WHAT IS CLAIMED IS:

1. A method for video encoding, comprising:

generating an initial feature representation from an input image to be encoded;

performing an iterative update of values of a plurality of elements in the initial feature representation, the iterative update comprising

generating a coded representation corresponding to a final feature representation based on the final feature representation, the final feature representation having been updated from the initial feature representation by a number of iterations of the iterative update, and

generating a reconstructed image corresponding to the final feature representation based on the coded representation; and

generating an encoded image corresponding to the final feature representation having updated values of the plurality of elements, wherein

one of (i) a rate-distortion loss corresponding to the final feature representation or (ii) the number of iterations of the iterative update satisfies a pre-determined condition, and

the rate-distortion loss is determined based on a bit consumption of the coded representation corresponding to the final feature representation and a difference between the input image and the reconstructed image corresponding to the final feature representation.

2. The method of claim 1, wherein

the generating the coded representation comprises:

adding noise to the final feature representation to generate a first representation corresponding to the final feature representation;

generating the coded representation based on a simulated entropy encoding of the first representation, and

the generating the reconstructed image comprises a simulated entropy decoding of the coded representation.

3. The method of claim 2, wherein the generating the coded representation based on the simulated entropy encoding further comprises:

determining a context model based on the final feature representation, and

generating the coded representation based on the simulated entropy encoding of the first representation with the determined context model.

4. The method of claim 1, wherein

the generating the encoded image comprises quantizing the final feature representation into a quantized representation, the encoded image being generated based on the quantized representation.

5. The method of claim 4, wherein the generating the encoded image comprises:

entropy encoding the quantized representation into the encoded image.

6. The method of claim 5, wherein the entropy encoding comprises arithmetic encoding the quantized representation into the encoded image.

7. The method of claim 5, wherein the entropy encoding the quantized representation further comprises:

determining a context model based on the final feature representation, and

entropy encoding, based on the determined context model, the quantized representation into the encoded image.

8. The method of claim 1, wherein

the performing comprises performing the iterative update of (i) the values of the plurality of elements in the initial feature representation and (ii) a pretrained parameter in a neural network (NN) used to generate the coded representation based on the final feature representation, the generating the coded representation comprises generating the coded representation corresponding to the final feature representation based on the final feature representation by the NN updated with a replacement parameter that replaces the pretrained parameter, the replacement parameter having been determined from the pretrained parameter by the number of iterations of the iterative update, and

the generating the encoded image comprises generating, based on the NN updated by the replacement parameter, the encoded image from the final feature representation.

9. The method of claim 1, wherein

the performing comprises performing the iterative update of (i) the values of the plurality of elements in the initial feature representation and (ii) a pretrained parameter in a neural network (NN) used to generate the reconstructed image based on the coded representation,

the generating the reconstructed image comprises generating the reconstructed image corresponding to the final feature representation based on the coded representation corresponding to the final feature representation by the NN updated with a replacement parameter that replaces the pretrained parameter, the replacement parameter having been determined from the pretrained parameter by the number of iterations of the iterative update, and

the method further comprises

generating neural network update information that is comprised in a coded bitstream, the neural network update information indicating the replacement parameter, and

transmitting the encoded image and the neural network update information in the coded bitstream.

10. The method of claim 1, further comprising:

transmitting the encoded image in a coded bitstream.

11. An apparatus for video encoding, comprising:

processing circuitry configured to:

generate an initial feature representation from an input image to be encoded;

perform an iterative update of values of a plurality of elements in the initial feature representation, the iterative update comprising

generating a coded representation corresponding to a final feature representation based on the final feature representation, the final feature representation having been updated from the initial feature representation by a number of iterations of the iterative update, and

generating a reconstructed image corresponding to the final feature representation based on the coded representation; and

generate an encoded image corresponding to the final feature representation having updated values of the plurality of elements, wherein

one of (i) a rate-distortion loss corresponding to the final feature representation or (ii) the number of iterations of the iterative update satisfies a pre-determined condition, and

the rate-distortion loss is determined based on a bit consumption of the coded representation corresponding to the final feature representation and a difference between the input image and the reconstructed image corresponding to the final feature representation.

12. The apparatus of claim 11, wherein the processing circuitry is configured to:

add noise to the final feature representation to generate a first representation corresponding to the final feature representation;

generate the coded representation based on a simulated entropy encoding of the first representation, and

perform a simulated entropy decoding of the coded representation.

13. The apparatus of claim 12, wherein the processing circuitry is configured to:

determine a context model based on the final feature representation, and

generate the coded representation based on the simulated entropy encoding of the first representation with the determined context model.

14. The apparatus of claim 11, wherein the processing circuitry is configured to:

quantize the final feature representation into a quantized representation, the encoded image being generated based on the quantized representation.

15. The apparatus of claim 14, wherein the processing circuitry is configured to:

entropy encode the quantized representation into the encoded image.

16. The apparatus of claim 15, wherein the processing circuitry is configured to arithmetic encode the quantized representation into the encoded image.

17. The apparatus of claim 15, wherein the processing circuitry is configured to:

determine a context model based on the final feature representation, and

entropy encode, based on the determined context model, the quantized representation into the encoded image.

18. The apparatus of claim 11, wherein the processing circuitry is configured to:

perform the iterative update of (i) the values of the plurality of elements in the initial feature representation and (ii) a pretrained parameter in a neural network (NN) used to generate the coded representation based on the final feature representation comprising

generating the coded representation corresponding to the final feature representation based on the final feature representation by the NN updated with a replacement parameter that replaces the pretrained parameter, the replacement parameter having been determined from the pretrained parameter by the number of iterations of the iterative update, and

generate, based on the NN updated by the replacement parameter, the encoded image from the final feature representation.

19. The apparatus of claim 11, wherein the processing circuitry is configured to:

perform the iterative update of (i) the values of the plurality of elements in the initial feature representation and (ii) a pretrained parameter in a neural network (NN) used to generate the reconstructed image based on the coded representation comprising

generating the reconstructed image corresponding to the final feature representation based on the coded representation corresponding to the final feature representation by the NN updated with a replacement parameter that replaces the pretrained parameter, the replacement parameter having been determined from the pretrained parameter by the number of iterations of the iterative update,

generate neural network update information that is comprised in a coded bitstream, the neural network update information indicating the replacement parameter, and

transmit the encoded image and the neural network update information in the coded bitstream.

20. A non-transitory computer-readable storage medium storing a program executable by at least one processor to perform:

generating an initial feature representation from an input image to be encoded;

performing an iterative update of values of a plurality of elements in the initial feature representation, the iterative update comprising

generating a coded representation corresponding to a final feature representation based on the final feature representation, the final feature representation having been updated from the initial feature representation by a number of iterations of the iterative update, and

generating a reconstructed image corresponding to the final feature representation based on the coded representation; and

generating an encoded image corresponding to the final feature representation having updated values of the plurality of elements, wherein

one of (i) a rate-distortion loss corresponding to the final feature representation or (ii) the number of iterations of the iterative update satisfies a pre-determined condition, and

the rate-distortion loss is determined based on a bit consumption of the coded representation corresponding to the final feature representation and a difference between the input image and the reconstructed image corresponding to the final feature representation.