

Frequency-Domain Analysis of Traces for the Detection of AI-based Compression

Sandra Bergmann¹, Denise Moussa^{1,2}, Fabian Brand¹, André Kaup¹ and Christian Riess¹

¹Friedrich-Alexander University Erlangen-Nürnberg, Germany

²Federal Criminal Police Office (BKA), Germany

{sandra.daniela.bergmann, denise.moussa, fabian.brand, andre.kaup, christian.riess}@fau.de

Abstract—The JPEG algorithm is the most popular compression method on the internet. Its properties have been extensively studied in image forensics for examining image origin and authenticity. However, the JPEG standard will in the near future be extended with AI-based compression. This approach is fundamentally different from the classic JPEG algorithm, and requires an entirely new set of forensics tools.

As a first step towards forensic tools for AI compression, we present a first investigation of forensic traces in HiFiC, the current state-of-the-art AI-based compression method. We investigate the frequency space of the compressed images, and identify two types of traces, which likely arise from GAN upsampling and in homogeneous areas. We evaluate the detectability on different patch sizes and unseen postprocessing, and report a detectability of 96.37%. Our empirical results also suggest that further, yet unidentified, compression traces can be expected in the spatial domain.

Index Terms—AI-based compression, frequency analysis

I. INTRODUCTION

Image forensics aims to reconstruct cues about origin and authenticity of an image. JPEG forensics is one of the most extensively studied subfields, since two decades [1]–[3].

However, “good, old-fashioned” JPEG forensics may require an update soon, since highly efficient AI-based image compression methods have recently emerged. In fact, the JPEG committee even decided to standardize the newly developed JPEG AI format, which is expected to become available in the next few years [4].

Different AI-based compression methods have been proposed. Early works are based on RNNs [5]. Many recent methods use autoencoders, a type of neural network that is trained to encode and reconstruct (decode) its input [6]. One of the most prominent methods is High-Fidelity Generative Image Compression (HiFiC), which currently constitutes the state-of-the-art. HiFiC is based on a convolutional encoder with hyper-prior and a generative adversarial network (GAN) as decoder. It can compress the data at three quality levels: high (HiFiC-Hi), middle (HiFiC-Mi) and low (HiFiC-Lo) [7].

Berthet *et al.* presented a first forensic experiment on AI compression [8]. They investigated the generalization ability of existing JPEG-based forensic image manipulation detectors to HiFiC. At the example of the CAT-Net [9], the authors show that JPEG-based detectors perform poorly on AI-compressed

inputs. This demonstrates the need for novel forensic methods for AI compression. In this spirit, Bhowmik *et al.* investigated source identification on AI-compressed images [10].

Despite those pioneering works, there is to the best of our knowledge no investigation of the actual traces of AI compression. We argue that knowledge about the fingerprint of AI compression methods can be a vital ingredient for building reliable and interpretable forensic detectors. In this work, we want to close this gap. We show that the state-of-the-art method HiFiC leaves characteristic artifacts in the frequency-domain, and we empirically investigate the detectability of these artifacts. Our specific contributions are:

- We expose two types of HiFiC artifacts in frequency domain. They likely stem from upsampling in the HiFiC GAN, and from homogeneous regions.
- We further characterize these artifacts, and show that they behave consistently and well predictably.
- We evaluate the artifact detectability under various types of unseen postprocessing. Furthermore, we present evidence that there are likely further traces in spatial domain that can be exploited for detecting AI compression.

The paper is organized as follows: Section 2 summarizes the related work. In Sec. 3, we briefly describe the HiFiC compression method. In Sec. 4, we analytically investigate HiFiC compression artifacts in the frequency domain. Section 5 shows our experiments with respect to the artifacts’ robustness. We provide conclusions and an outlook in Sec. 6.

II. RELATED WORK

The compression history of an image can provide cues about its provenance and potentially also about manipulations. The literature on compression artifacts in traditional JPEG images is vast, such that this brief review must necessarily omit many excellent works.

Many works aim to distinguish single-compressed from double-compressed images [1], [2]. Later works also addressed very difficult cases, for example when the same quantization table is used for recompression [11], when the JPEG grid of the first and second compression are not aligned [12], or when resampling occurs between both compression steps [3]. A popular statistics for detecting double compression is Benford’s Law, which uses the distribution of the first digits of JPEG coefficients [13], [14]. With the rise of deep learning applications, several deep learning works also address double

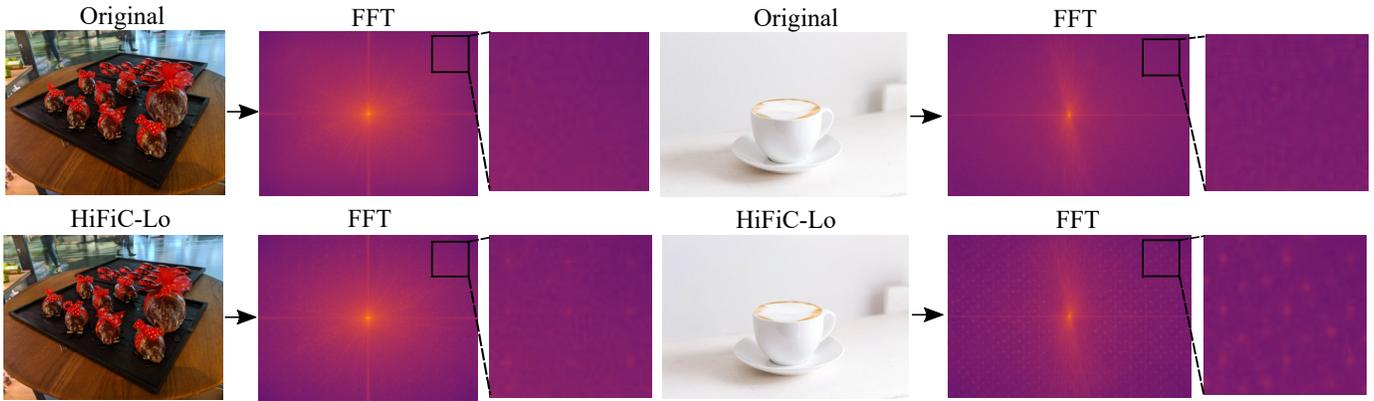


Fig. 1: Example images from the CLIC2020 dataset. AI-compressed images on the bottom exhibit frequency-space artifacts that are not found in the original images on top. Left are upsampling artifacts, right are homogeneous-area artifacts.

compression detection [12], [15], [16]. For example, a state-of-the-art network is CAT-Net [9].

JPEG compression also leaves other traces in the image that can be detected. For example, rounding errors can occur due to the operator used in JPEG compression, which leave artifacts in the image and can be localized [17]. Another work has dealt with the so-called “dimples” JPEG artifacts, which are caused by the mathematical operator to convert the discrete cosine transform (DCT) coefficients from float to integer [18]. JPEG library related artifacts can also arise during compression. Furthermore, the JPEG library libjpeg leaves a high frequency periodic pattern in chroma subsampling [19]. Since the literature shows that JPEG compression leaves traces in the image, it is important to investigate this for AI-based compression as well.

Another line of research picked up frequency-related artifacts in image forensics, namely the detection of GAN-generated images. The number and quality of such images is improving so much that a human can no longer identify between real and fake [20]. The detection of GAN-generated images often involves analyzing the frequency domain of the image [21]. GAN generators leave fingerprints in their images that can be used to detect a GAN-generated image and even to distinguish the originating generator network from these traces [22], [23]. Frank *et al.* link the fingerprints in the frequency domain to the GAN upsampling operators [22]. These fingerprints are not limited to GANs. For example, it has been shown that the recently developed diffusion models exhibit similar artifacts in the frequency domain [24], [25]. Ricker *et al.* attribute these artifacts to a training objective that performs suboptimally at processing high frequencies [25].

The AI-based compression algorithm HiFiC uses a conditional GAN, with an additional latent space. Hence, we use the GAN fingerprinting works as a guide to forensically examine HiFiC compression. Our analysis uses the fast Fourier transform (FFT) to expose two salient compression artifacts. While we hope that these findings will be of forensic use, we also note that this analysis only draws a partial picture of HiFiC artifacts, because our empirical evaluation also indicates

that there may be further traces in spatial domain.

III. HIGH-FIDELITY GENERATIVE IMAGE COMPRESSION

HiFiC compression uses internally a conditional GAN. Like standard GANs, this architecture optimizes during training a generator subnetwork and a discriminator subnetwork. Additionally, a generative model is learned with a conditional probability that uses for each data point an additional piece of information, like a class or a label [7].

HiFiC compresses an image \mathbf{x} with a quantized latent encoder $\mathbf{y} = E(\mathbf{x})$. Duan *et al.* showed that the latent space \mathbf{y} behaves similarly to an orthogonal transform like the DCT to capture spatial dependencies in the image, and in particular to model the amount of texture in a region [26]. To additionally guide this representation, additional hyperpriors \mathbf{z} are introduced as side information for the entropy model of the encoder [6].

Decompression is done with a decoder that is conditioned on \mathbf{z} . The specific implementation uses a GAN generator that reconstructs from \mathbf{y} the image $\mathbf{x}' = G(\mathbf{y})$. Note that even though the visual quality of the decompressed image is overall quite high, the overall compression method is lossy, i.e., $\mathbf{x}' \neq \mathbf{x}$ [7].

HiFiC is published with results on images from the CLIC2020 [27], DIV2k [28] and Kodak [29] datasets¹. Throughout this work, we use the provided implementation and compressed images in order to characterize the original method as closely as possible.

IV. FREQUENCY DOMAIN ANALYSIS

The frequency domain of HiFiC-compressed images reveals various artifacts that do not appear in natural images. These artifacts are shown in two example images from the CLIC2020 dataset [27] in Fig. 1. On top, images and spectra of the uncompressed source are shown. On bottom, the images and spectra of their HiFiC-Lo compressed counterparts are shown. It can be observed that the perceptual quality of the compressed image is quite high in the spatial domain. However,

¹<https://hific.github.io/raw/index.html>

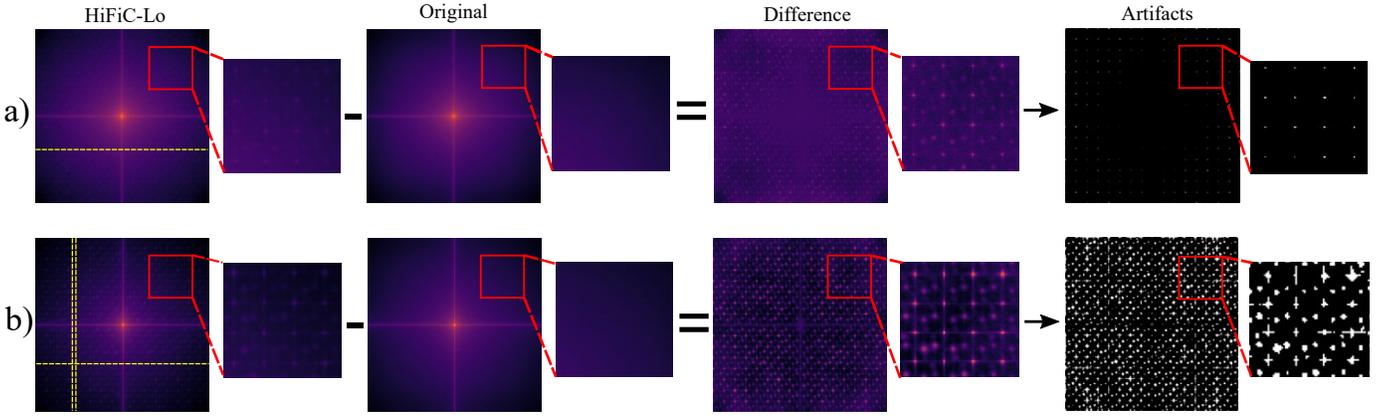


Fig. 2: Analysis of the frequency domain. (a) Difference of spectra from original and HiFiC-Lo compressed images exhibits grid artifacts. (b) Difference of spectra taken exclusively from homogenous image regions exhibits a second type of artifact. The yellow lines indicate the slices for the plots in Fig. 3 (a) and Fig. 4.

the frequency domain exhibits a grid pattern. The mostly homogeneous image on the right exhibits an additional artifact.

It is highly plausible that an upsampling operation is the cause of the grid pattern artifacts, because the generator of a GAN is used for decoding AI-compressed images. To confirm this assumption and to localize the grid pattern, further investigations will be performed in the following Sec. IV-A.

The second artifact is shown in Fig. 1 in the homogeneous image on the right. To the best of our knowledge, this artifact is not yet known in the literature. We hypothesize that this artifact is particular to the compression of this network in low-texture regions. For example, it may be the case that this artifact arises for particular values of the hyperprior \mathbf{z} that captures different spatial dependencies in the image. The decoder is conditioned on \mathbf{z} , which may ultimately introduce the artifact upon reconstruction of the image. We analyze this artifact by examining local image regions of images with many homogeneous areas in Sec. IV-B.

A. Analysis of Upsampling Artifacts

The described artifacts exhibit a relatively weak signature in frequency domain. This signature can be slightly enhanced by averaging the spectra of multiple images. To do so, we select 45 images from the CLIC2020 dataset [27]. A 1000×1000 pixels patch is cropped from each image, converted to grayscale, and represented by the log-magnitude of its Fourier coefficients (here and throughout the paper, the number of used Fourier coefficients corresponds to the number of pixels).

Such mean spectra are created for the images in uncompressed format and for the images with HiFiC-Lo compression (to create very pronounced artifact signatures). Subsequently, these two mean spectra are subtracted to highlight the artifacts. Figure 2 shows these spectra and the resulting difference spectra. Additionally, binarizations of the difference spectra are shown that highlight the artifacts. Figure 2 a) shows the upsampling artifacts. The binary image indicates that the peaks of the grid pattern occur at certain intervals. For a $N \times N$ image, the artifacts are periodic and appear in the

zero-centered frequency domain at the locations $(u, v) = (i \cdot (N/16), j \cdot (N/16))$ for $-8 \leq i, j \leq 8$, where u and v denote the horizontal and vertical frequencies, respectively.

Figure 3 further characterizes the observed peaks and the behavior during recompression. Figure 3 a) plots in 1-D the HiFiC-Lo and HiFiC-Hi coefficients through the average FFT spectra along the horizontal yellow line of Fig. 2 a). It can be seen that HiFiC-Lo creates in most cases higher peaks than HiFiC-Hi. This observation is confirmed in the other line plots and suggests that HiFiC-Lo creates artifacts that are structurally similar but stronger than artifacts of HiFiC-Hi. Additionally, the line plot also confirms that the grid pattern occurs at the previously noted periodic intervals.

One may wonder whether the structure of the artifacts changes when iterating the compression, which can be seen as an analogy to the study of double-JPEG compression in image forensics. Repeated AI compression is examined in Fig. 3 b) to Fig. 3 e). These plots show a horizontal slice through the averaged FFT spectra of 260×260 pixel patches from 20 images. They show three cases, namely a HiFiC-Hi image recompressed with HiFiC-Hi, a HiFiC-Hi image recompressed with HiFiC-Lo, and a HiFiC-Lo image recompressed as HiFiC-Hi. In the first case, Fig. 3 b) shows that the spectra including the peaks are almost identical between HiFiC-Hi images and their HiFiC-Hi-recompressed version. In the second case, Fig. 3 c) shows that recompressing a HiFiC-Hi image with HiFiC-Lo enlarges the peaks. In the third case, Fig. 3 d) shows that HiFiC-Hi recompressed with HiFiC-Lo is virtually identical to HiFiC-Lo, which likely makes these spectra indistinguishable. This is consistent with the previous observations in Fig. 3 b) and c). Fig. 3 e) shows HiFiC-Lo followed by HiFiC-Hi, and in comparison only HiFiC-Lo and only HiFiC-Hi. It is somewhat surprising that the height of the peaks for the double compression are between the two single compressions. One might have expected that the primary Lo compression strongly biases the subsequent Hi compression. However, this is not the case here, hence both compressions

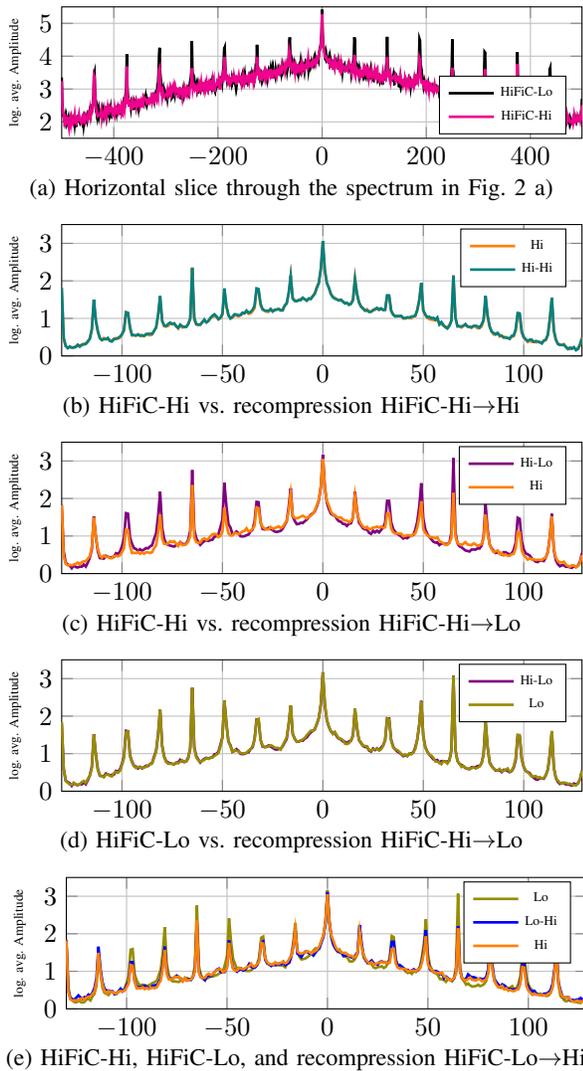


Fig. 3: Line spectra for upsampling artifacts after compressions HiFiC-Hi, HiFiC-Lo, and their recompression combinations. The x-axis shows horizontal frequency u .

must be treated as similar-but-different operations.

B. Analysis of Homogeneous Artifacts

Images with many homogeneous structures exhibit another type of artifact. In order to analyze these more precisely, we hand-selected 20 mostly homogeneous images from the datasets by Mentzer *et al.* [7]. Homogeneous patches of 260×260 patches are cut from the images. These patches are averaged and subtracted analogously to the previous processing to expose the difference in spectra of AI-compressed and original patches. The resulting spectra are shown in Fig. 2 b). The artifact pattern is almost uniformly distributed across the entire frequency domain. Each pattern consists of a pair of nearby peaks in horizontal direction.

Line plots for this pattern are shown in Fig. 4. The line plots are extracted along the yellow lines in Fig. 2 b). Horizontal FFT coefficients are shown in Fig. 4 a). The vertical FFT

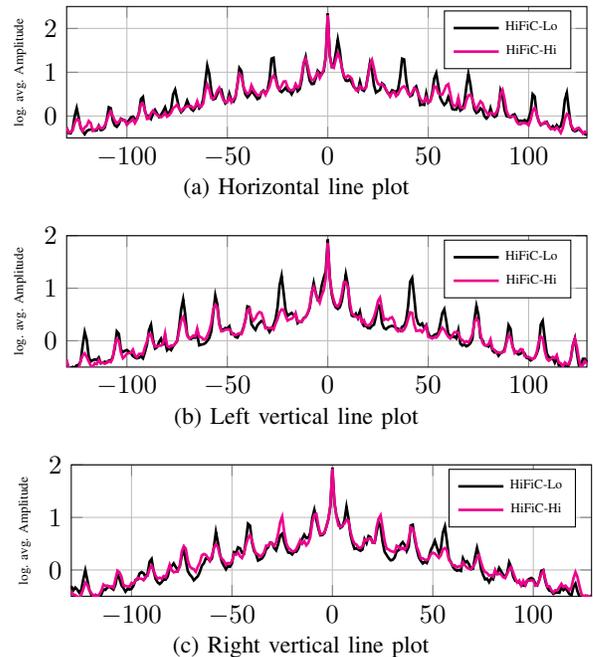


Fig. 4: Homogeneous-region artifacts in horizontal and vertical slices of the spectra in Fig. 2 b). The x-axis shows in a) horizontal frequency u , in b) and c) vertical frequency v .

coefficients are shown in Fig. 4 b) and c). Their position in the spectrum are chosen to intersect the artifact pattern in the pair of nearby peaks. We observe in all plots that the peaks of the artifacts are evenly spaced. HiFiC-Lo creates larger amplitudes, analogously to the artifacts in the previous Section. We also investigated the impact of recompression on these artifacts. The findings are analogous to the artifacts from the previous Section and are not explicitly plotted here.

We note that this pattern is not strictly limited to homogeneous patches. The difference image in Fig. 2 a) also exhibits weak traces of this pattern. However, those traces are mostly overlaid by frequency components from texture.

C. Baseline Detectability for Various Patch Sizes

The previous investigation suggests that distinguishing original and HiFiC-compressed images should in principle be well feasible. We hence perform a baseline detectability evaluation on patches of different size, to empirically test for a lower bound of detection. We use a linear regression classifier on 2000 randomly selected non-overlapping square patches from 45 images of CLIC2020. The patches have 32, 64, 128, and 256 pixels per dimension. Two copies of each image are in uncompressed and in HiFiC-Hi compressed format. 1500 of these patches are used for training and 500 for testing (where patches from the same root image are either in training or in test to avoid information leakage). Each patch is converted to grayscale, and the log-magnitude of its Fourier coefficients is normalized to zero mean and unit variance.

The resulting accuracies are shown in Table I. As expected, smaller patch sizes are more difficult to classify. However,

TABLE I: Logistic Regression on different patch sizes

Patch Size	32×32	64×64	128×128	256×256
Accuracy (%)	72.90	81.48	91.59	93.68

already patches of 128×128 pixels are classified with an accuracy of more than 90%, patches of 256×256 pixels achieve an accuracy of 93.68%. When evaluating the same classifier on patches with HiFiC-Lo compression, then the accuracy increases even to 94.98% due to the stronger artifacts as shown in the previous Sections.

V. ROBUSTNESS OF AI-BASED COMPRESSION ARTIFACTS

We use more complex classifiers than logistic regression to empirically investigate the robustness of the artifacts. Our classifiers of choice are here InceptionV3 [30], ResNet50 [31] and EfficientNetB2 [32].

A. Datasets and Experimental Setup

We use the images by Mentzer *et al.*, which comprise 428 images from the CLIC2020 dataset, 45 images from the DIV2k dataset, and 24 images from the Kodak dataset in uncompressed, HiFiC-Hi, and HiFiC-Lo format. For training and validation, we use 80% and 10% of the CLIC2020 images uncompressed and HiFiC-Hi compressed. The remaining data is used for testing, in particular DIV2k and Kodak images for testing on unseen data. We extract 260×260 non-overlapping patches, i.e., slightly larger than 256×256 , to match the expected input size of EfficientNet. The patches are processed analogously to the previous experiment. Additionally, we train and evaluate on 260×260 patches in spatial domain without any additional processing of the data. Training is done with the Adam optimizer with learning rate $l = 0.0001$ and default parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-7}$. We train the networks for 10 epochs with a batch size of 16.

B. Baseline Neural Network Evaluation

The accuracies for the baseline classification whether an image was AI-compressed are presented in Table II. All networks achieve on all datasets an accuracy greater than 90%, and in particular they outperform the logistic regression by several percent. EfficientNetB2 achieves overall the highest accuracies. Although the performance differences are not significant, we use this network for the remaining experiments.

As a sidenote, we tested the (HiFiC-Hi)-trained EfficientNetB2 also on HiFiC-Lo images. When using frequencies as input, the accuracy even further increases due to the stronger frequency-space artifacts in HiFiC-Lo. Conversely, when using pixels as input, the accuracy decreases due to the unseen image quality. We use this as further empirical evidence for the benefits of detecting HiFiC compression in frequency domain.

Analogously, we verified on the same experimental setup that recompression does not negatively impact the detectability of HiFiC compression. If HiFiC-Hi images are recompressed with HiFiC-Hi or HiFiC-Lo, EfficientNetB2 achieves accuracies of 99.57% and 96.92% in pixel inputs and accuracies

TABLE II: Detection Accuracy (%) of different Architectures

Dataset	InceptionV3		ResNet50		EfficientNetB2	
	Spatial	Freq.	Spatial	Freq.	Spatial	Freq.
CLIC2020	97.42	94.48	96.73	95.20	99.10	96.37
DIV2k	90.57	91.53	91.74	95.89	95.54	95.29
Kodak	94.69	91.12	97.46	92.84	95.77	95.52

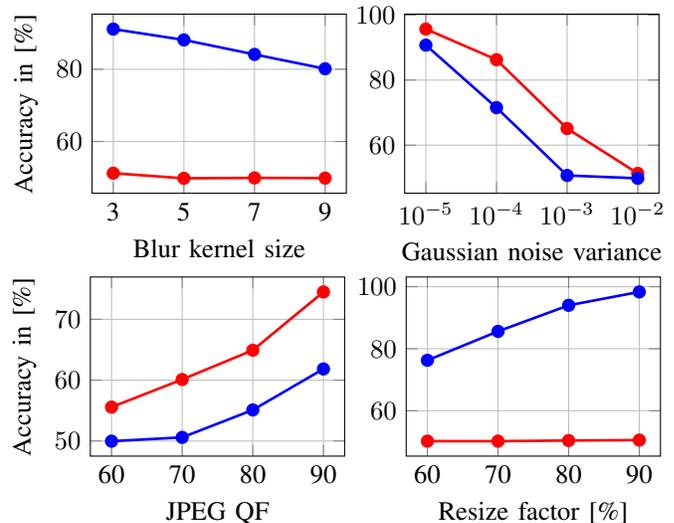


Fig. 5: Accuracy for different postprocessing for frequency inputs (red) and pixel input (blue). Top left: Gaussian blur. Top right: Gaussian noise. Bottom left: JPEG compression. Bottom right: Image scaling.

of 96.94% and 97.93% for frequency inputs. These numbers are even higher than the results in Table II, which confirms that some combinations of recompressions further enhance the compression signature.

C. Robustness to Unseen Postprocessing

We test the robustness to unseen postprocessing on original images and on images with HiFiC-Hi compression, which leaves weaker traces than HiFiC-Lo. We evaluate standard postprocessing from the literature, namely Gaussian blur, Gaussian noise, JPEG compression and image rescaling.

We use kernel sizes 3, 5, 7, and 9 for Gaussian blur. Gaussian noise is added with variances of 10^{-5} , 10^{-4} , 10^{-3} and 10^{-2} . JPEG compression is applied with qualities from 60 to 90 in steps of 10 using libjpeg. For rescaling, we downsample the image to 60%, 70%, 80%, and 90% of its original size. All these postprocessing operations are applied to both original and AI-compressed images, and only at test time, not during training in order to gain insights into the robustness of the features. Figure 5 shows the resulting accuracies for frequency input (red) and pixel input (blue).

Gaussian blur (top left) acts as a low-pass filter that erases most of the frequency artifacts. Surprisingly, spatial domain still provides noticeable traces that the classifier can exploit, even up to a kernel size of 9.

Gaussian noise (top right) also strongly impacts the frequency domain. However, even though both frequency and spatial feature performance deteriorates, the frequency domain features consistently achieve higher accuracies.

JPEG compression (bottom left) also strongly impacts the frequency features. The spatial features, however, are more affected by this operation than the frequency-domain features.

Downsampling (bottom right) distorts and stretches HiFiC's frequency signature, and the accuracy immediately collapses. In comparison, spatial features are surprisingly robust. Since we feed raw RGB to the spatial network, we hypothesize that correlations between color channels might add further cues about AI compression.

In summary, this postprocessing experiment exhibits distinct advantages for both the spatial features and the frequency features. The behavior of the frequency features is in line with the observations in previous Sections. The strong performance of spatial features under unseen Gaussian blur and downsampling indicates that there are further compression cues in spatial domain. These cues are different from the frequency features and their specific characteristic is subject for future work.

VI. CONCLUSIONS

In this work, we examine AI-compressed images for their forensic traces. We argue that such types of compression will gain relevance with the upcoming new standard for JPEG-AI. In our study, two types of frequency-domain artifacts are observed. One feature resembles GAN upsampling artifacts. The other feature is different, and is primarily observed in homogeneous regions. We characterize the frequency signature of both artifacts and empirically test their robustness. Overall, these first results are encouraging. We furthermore observe that there may be further compression traces in spatial domain that can potentially be exploited in future work.

REFERENCES

- [1] J. Lukás and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," *Proc. Digital Forensic Research Workshop*, 2003.
- [2] A. C. Popescu and H. Farid, "Statistical tools for digital forensics," in *International Workshop on Information Hiding*, 2004, pp. 128–147.
- [3] T. Bianchi, A. Piva, and F. Pérez-González, "Near optimal detection of quantized signals and application to JPEG forensics," in *IEEE International Workshop on Information Forensics and Security*, 2013, pp. 168–173.
- [4] J. Ascenso, P. Akyazi, F. Pereira, and T. Ebrahimi, "Learning-based image coding: early solutions reviewing and subjective quality evaluation," in *Optics, Photonics and Digital Technologies for Imaging Applications VI*, vol. 11353. SPIE, 2020, p. 113530S.
- [5] G. Toderici, D. Vincent, N. Johnston, S. Jin Hwang, D. Minnen, J. Shor, and M. Covell, "Full resolution image compression with recurrent neural networks," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [6] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," in *International Conference on Learning Representations*, 2018.
- [7] F. Mentzer, G. D. Toderici, M. Tschannen, and E. Agustsson, "High-fidelity generative image compression," *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [8] A. Berthet and J.-L. Dugelay, "AI-based compression: A new unintended counter attack on JPEG-related image forensic detectors?" in *IEEE International Conference on Image Processing*, 2022, pp. 3426–3430.
- [9] M.-J. Kwon, I.-J. Yu, S.-H. Nam, and H.-K. Lee, "CAT-Net: Compression artifact tracing network for detection and localization of image splicing," in *IEEE Winter Conference on Applications of Computer Vision*, 2021, pp. 375–384.
- [10] D. Bhowmik, M. Elawady, and K. Nogueira, "Security and forensics exploration of learning-based image coding," in *2021 International Conference on Visual Communications and Image Processing*, 2021, pp. 1–5.
- [11] J. Yang, J. Xie, G. Zhu, S. Kwong, and Y.-Q. Shi, "An effective method for detecting double JPEG compression with the same quantization matrix," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 11, pp. 1933–1942, 2014.
- [12] M. Barni, L. Bondi, N. Bonettini, P. Bestagini, A. Costanzo, M. Maggini, B. Tondi, and S. Tubaro, "Aligned and non-aligned double JPEG detection using convolutional neural networks," *Journal of Visual Communication and Image Representation*, vol. 49, pp. 153–163, 2017.
- [13] D. Fu, Y. Q. Shi, and W. Su, "A generalized Benford's law for JPEG coefficients and its applications in image forensics," in *Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, 2007, p. 65051L.
- [14] C. Pasquini, F. Pérez-González, and G. Boato, "A Benford-Fourier JPEG compression detector," in *IEEE International Conference on Image Processing*, 2014, pp. 5322–5326.
- [15] Q. Wang and R. Zhang, "Double JPEG compression forensics based on a convolutional neural network," in *EURASIP J. on Info. Security*, 2016.
- [16] J. Park, D. Cho, W. Ahn, and H.-K. Lee, "Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network," in *European Conference on Computer Vision*, 2018.
- [17] S. Agarwal and H. Farid, "Photo forensics from rounding artifacts," in *ACM Workshop on Information Hiding and Multimedia Security*, 2020, p. 103–114.
- [18] —, "Photo forensics from JPEG dimples," in *IEEE Workshop on Information Forensics and Security*, 2017, pp. 1–6.
- [19] B. Lorch and C. Riess, "Image forensics from chroma subsampling of high-quality JPEG images," in *ACM Workshop on Information Hiding and Multimedia Security*, 2019.
- [20] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 2019.
- [21] O. Giudice, L. Guarnera, and S. Battiato, "Fighting deepfakes by detecting GAN DCT anomalies," *Journal of Imaging*, vol. 7, no. 8, 2021.
- [22] J. Frank, T. Eisenhofer, L. Schönherr, A. Fischer, D. Kolossa, and T. Holz, "Leveraging frequency analysis for deep fake image recognition," in *International Conference on Machine Learning*, vol. 119, 13–18 Jul 2020.
- [23] F. Marra, D. Gragnaniello, L. Verdoliva, and G. Poggi, "Do GANs leave artificial fingerprints?" in *IEEE Conference on Multimedia Information Processing and Retrieval*, 2019, pp. 506–511.
- [24] R. Corvi, D. Cozzolino, G. Zingarini, G. Poggi, K. Nagano, and L. Verdoliva, "On the detection of synthetic images generated by diffusion models," 2022. [Online]. Available: <https://arxiv.org/abs/2211.00680>
- [25] J. Ricker, S. Damm, T. Holz, and A. Fischer, "Towards the detection of diffusion model deepfakes," 2022. [Online]. Available: <https://arxiv.org/abs/2210.14571>
- [26] Z. Duan, M. Lu, Z. Ma, and F. Zhu, "Opening the black box of learned image coders," in *Picture Coding Symposium*, 2022, pp. 73–77.
- [27] G. Toderici, L. Theis, N. Johnston, E. Agustsson, F. Mentzer, J. Balle, W. Shi, and R. Timofte. (2020) CLIC 2020: Challenge on learned image compression. [Online]. Available: <http://compression.cc/>
- [28] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017.
- [29] (1999) Kodak lossless true color image suite. [Online]. Available: <https://r0k.us/graphics/kodak/>
- [30] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [32] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, vol. 97, 2019, pp. 6105–6114.