1 Full Benchmark Tables

This material includes all the experiments conducted on PhysBench.

Among UBFC-PHYS[6], there are some very noisy samples, we have excluded some of them but cannot guarantee that the remaining samples are of high quality. In the dark skin samples of MMPD[9], there are also some noisy sports scene samples, but we did not do any additional processing. Apart from this, all benchmark tests strictly use the same data to ensure fairness.

The samples excluded in UBFC-PHYS are as follows:

s3-T1 s8-T1 s9-T1 s26-T1 s28-T1 s30-T1 s31-T1 s32-T1 s33-T1 s40-T1 s52-T1 s53-T1 s54-T1 s56-T1 s1-T2 s4-T2 s6-T2 s8-T2 s9-T2 s11-T2 s12-T2 s13-T2 s14-T2 s19-T2 s21-T2 s22-T2 s25-T2 s26-T2 s27-T2 s28-T2 s31-T2 s32-T2 s33-T2 s35-T2 s38-T2 s39-T2 s41-T2 s42-T2 s45-T2 s47-T2 s48-T2 s52-T2 s53-T2 s55-T2 s5-T3 s8-T3 s9-T3 s10-T3 s13-T3 s14-T3 s17-T3 s22-T3 s25-T3 s26-T3 s28-T3 s30-T3 s32-T3 s33-T3 s35-T3 s37-T3 s40-T3 s47-T3 s48-T3 s49-T3 s50-T3 s52-T3 s53-T3

1.1 Quality of datasets

RLAP: High synchronicity, low compression, low noise.
RLAP-rPPG: High synchronicity, uncompressed, low noise.
PURE[8]: High synchronicity, uncompressed, low noise.
UBFC-rPPG[1]: A part of videos are not synchronized, uncompressed, low noise.
MMPD[9]: Tiny offset, high compression, dark skin tone and motion samples may have high noise.
MMPD-Simplest: Tiny offset, high compression, low noise.
COHFACE[3]: High synchronicity, very high compression, low noise.
UBFC-PHYS[6]: Low synchronicity, low compression, high noise.

1.2 Experimental Platform

System: Windows 11
Python: 3.9
TensorFlow: 2.10
CUDA: 11.3
CUDNN: 8.2
GPU: Nvidia RTX 3090 24G
CPU: AMD Ryzen 9 5950X 16-Core
RAM: 128G

1.3 Carbon Emissions

Thanks to the efficient program architecture, experiments using PhysBench are usually several times to tens of times faster than similar frameworks. This means that compared to using other frameworks, PhysBench can save time and reduce carbon emissions.

- PhysNet training time on PURE (32 batch size):
  4s per epoch on PhysBench, 25s per epoch on rPPG-Toolbox.
- PhysNet testing time on UBFC-rPPG:
  a total of 2s on PhysBench, 47s on rPPG-Toolbox.

The total power of the platform is approximately 550w, it takes about 72 hours to complete all experiments, and the carbon dioxide emissions are about 39.5 kg.

2 Training on RLAP

All models in the table are trained on the complete RLAP dataset, where the training set's fold is 1, 2, 3, validation set's fold is 0 and test set's fold is 4. The specific division can be downloaded from the Github homepage for nutrition labels.

2.1 Evaluation of Heart Rate

Table 1: The HR experimental results on RLAP, PURE, UBFC-rPPG

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	1.52	4.40	0.906	1.76	4.87	0.877	2.80	8.31	0.937	1.06	1.51	0.997
TS-CAN[4]	1.23	3.59	0.937	1.23	3.82	0.922	2.12	6.67	0.960	0.99	1.44	0.997
EfficientPhys[5]	1.05	3.41	0.943	1.00	3.39	0.939	1.33	5.97	0.968	1.03	1.45	0.997
PhysNet[10]	1.12	4.13	0.916	1.04	3.80	0.923	0.51	0.91	0.999	0.92	1.46	0.997
PhysFormer[11]	1.56	6.28	0.803	0.78	2.83	0.957	1.63	9.45	0.941	1.06	1.53	0.997
Seq-rPPG	1.07	4.15	0.917	0.81	2.97	0.953	0.37	0.63	1.000	0.87	1.40	0.997
NoobHeart	1.79	5.85	0.832	1.57	4.71	0.883	0.45	0.70	1.000	1.14	1.69	0.996

Table 2: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	13.7	23.1	0.201	1.03	1.46	0.987	2.75	8.63	0.733	14.5	24.6	0.241
TS-CAN[4]	12.3	22.1	0.258	0.95	1.40	0.989	2.28	7.81	0.774	15.0	24.5	0.267
EfficientPhys[5]	10.7	20.7	0.342	1.59	5.41	0.821	3.94	12.0	0.528	15.2	25.3	0.259
PhysNet[10]	13.1	24.1	0.176	0.97	1.45	0.988	19.6	26.9	-0.45	14.8	24.8	0.280
PhysFormer[11]	12.8	22.1	0.251	1.70	4.13	0.890	20.0	26.1	-0.37	14.7	24.7	0.278
Seq-rPPG	14.6	25.2	0.275	1.52	3.93	0.915	16.1	25.7	-0.12	15.1	25.2	0.263
NoobHeart	21.9	30.0	0.183	2.78	6.31	0.763	25.0	29.5	-0.36	14.4	24.4	0.283

2.2 Evaluation of Heart Rate Variability (SDNN)

Table 3: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	57.6	64.2	0.338	86.0	92.0	0.297	30.0	37.8	0.648
TS-CAN[4]	50.1	59.3	0.395	61.4	74.1	0.293	25.6	31.8	0.588
EfficientPhys[5]	43.7	53.7	0.356	28.0	44.0	0.468	10.1	15.4	0.827
PhysNet[10]	36.4	43.8	0.306	22.5	35.7	0.560	12.2	14.9	0.887
PhysFormer[11]	28.8	34.4	0.450	21.6	32.0	0.576	8.37	11.1	0.921
Seq-rPPG	14.4	22.1	0.424	9.51	15.8	0.872	4.73	8.25	0.911
NoobHeart	52.3	57.3	0.488	50.8	58.1	0.657	33.1	36.5	0.697

3 Training on UBFC-rPPG

All models are trained on UBFC-rPPG, with the first 80% used for training and the last 20% used for validation.

3.1 Evaluation of Heart Rate

Table 4: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$
DeepPhys[2]	12.6	18.4	0.048	11.8	17.2	0.168	10.0	17.9	0.627	-	-	-
TS-CAN[4]	10.0	16.6	0.217	9.50	15.1	0.347	4.79	11.9	0.773	-	-	-
EfficientPhys[5]	4.86	10.5	0.521	5.14	10.4	0.581	9.35	20.4	0.570	-	-	-
PhysNet[10]	6.15	13.4	0.408	6.24	13.5	0.375	8.82	19.4	0.694	-	-	-
PhysFormer[11]	11.7	22.3	-0.12	11.7	21.5	-0.035	17.6	29.0	0.398	-	-	-
Seq-rPPG	18.4	31.7	0.007	18.4	30.8	0.079	25.6	43.4	0.248	-	-	-
NoobHeart	5.38	12.9	0.456	6.89	14.5	0.373	2.30	10.6	0.893	-	-	-

Table 5: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$
DeepPhys[2]	17.1	24.2	0.174	8.25	12.1	0.214	14.1	19.4	0.078	19.0	27.0	0.004
TS-CAN[4]	16.5	24.7	0.140	3.15	6.38	0.737	11.0	16.7	0.222	16.0	25.2	0.188
EfficientPhys[5]	14.8	23.8	0.120	2.83	8.54	0.609	7.07	15.0	0.361	15.0	24.6	0.272
PhysNet[10]	12.3	21.9	0.226	4.70	12.6	0.370	5.15	10.6	0.577	20.1	31.2	0.010
PhysFormer[11]	16.5	24.7	0.107	16.2	23.5	-0.05	7.36	11.9	0.347	32.6	43.0	-0.20
Seq-rPPG	25.7	34.5	0.063	24.6	34.4	-0.02	8.18	15.8	0.326	36.9	47.0	-0.11
NoobHeart	27.4	37.0	-0.06	3.44	8.21	0.691	5.44	14.6	0.398	12.6	21.8	0.361

3.2 Evaluation of Heart Rate Variability (SDNN)

Table 6: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$	MAE↓	RMSE↓	$\rho\uparrow$
DeepPhys[2]	90.7	93.4	-0.22	85.9	91.3	0.136	-	-	-
TS-CAN[4]	90.8	93.3	0.117	95.1	97.8	0.152	-	-	-
EfficientPhys[5]	82.2	85.9	0.227	86.2	94.4	0.205	-	-	-
PhysNet[10]	62.5	65.7	0.297	60.3	67.5	0.326	-	-	-
PhysFormer[11]	67.7	71.0	0.070	71.5	76.9	0.078	-	-	-
Seq-rPPG	76.5	80.7	0.340	83.0	91.5	0.060	-	-	-
NoobHeart	83.9	88.4	0.340	91.0	98.8	0.236	-	-	-

4 Training on PURE

All models are trained on PURE, with the first 80% used for training and the last 20% used for validation.

4.1 Evaluation of Heart Rate

Table 7: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	5.93	11.9	0.442	5.68	11.3	0.498	-	-	-	3.12	8.78	0.897
TS-CAN[4]	4.08	9.37	0.630	5.04	10.7	0.527	-	-	-	1.03	1.63	0.996
EfficientPhys[5]	2.10	5.90	0.841	2.68	7.52	0.760	-	-	-	0.98	1.48	0.997
PhysNet[10]	1.79	6.35	0.801	1.17	3.75	0.924	-	-	-	1.02	1.65	0.996
PhysFormer[11]	1.74	5.77	0.836	1.44	4.78	0.878	-	-	-	1.66	3.11	0.988
Seq-rPPG	5.82	15.1	0.408	6.49	16.2	0.380	-	-	-	1.16	1.86	0.994
NoobHeart	2.86	8.76	0.681	3.63	10.7	0.511	-	-	-	0.97	1.45	0.997

Table 8: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	16.1	24.8	0.136	4.82	13.5	0.361	7.21	14.3	0.386	14.9	24.6	0.255
TS-CAN[4]	13.0	21.9	0.242	2.00	4.89	0.857	9.01	16.2	0.252	15.3	25.1	0.270
EfficientPhys[5]	11.1	20.4	0.374	1.96	5.58	0.804	4.65	10.9	0.605	15.8	26.1	0.237
PhysNet[10]	14.4	24.3	0.154	2.15	4.31	0.890	24.0	28.7	-0.24	14.6	24.6	0.273
PhysFormer[11]	12.8	22.6	0.260	2.74	5.53	0.827	21.0	27.2	-0.20	14.4	24.5	0.268
Seq-rPPG	16.0	27.6	0.212	0.93	1.28	0.990	7.78	15.3	0.376	15.4	26.1	0.198
NoobHeart	24.6	34.0	-0.02	9.27	15.7	0.295	21.1	28.3	-0.39	14.1	24.1	0.283

4.2 Evaluation of Heart Rate Variability (SDNN)

Table 9: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	90.6	95.0	0.119	-	-	-	50.5	60.4	0.295
TS-CAN[4]	91.2	94.3	0.164	-	-	-	36.6	47.0	0.258
EfficientPhys[5]	72.1	77.4	0.374	-	-	-	15.3	22.7	0.688
PhysNet[10]	47.1	53.6	0.229	-	-	-	21.3	26.8	0.650
PhysFormer[11]	55.2	61.6	0.332	-	-	-	21.0	24.6	0.766
Seq-rPPG	76.8	81.6	0.279	-	-	-	17.8	28.2	0.502
NoobHeart	80.2	85.1	0.359	-	-	-	36.4	40.4	0.623

5 Training on RLAP (As Large as PURE)

Trained on the RLAP training set, but only took the first 3% of the data, which means its scale is similar to PURE.

5.1 Evaluation of Heart Rate

Table 10: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	6.28	12.7	0.387	11.1	21.4	0.347	9.15	23.6	0.468	2.11	5.45	0.955
TS-CAN[4]	2.54	6.98	0.780	3.59	10.0	0.610	3.20	9.52	0.920	1.00	1.44	0.997
EfficientPhys[5]	1.08	3.74	0.932	0.97	3.50	0.940	2.52	8.89	0.930	1.01	1.48	0.997
PhysNet[10]	2.21	7.21	0.740	1.28	3.88	0.919	1.58	8.54	0.931	1.39	2.01	0.995
PhysFormer[11]	2.07	7.75	0.706	0.99	3.10	0.948	3.89	15.1	0.77	1.24	1.86	0.996
Seq-rPPG	1.71	5.75	0.852	1.48	4.92	0.880	0.42	0.60	1.000	0.92	1.46	0.997
NoobHeart	3.28	9.54	0.645	3.84	10.6	0.539	0.58	0.91	0.999	2.01	6.90	0.937

Table 11: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	15.0	23.5	0.234	2.31	5.31	0.828	9.04	15.5	0.363	16.0	25.3	0.222
TS-CAN[4]	12.9	21.9	0.246	2.37	6.87	0.715	6.98	13.3	0.464	14.7	24.7	0.270
EfficientPhys[5]	10.2	19.9	0.407	1.67	5.55	0.815	7.73	16.5	0.350	15.1	25.4	0.253
PhysNet[10]	13.3	22.9	0.209	1.39	2.20	0.971	17.3	25.3	-0.45	14.8	24.8	0.278
PhysFormer[11]	2.97	7.06	0.700	16.6	26.3	0.09	24.3	29.3	-0.38	14.2	24.0	0.302
Seq-rPPG	14.9	24.5	0.294	3.80	8.38	0.689	10.8	19.5	0.051	15.2	25.2	0.258
NoobHeart	22.8	33.5	0.038	3.39	9.76	0.531	14.0	23.9	-0.17	13.0	22.5	0.378

5.2 Evaluation of Heart Rate Variability (SDNN)

Table 12: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	95.1	100	-0.136	91.9	96.2	0.216	54.5	61.4	0.497
TS-CAN[4]	83.7	87.4	0.177	89.0	94.8	0.276	40.4	47.9	0.471
EfficientPhys[5]	47.0	58.9	0.375	27.6	46.4	0.395	9.66	14.6	0.842
PhysNet[10]	44.6	49.2	0.490	37.0	50.7	0.380	17.1	20.8	0.809
PhysFormer[11]	33.4	38.0	0.499	37.4	49.4	0.316	15.3	18.8	0.834
Seq-rPPG	54.2	63.5	0.260	32.0	48.2	0.440	13.0	20.6	0.664
NoobHeart	76.8	81.5	0.306	61.8	75.4	0.336	34.6	41.8	0.560

6 Training on RLAP (H.264 Format)

Use videos in H.264 format from RLAP to train the model, these videos are provided simultaneously with uncompressed videos and can be found in the dataset.

6.1 Evaluation of Heart Rate

Table 13: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	12.4	17.9	0.137	22.9	29.5	0.168	24.3	34.1	0.142	16.6	26.4	0.372
TS-CAN[4]	6.62	13.5	0.377	6.63	12.9	0.450	11.4	20.9	0.607	3.20	7.85	0.899
EfficientPhys[5]	15.3	21.7	-0.00	15.1	21.2	0.107	29.5	38.9	0.011	39.0	42.6	-0.10
PhysNet[10]	2.07	6.43	0.798	2.32	7.24	0.742	7.29	16.9	0.741	0.86	1.23	0.998
PhysFormer[11]	3.14	8.66	0.625	4.06	11.1	0.446	16.4	26.8	0.425	1.38	2.49	0.991
Seq-rPPG	5.70	13.9	0.414	5.58	14.1	0.422	6.98	20.3	0.612	1.83	4.45	0.977
NoobHeart	3.88	10.4	0.609	5.34	12.5	0.461	0.74	1.14	0.999	2.04	6.97	0.936

Table 14: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	18.9	26.4	0.100	12.5	22.7	-0.02	11.7	18.0	0.226	18.3	26.0	0.047
TS-CAN[4]	15.2	23.2	0.163	4.93	10.1	0.508	2.67	7.61	0.798	17.3	27.4	0.134
EfficientPhys[5]	17.7	25.1	0.141	10.5	18.4	0.053	13.0	19.4	0.059	22.9	31.0	-0.12
PhysNet[10]	12.0	20.8	0.226	3.52	8.60	0.608	4.89	12.8	0.430	16.3	26.8	0.170
PhysFormer[11]	11.9	20.1	0.180	6.62	14.3	-0.00	7.89	14.9	0.138	20.7	30.8	-0.03
Seq-rPPG	17.3	27.1	0.176	2.17	5.99	0.793	5.92	14.1	0.379	15.2	26.5	0.189
NoobHeart	24.3	34.9	-0.01	2.02	7.55	0.684	6.27	15.4	0.396	12.8	22.3	0.399

6.2 Evaluation of Heart Rate Variability (SDNN)

Table 15: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	92.7	94.8	0.079	94.3	96.8	0.108	89.0	92.5	0.256
TS-CAN[4]	87.9	91.0	0.111	97.4	101	-0.11	71.1	76.9	0.326
EfficientPhys[5]	92.4	96.3	0.027	96.2	99.2	-0.00	27.3	33.5	0.084
PhysNet[10]	55.0	59.8	0.365	68.5	78.0	0.088	21.0	27.4	0.687
PhysFormer[11]	44.5	49.8	0.411	62.1	69.4	0.106	17.2	25.6	0.707
Seq-rPPG	73.2	78.6	0.378	81.9	90.6	0.157	33.9	41.8	0.521
NoobHeart	77.5	82.9	0.364	88.4	98.0	0.300	34.6	42.6	0.485

7 Training on RLAP (Random Offset 0s to 0.2s)

Training was conducted using RLAP, but with the addition of a slight random offset. Although this asynchrony is not significant, less than 0.2 seconds, it still severely degraded performance.

7.1 Evaluation of Heart Rate

Table 16: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP			RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	11.1	17.8	0.097	21.0	28.2	-0.01	21.6	33.1	0.277	18.4	26.4	0.323
TS-CAN[4]	5.01	12.4	0.486	7.69	16.1	0.360	5.78	14.2	0.829	1.78	5.37	0.954
EfficientPhys[5]	1.68	5.47	0.863	2.04	6.78	0.798	3.92	12.7	0.865	0.89	1.27	0.998
PhysNet[10]	1.90	7.48	0.727	1.04	3.53	0.933	2.48	12.3	0.850	1.40	2.41	0.992
PhysFormer[11]	2.53	9.01	0.620	1.59	5.62	0.842	4.37	17.3	0.691	1.52	2.12	0.995
Seq-rPPG	3.55	10.7	0.507	2.54	7.23	0.759	1.90	10.8	0.887	1.20	1.73	0.995
NoobHeart	3.88	10.4	0.609	5.34	12.5	0.461	2.69	11.2	0.878	2.04	6.97	0.936

Table 17: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	MMPD			MMPD-Simplest			COHFACE			UBFC-PHYS		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	18.8	26.4	0.147	10.3	14.1	0.046	13.5	18.9	0.273	17.2	23.9	0.214
TS-CAN[4]	12.7	21.6	0.261	3.00	8.45	0.586	9.36	16.4	0.265	15.8	25.4	0.277
EfficientPhys[5]	13.3	23.4	0.213	5.34	15.4	0.185	2.66	9.27	0.703	15.7	26.1	0.207
PhysNet[10]	16.8	27.3	0.067	2.25	5.65	0.802	24.4	29.7	-0.38	14.6	24.5	0.290
PhysFormer[11]	19.7	29.6	0.012	2.47	6.63	0.743	26.6	30.5	-0.30	14.0	23.8	0.302
Seq-rPPG	20.0	30.0	0.032	4.07	8.36	0.728	18.5	26.1	-0.24	14.6	24.5	0.285
NoobHeart	24.3	34.9	-0.01	2.02	7.55	0.684	6.27	15.4	0.396	12.8	22.3	0.399

7.2 Evaluation of Heart Rate Variability (SDNN)

Table 18: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	RLAP-rPPG			PURE			UBFC-rPPG		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	90.4	93.3	0.106	94.8	97.5	0.194	99.2	103.4	-0.01
TS-CAN[4]	90.8	94.7	-0.03	88.7	93.7	0.158	57.7	69.6	0.263
EfficientPhys[5]	77.7	82.8	0.306	77.7	86.2	0.346	27.9	40.7	0.557
PhysNet[10]	33.3	40.1	0.304	26.6	40.0	0.417	16.3	19.6	0.838
PhysFormer[11]	33.4	38.9	0.335	28.7	44.8	0.274	17.8	21.6	0.783
Seq-rPPG	52.6	61.5	0.339	30.1	41.6	0.419	17.3	30.5	0.467
NoobHeart	77.5	82.9	0.364	88.4	98.0	0.300	34.6	42.6	0.485

8 Training on RLAP (Seq-rPPG Hyperparameter Analysis)

In 8x8 and 32x32 resolutions, we conducted ablation experiments for Convolution and Spectral Transformation(ST). The results are shown in Table 19 20 21. The performance of the 8x8 input is similar to that of the 32x32, but the computation and memory overheads are much smaller for the former than for the latter, and its privacy is also much stronger than that of the latter. Therefore, we finally release the 8x8-Conv-ST model.

8.1 Evaluation of Heart Rate

Table 19: The HR experimental results on RLAP, PURE, UBFC-rPPG.

Method	**RLAP**			**RLAP-rPPG**			**PURE**			**UBFC-rPPG**		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
8x8-Conv-ST	1.07	4.15	0.917	0.81	2.97	0.953	0.37	0.63	1.000	0.87	1.40	0.997
8x8-Conv	1.56	5.50	0.863	1.69	6.33	0.806	2.19	13.0	0.863	1.00	1.37	0.997
8x8-ST	4.40	14.1	0.416	3.06	10.1	0.640	0.44	0.79	0.999	1.03	1.71	0.996
32x32-Conv-ST	1.56	5.31	0.868	1.52	5.19	0.871	0.44	0.77	1.000	0.83	1.47	0.997
32x32-Conv	1.53	5.26	0.875	1.74	6.33	0.810	0.38	0.50	1.000	0.84	1.29	0.998
32x32-ST	3.06	10.8	0.613	1.99	6.47	0.806	0.53	1.04	0.999	0.96	1.55	0.997

Table 20: The HR experimental results on MMPD, COHFACE, UBFC-PHYS.

Method	**MMPD**			**MMPD-Simplest**			**COHFACE**			**UBFC-PHYS**		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
8x8-Conv-ST	14.6	25.2	0.275	1.52	3.93	0.915	16.1	25.7	-0.12	15.1	25.2	0.263
8x8-Conv	15.6	27.1	0.168	2.40	8.79	0.628	3.51	10.2	0.640	14.9	24.9	0.267
8x8-ST	19.9	30.5	0.174	6.08	15.0	0.172	26.7	32.5	-0.35	15.3	25.9	0.224
32x32-Conv-ST	16.8	27.6	0.206	1.51	3.81	0.916	11.6	21.3	-0.04	14.9	25.2	0.238
32x32-Conv	14.7	26.4	0.185	0.91	1.30	0.990	3.61	10.45	0.639	15.1	25.4	0.244
32x32-ST	18.5	29.7	0.157	3.62	10.4	0.506	21.1	28.7	-0.29	13.5	23.2	0.382

8.2 Evaluation of Heart Rate Variability (SDNN)

Table 21: The HRV experimental results on RLAP, PURE, UBFC-rPPG.

Method	**RLAP-rPPG**			**PURE**			**UBFC-rPPG**		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
8x8-Conv-ST	14.4	22.1	0.424	9.51	15.8	0.872	4.73	8.25	0.911
8x8-Conv	64.1	71.6	0.275	49.1	64.1	0.318	20.9	26.2	0.720
8x8-ST	25.3	36.2	0.217	13.1	19.5	0.746	8.25	10.6	0.880
32x32-Conv-ST	37.0	49.4	0.337	12.4	22.9	0.740	6.83	12.6	0.818
32x32-Conv	59.1	68.2	0.248	38.6	50.0	0.463	17.3	25.3	0.600
32x32-ST	28.9	40.7	21.9	14.1	24.6	0.673	5.42	7.59	0.927

9 Testing on MMPD (Different Skin Tone)

These experiments are based on different subsets of MMPD, and in order to discuss the impact of skin type, all low-light and motion samples are excluded. According to Nowara et al. [7], compression algorithms cause more severe damage to physiological signals of dark skin samples, and since MMPD uses H.264 compression, it cannot fully judge the algorithm bias on skin type. Based on the open-source program PhysRecorder, we are collaborating with another team to collect lossless video samples of dark skin; this issue will be thoroughly resolved in the future.

Skin tone 3 is the lightest while skin tone 6 is darkest.

Table 22 shows the results trained on RLAP, while Table 23 shows the results trained on PURE.

9.1 Evaluation of Heart Rate

Table 22: The HR experimental of training on RLAP.

Method	**Type 3**			**Type 4**			**Type 5**			**Type 6**		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	1.02	1.40	0.989	16.9	32.2	-0.54	5.85	7.29	0.917	6.93	10.2	0.814
TS-CAN[4]	0.95	1.34	0.990	16.6	32.4	-0.64	5.13	6.64	0.910	3.96	6.63	0.929
EfficientPhys[5]	0.92	1.30	0.990	17.5	33.0	-0.53	2.34	3.49	0.966	2.97	6.20	0.944
PhysNet[10]	0.99	1.43	0.988	22.3	33.9	-0.45	5.49	10.1	0.760	10.9	22.1	0.030
PhysFormer[11]	1.40	1.95	0.981	7.16	13.3	0.941	1.98	3.26	0.979	12.2	19.8	-0.39
Seq-rPPG	1.17	1.51	0.989	7.16	13.8	0.925	5.40	10.3	0.758	15.6	23.3	0.292
NoobHeart	2.99	8.88	0.554	29.7	39.7	-0.97	9.99	15.4	0.434	17.6	23.5	0.264

Table 23: The HR experimental of training on PURE.

Method	**Type 3**			**Type 4**			**Type 5**			**Type 6**		
	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑	MAE↓	RMSE↓	ρ↑
DeepPhys[2]	1.33	1.90	0.979	20.3	33.1	-0.75	4.50	7.61	0.873	8.55	18.6	0.137
TS-CAN[4]	1.21	1.76	0.986	20.3	33.2	-0.42	9.18	19.1	-0.20	19.2	23.7	-0.51
EfficientPhys[5]	1.68	2.49	0.963	17.8	32.5	-0.48	12.3	17.6	-0.17	1.35	1.94	0.996
PhysNet[10]	3.37	6.58	0.746	11.2	15.6	0.630	8.19	12.6	0.649	16.3	18.5	0.484
PhysFormer[11]	4.04	7.09	0.930	10.7	15.4	0.653	2.52	3.78	0.969	4.95	7.80	0.917
Seq-rPPG	0.99	1.26	0.991	18.1	32.8	-0.48	5.31	9.92	0.774	21.6	30.8	-0.10
NoobHeart	13.2	18.9	0.890	21.9	26.9	-0.46	10.8	16.7	0.353	25.9	31.5	-0.01

References

[1] Serge Bobbia, Richard Macwan, Yannick Benezeth, Alamin Mansouri, and Julien Dubois. Unsupervised skin tissue segmentation for remote photoplethysmography. Pattern Recognition Letters, 124:82–90, 2019.

[2] Weixuan Chen and Daniel McDuff. Deepphys: Video-based physiological measurement using convolutional attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), September 2018.

[3] Guillaume Heusch, André Anjos, and Sébastien Marcel. A reproducible study on remote heart rate measurement. arXiv, September 2017. URL https://arxiv.org/abs/1709.00962.

[4] Xin Liu, Josh Fromm, Shwetak Patel, and Daniel McDuff. Multi-task temporal shift attention networks for on-device contactless vitals measurement. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, Advances in Neural Information Processing Systems, volume 33, pages 19400–19411. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/file/e1228be46de6a0234ac22ded31417bc7-Paper.pdf.

[5] Xin Liu, Brian L. Hill, Ziheng Jiang, Shwetak Patel, and Daniel McDuff. Efficientphys: Enabling simple, fast and accurate camera-based vitals measurement, 2021. URL https://arxiv.org/abs/2110.04447.

[6] Rita Meziati Sabour, Yannick Benezeth, Pierre De Oliveira, Julien Chappe, and Fan Yang. Ubfc-phys: A multimodal database for psychophysiological studies of social stress. IEEE Transactions on Affective Computing, pages 1–1, 2021. doi: 10.1109/TAFFC.2021.3056960.

[7] Ewa M. Nowara, Daniel McDuff, and Ashok Veeraraghavan. A meta-analysis of the impact of skin type and gender on non-contact photoplethysmography measurements. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 1148–1155, 2020. doi: 10.1109/CVPRW50498.2020.00150.

[8] Ronny Stricker, Steffen Müller, and Horst-Michael Gross. Non-contact video-based pulse rate measurement on a mobile service robot. In The 23rd IEEE International Symposium on Robot and Human Interactive Communication, pages 1056–1062, 2014. doi: 10.1109/ROMAN.2014.6926392.

[9] Jiankai Tang, Kequan Chen, Yuntao Wang, Yuanchun Shi, Shwetak Patel, Daniel McDuff, and Xin Liu. Mmpd: Multi-domain mobile video physiology dataset, 2023.

[10] Zitong Yu, Xiaobai Li, and Guoying Zhao. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks. arXiv preprint arXiv:1905.02419, 2019.

[11] Zitong Yu, Yuming Shen, Jingang Shi, Hengshuang Zhao, Philip Torr, and Guoying Zhao. Physformer: Facial video-based physiological measurement with temporal difference transformer. In CVPR, 2022.