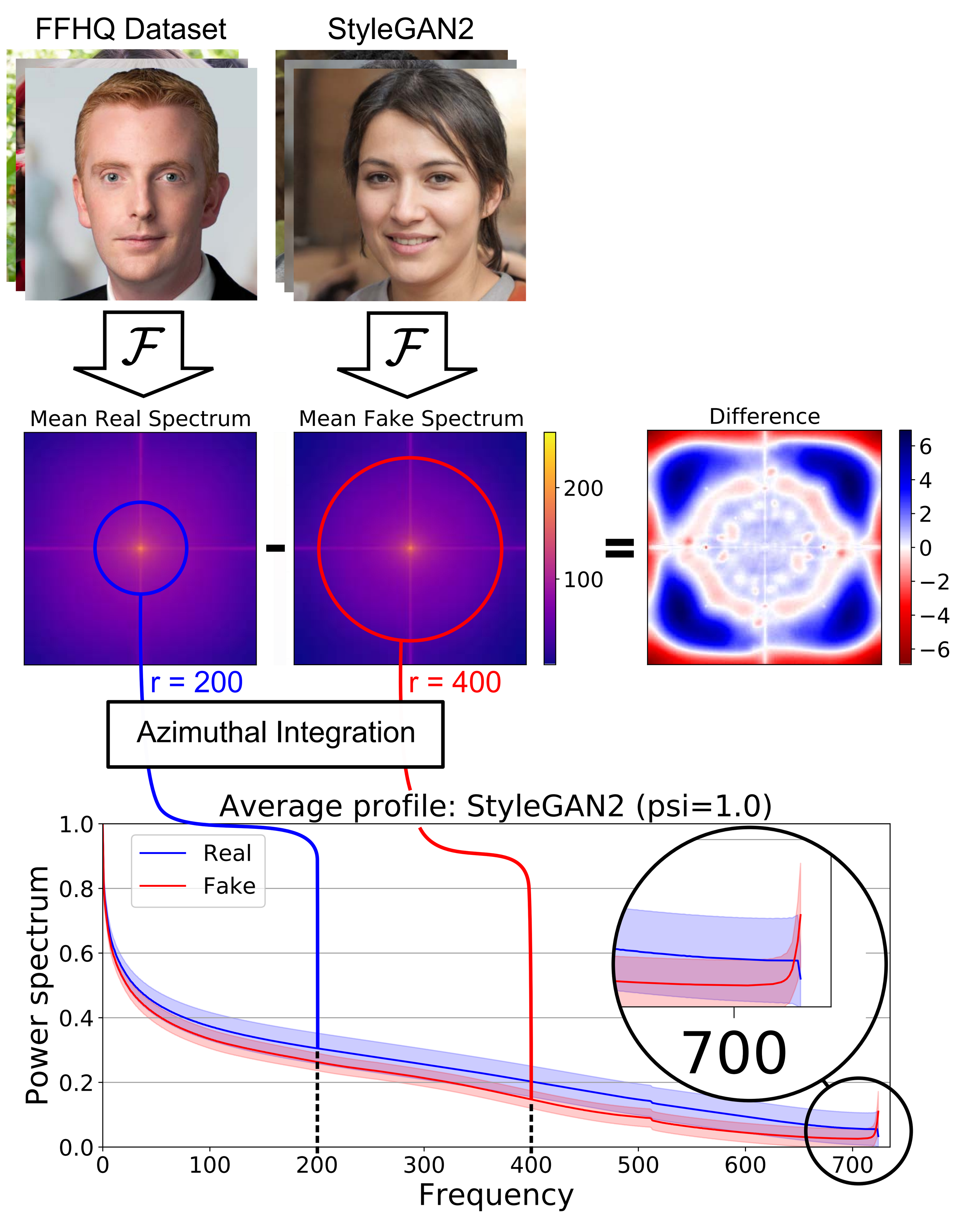
Spectral Distribution Aware Image Generation

Steffen Jung, Margret Keuper

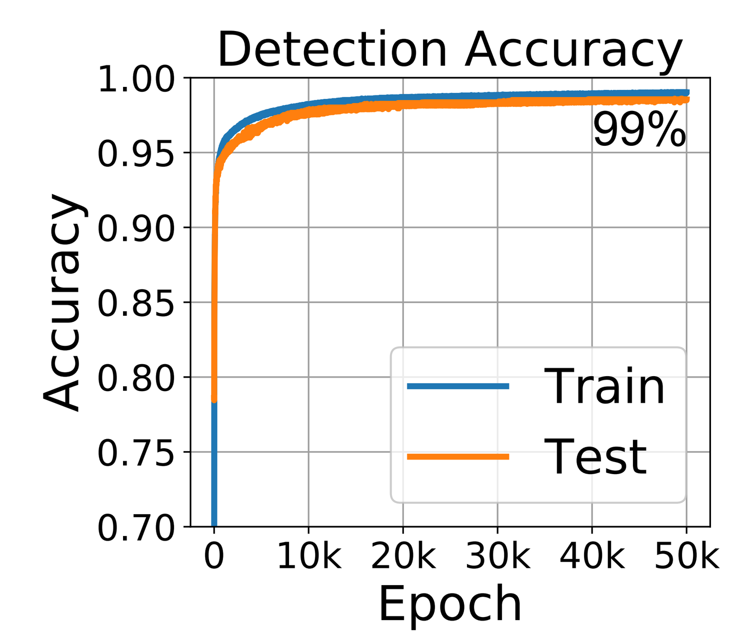
1) The Problem: Spectral Fidelity

Commonly used generative adversarial networks are not able to learn the distribution of training images in the spectral domain. Projecting the 2D Fourier transformed images into a 1D representation (spectral profiles) via azimuthal integration (AI) reveals that generated images contain high frequency artifacts.



2) Fake Detection in the Spectral Domain

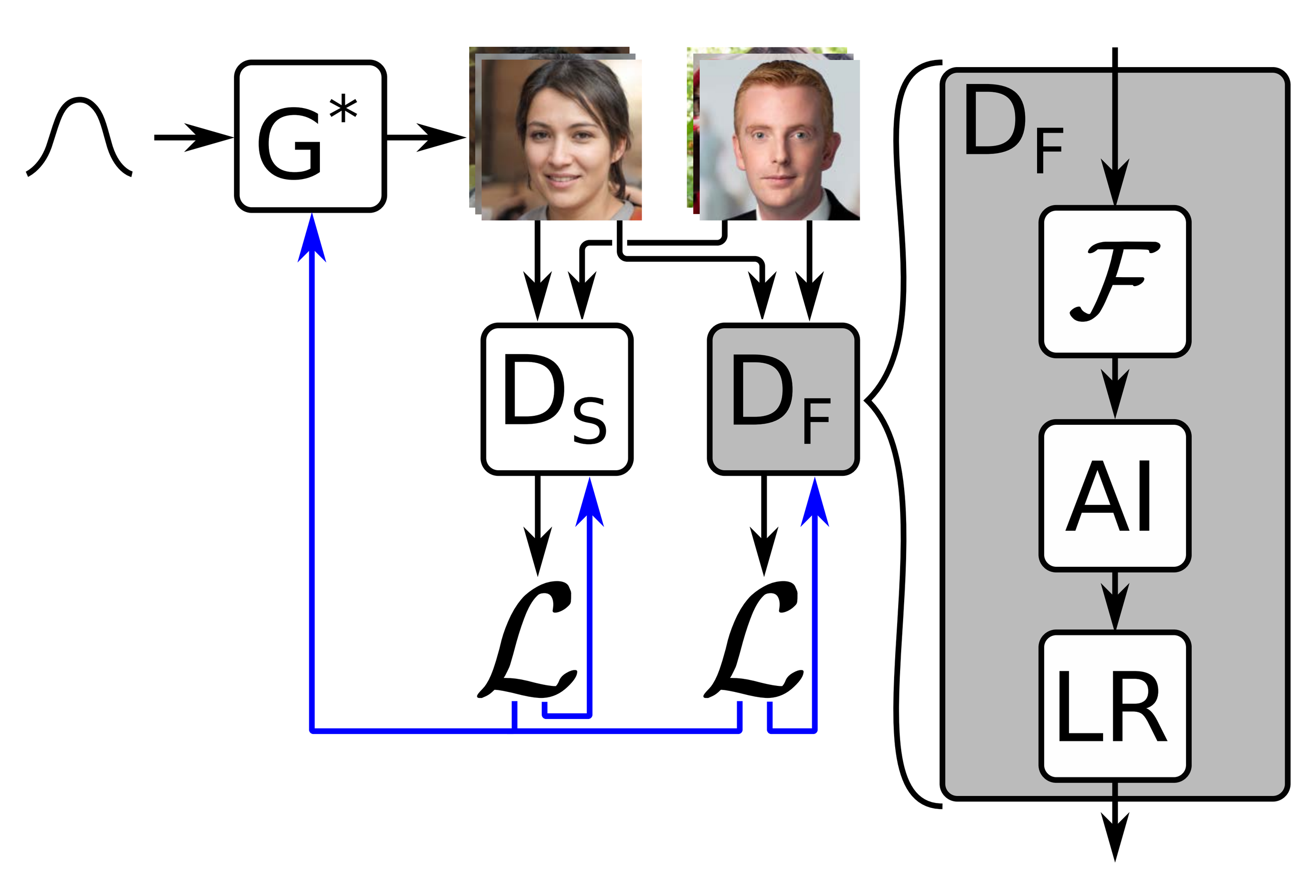
Recent methods [1] in fake detection exploit the flawed spectral distribution to achieve impressive results. Training a logistic regression with spectral profiles of images generated by StyleGAN2 and the respective training data achieves a detection accuracy of above 99%. To increase spectral fidelity, [1] provide a penalty term for the generator network. However, our experiments show that this penalty term leads to unstable training and decreased image quality.



[1] R. Durall, M. Keuper and J. Keuper. **Watch your Up-Convolution: CNN Based Generative Deep Neural Networks are Failing to Reproduce Spectral Distributions.** In CVPR 2020.

3) Our Approach: Spectral Discriminator

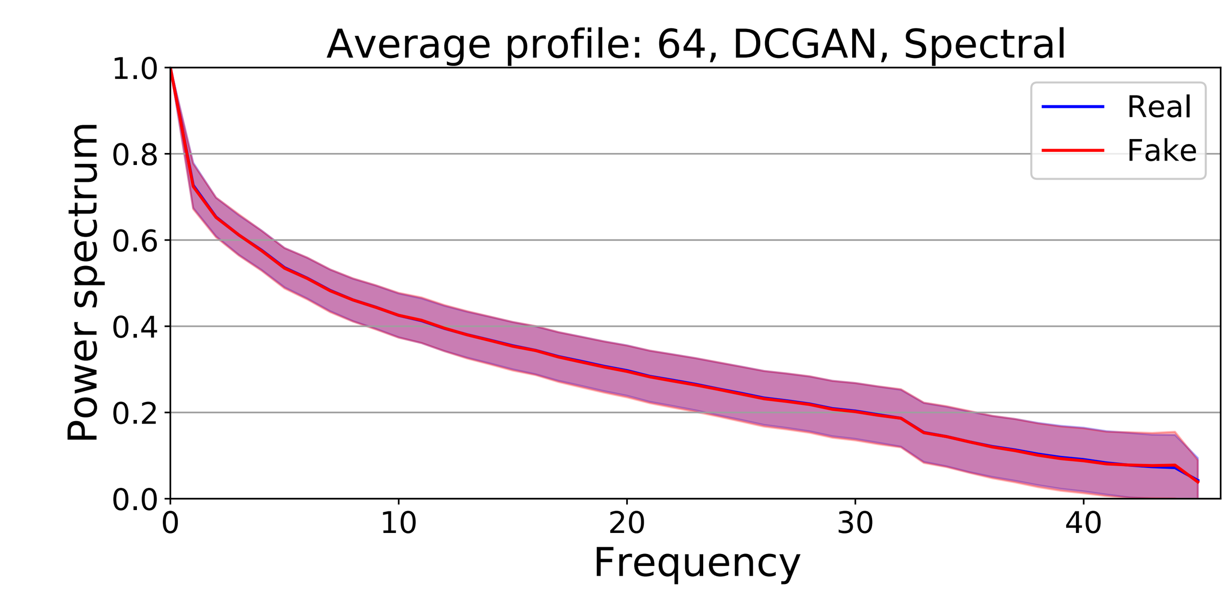
Our approach to increase spectral fidelity incorporates an additional discriminator into the training process. This spectral discriminator (D_F) transforms images into spectral profiles using them as input to a single fully-connected layer. Hence, our approach is lightweight, adding only O(n) additional parameters (n = longest image side). The discriminator can be combined with any spatial discriminator (D_S), and therefore provides a versatile plug-in method to increase spectral fidelity.

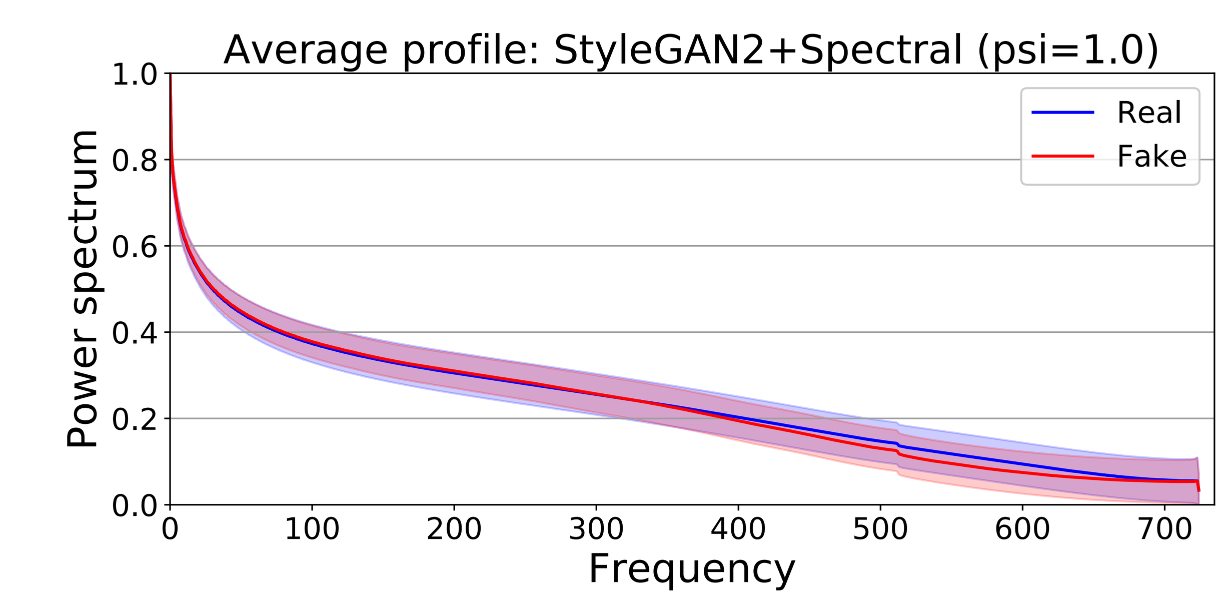


4) Results

To show the effectiveness and versatility of our approach we train models using different loss functions (DCGAN, LSGAN, WGAN, WGAN-GP) and different image resolutions (64^2 in table below, $128^2/256^2$ in paper). We additionally finetune a StyleGAN2 model to show scalability to SOTA architectures. We evaluate the spatial image quality via Fréchet inception distance (FID) and the spectral fidelity using two new metrics. First, we compute the L_1 distance between means of the spectral profiles of training and generated images and call this metric spectral difference (SD). Second, we train a logistic regression as described in 2) and translate the resulting detection accuracy into a score. We call this metric cloaking score (CS). In the paper we also show how our approach affects different fake detection methods.

	Model	FID↓	SD↓	CS↑
64	DCGAN	**15.257**	1.293	0.16
	DCGAN + [1]	29.875	0.311	0.25
	DCGAN, ours	15.591	**0.042**	**0.84**
	LSGAN	15.518	0.468	0.04
	LSGAN, ours	**15.515**	**0.041**	**0.86**
	WGAN	**47.704**	1.291	0.01
	WGAN, ours	47.948	**0.029**	**0.85**
	WGAN-GP	**39.404**	0.575	0.18
	WGAN-GP, ours	39.441	**0.252**	**0.54**
1024	StyleGAN2	**2.733**	32.250	0.09
	StyleGAN2, ours	3.326	**5.978**	**0.24**





5) Conclusion

Our approach increases spectral fidelity and provides:

- **stable** training
- **lightweight**: small number of additional parameters
- **versatililty**: different loss functions, scales to high resolutions