Supplementary Material for
Detecting Deepfakes with Self-Blended Images

Kaede Shiohara Toshihiko Yamasaki
The University of Tokyo
{shiohara, yamasaki}@cvm.t.u-tokyo.ac.jp

1. Additional Experiments

To show the validity and generality of our proposed method, we conduct some additional experiments using BI [9]. We implement it following the author's code [3].

Landmark Similarity. In BI, 100 images with the closest facial landmarks to the base image are used as candidate source images. We train EfficientNet-B4 [12] (EFNB4) on all 100 ($BI_{1\text{-}100}$) and on the top 20 ($BI_{1\text{-}20}$). As shown in Table 8(a), the model trained on $BI_{1\text{-}20}$ outperforms the model trained on original BI on CDF and DFDCP, and is on par with on DFD, DFDC, and FFIW. This result indicates that, at least in the landmark similarity, easy samples with low similarity do not contribute to the model generality.

Joint Training of SBIs and BI. To explore the best practice for more general deepfake detection, we train EFNB4 on the joint dataset of SBIs and BI that are each sampled with the probability of 0.5. The results of joint-training are basically lower than that of our proposed SBIs as shown in Table 8(b).

Source-Target Augmentation in BI. To demonstrate the superiority of our two ideas of (1) blending identical images and (2) augmenting source and target images, we incorporate our source-target augmentation into BI. As shown in Table 8(c), the results are lower than that of SBIs on four out of the five test sets, although they are better than that of the original BI, which indicates our two ideas are both important for general deepfake detection.

2. Comprehensive Results

We provide comprehensive results of our method including video-level area under the receiver operating characteristic curve (AUC) and average precision (AP). We also describe the number of real and fake videos. We additionally evaluate our method on FaceShifter [8] (FSh) and DeeperForensics1.0 [7] (DF1.0) datasets. On DF1.0, we use c23 (lightly compressed) real videos of FF++ following the convention. The result is given in Table 9.

Method	Test Set AUC (%)				
	CDF	DFD	DFDC	DFDCP	FFIW
(a) Effect of Landmark Similarity					
$BI_{1\text{-}100}$ (original)	69.40	**97.50**	**66.55**	68.71	**85.69**
$BI_{1\text{-}20}$	**71.44**	97.27	65.27	**68.80**	85.35
(b) Effect of Joint-Training of SBIs and BI					
SBIs (Ours)	**93.18**	97.56	**72.42**	**86.15**	84.83
BI	69.40	97.50	66.55	68.71	**85.69**
SBIs + BI	89.36	**98.34**	71.87	82.92	83.53
(c) Effect of Source-Target Augmentation in BI					
SBIs (Ours)	**93.18**	97.56	**72.42**	**86.15**	84.83
BI	69.40	97.50	66.55	68.71	**85.69**
BI w/ S-T Aug.	78.99	**99.04**	71.98	74.25	83.14

Table 8. **Cross-dataset evaluation results of additional experiments**.

Test Set			Metrics	
Database	#Real	#Fake	AUC(%)	AP(%)
DF [2]	140	140	99.99	99.99
F2F [14]	140	140	99.88	99.89
FS [4]	140	140	99.91	99.91
NT [13]	140	140	98.79	99.15
FF++ [11]	140	560	99.64	99.92
DFD [1]	363	3068	97.56	99.70
FSh [8]	140	140	98.27	98.24
DF1.0 [7]	140	140	83.14	85.06
CDF [10]	178	340	93.18	96.35
DFDC [5]	2500	2500	72.42	75.17
DFDCP [6]	276	501	86.15	91.37
FFIW [15]	250	250	84.83	84.30

Table 9. **Comprehensive results and statistical details.**

References

[1] Contributing data to deepfake detection research. https://ai.googleblog.com/2019/09/contributing-data-to-deepfake-detection.html. Accessed: 2021-11-13. 1

[2] Deepfakes. https://github.com/deepfakes/faceswap. Accessed: 2021-11-13. 1

[3] Face x-ray. https://github.com/AlgoHunt/Face-Xray. Accessed: 2021-11-13. 1

[4] Faceswap. https://github.com/MarekKowalski/FaceSwap/. Accessed: 2021-11-13. 1

[5] Brian Dolhansky, Joanna Bitton, Ben Pflaum, Jikuo Lu, Russ Howes, Menglin Wang, and Cristian Canton Ferrer. The deepfake detection challenge dataset. *arXiv:2006.07397*, 2020. 1

[6] Brian Dolhansky, Russ Howes, Ben Pflaum, Nicole Baram, and Cristian Canton Ferrer. The deepfake detection challenge (dfdc) preview dataset. *arXiv:1910.08854*, 2019. 1

[7] Liming Jiang, Ren Li, Wayne Wu, Chen Qian, and Chen Change Loy. DeeperForensics-1.0: A large-scale dataset for real-world face forgery detection. In *CVPR*, 2020. 1

[8] Lingzhi Li, Jianmin Bao, Hao Yang, Dong Chen, and Fang Wen. Advancing high fidelity identity swapping for forgery detection. In *CVPR*, 2020. 1

[9] Lingzhi Li, Jianmin Bao, Ting Zhang, Hao Yang, Dong Chen, Fang Wen, and Baining Guo. Face x-ray for more general face forgery detection. In *CVPR*, 2020. 1

[10] Yuezun Li, Xin Yang, Pu Sun, Honggang Qi, and Siwei Lyu. Celeb-df: A large-scale challenging dataset for deepfake forensics. In *CVPR*, 2020. 1

[11] Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Niessner. Faceforensics++: Learning to detect manipulated facial images. In *ICCV*, 2019. 1

[12] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *ICML*, pages 6105–6114, 2019. 1

[13] Justus Thies, Michael Zollhöfer, and Matthias Nießner. Deferred neural rendering: Image synthesis using neural textures. *ACM TOG*, 2019. 1

[14] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner. Face2face: Real-time face capture and reenactment of rgb videos. In *CVPR*, pages 2387–2395, 2016. 1

[15] Tianfei Zhou, Wenguan Wang, Zhiyuan Liang, and Jianbing Shen. Face forensics in the wild. In *CVPR*, 2021. 1