λ	LFW	CFP-FP	AgeDB	CPLFW	VGG2-FP
0.1	98.43	93.53	87.63	84.12	91.62
1	99.23	95.57	91.97	87.03	93.28
10	**99.43**	**97.50**	**94.43**	**90.55**	**94.56**

Table 4: Results of our models trained with different weights for KL divergence (δ is 16). Training data: CASIA-WebFace.

δ	LFW	CFP-FP	AgeDB	CPLFW	VGG2-FP
8	99.30	97.11	93.75	89.33	94.38
16	**99.43**	**97.50**	**94.43**	**90.55**	**94.56**
32	99.33	97.04	93.47	89.62	93.70

Table 5: Results of our models trained with different values for degree of freedom (λ is 10). Training data: CASIA-WebFace.

of different models are shown in Fig. **??**. We can see that our proposed RTS can better discriminate the in-distribution and out-of-distribution data. Both quantitative and distributions results indicate that RTS is an effective technique to model uncertainty and complete OOD detection task.

4.5 Face Recognition Performance

Face verification accuracy on benchmarks. We compare RTS with the state-of-the-art uncertainty methods including PFE, GODIN, DUL and Relaxed Softmax, and related non-uncertainty methods including Baseline (ArcFace), MagFace and AdaFace. The results of recognition is shown in Table **??**. We can see that the recognition performance of RTS is comparable with the state-of-the-art methods on all test sets. This indicates that, besides the ability to reveal uncertainty of images and detect out-of-distribution data, the model trained with RTS can achieve competitive performances in face recognition task.

Balance Between FR and OOD Performance. Fig. **??** shows the performance between face recognition (average verification accuracy of all benchmarks) and out-of-distribution detection (AUC) of non-uncertainty and uncertainty models. The comprehensive performances of MagFace, AdaFace and Relaxed Softmax are both close to that of RTS. While MagFace has a great many hyper parameters needed to be adjusted manually and is difficult to reproduce good enough results. Besides, the convergence of Relaxed Softmax is very sensitive to the margin in its prediction head. In comparison, RTS has less hyper parameters and is easier to converge. Our method achieves the best performance in OOD detection task and comparably high enough face verification accuracy, demonstrating that RTS is a unified framework for uncertainty estimation and face recognition.

4.6 Ablation Study

Effects of KL divergence. The KL divergence loss works as a regularization term to prevent the uncertainty scale from growing infinitely. When the weight $\lambda < 0.1$, the model have difficulty in converging, and the performance also deteriorates at last. For large $\lambda(> 10)$, the model tends to predict nearly constant variance $v(\boldsymbol{x})$, which has little effects in modeling data uncertainty. We conduct experiments on models with $\lambda \in \{0.1, 1, 10\}$. The results are shown in Table **??**. Through experiments, we find that the model achieves the best performance when $\lambda = 10$. Thus, we set $\lambda = 10$ for our RTS model.

Effects of degree of freedom. We study the effects of degree of freedom δ, which is a hyperparameter of RTS. Intuitively, δ determines the shape of density of random temperature, t. The results are shown in Table **??**. From experimental results, we can see that RTS achieves the best performance when $\delta = 16$. Thus, we set $\delta = 16$ for our RTS model.

5 Conclusion

In this paper, we first analysis the connection between temperature scaling and uncertainty modeling in the classification model. Taking a probabilistic view, the temperature scalar is exactly the scale of uncertainty noise implicitly added in the softmax function. Based on this observation, a unified framework, Random Temperature Scaling (RTS), is proposed for uncertainty estimation and face recognition by modeling the uncertainty level by a stochastic distribution.

Experiments show that RTS can adjust the learning strength of different quality samples for stability and accuracy during training. The magnitude of variance in RTS acts as a metric to reveal the image quality and can be used to detect uncertain, low-quality and even OOD samples in testing phase. Face recognition models trained with RTS have higher security and reliability by rejecting untrusted images, especially when deployed in real-world face recognition systems. RTS achieves top performance on both FR and OOD detection tasks. Moreover, models trained with RTS performs robustly on datasets with noise. The proposed module is light-weight and only adds negligible computation cost to the original face recognition model.

Appendix

Cum sit harum inventore facilis dicta omnis odio illum eveniet porro dolores, cupiditate ad temporibus ratione enim.Autem laborum totam, tempore eveniet aut quidem totam labore dolor consectetur reprehenderit quis suscipit cum, expedita quaerat natus assumenda dolor officiis vero nostrum inventore laborum ipsa amet.Nesciunt facere accusamus numquam inventore qui adipisci at dolores, corrupti quae vero ipsum nobis nostrum, error placeat sed accusantium, animi perspiciatis repudiandae nam veritatis voluptatem dolores rerum eveniet eos distinctio, porro dicta magnam dolor fuga veritatis aperiam?Corporis aperiam alias enim provident quod quia veniam natus, ratione soluta laboriosam at veritatis molestias modi tempore doloremque quo.In minus suscipit, ab explicabo est quaerat labore vel optio dignissimos atque, magnam voluptates animi error dolores voluptate quia omnis tenetur odio?Ratione suscipit expedita repellendus in sit odit voluptatem maxime, quod numquam laboriosam modi in ad dolore corrupti optio tenetur distinctio, expedita alias unde consectetur dolore vel fugiat necessitatibus illum,