3] Xuan Ju, Ailing Zeng, Jianan Wang, Qiang Xu, and Lei Zhang. Human-art: A versatile human-centric dataset bridging natural and artificial scenes. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 618–629, 2023. 1

[4] Saeed Mahloujifar, Chuan Guo, G Edward Suh, and Kamalika Chaudhuri. Machine learning with privacy for protected attributes. In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 2640–2657. IEEE, 2025. 2

[5] Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pages 263–275. IEEE, 2017. 1

[6] Di Wang and Jinhui Xu. Differentially private empirical risk minimization with smooth non-convex loss functions: A non-stationary view. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1182–1189, 2019. 2

[7] Yingxue Zhou, Zhiwei Steven Wu, and Arindam Banerjee. Bypassing the ambient dimension: Private sgd with gradient subspace identification. *arXiv preprint arXiv:2007.03813*, 2020. 2

[8] Yingxue Zhou, Steven Wu, and Arindam Banerjee. Bypassing the ambient dimension: Private {sgd} with gradient subspace identification. In *International Conference on Learning Representations*, 2021. 2