Claims

1. A method for encoding a picture using a neural network, NN, wherein the NN comprises a plurality of layers comprising at least one downsampling layer (2110, 2120, 2130, 2140) that applies a downsampling to an input (2101, 2102, 2103, 2104) representing a matrix having a size S in at least one dimension, the method comprising, applying (2321), before processing (2311) the input with the at least one downsampling layer, a rescaling to the input, wherein the rescaling comprises changing the size S in the at least one dimension so that $\bar{S} = n \cdot r$, with $n \in \mathbb{N}$ and $r > 1, r \in \mathbb{N}$, where r is the downsampling ratio of the layer; after the rescaling, processing (2311) the input by the downsampling layer (2110, 2120, 2130, 2140) and providing (2312) an output with the size $S_1 < S$; and providing, after processing the picture using the NN, a bitstream as output of the NN.

2. The method according to claim 1, wherein the NN comprises a number of $M \in \mathbb{N}$ downsampling layers, wherein each downsampling layer m (2110, 2120, 2130, 2140), $m \leq M, m \in \mathbb{N}$ has an associated downsampling ratio $r_m > 1, r_m \in \mathbb{N}$ and the method further comprises: before processing (2311) an input (2101, 2102, 2103, 2104) representing a matrix having a size S_m in at least one dimension with a downsampling layer m, applying, if the size S_m is not an integer multiple of the downsampling ratio r_m of the downsampling layer m, a rescaling to the input, wherein the rescaling comprises changing the size S_m in the at least one dimension so that $\overline{S_m} = n \cdot r_m, n \in \mathbb{N}$; and/or

   wherein, before applying (2321) the rescaling to the input (2101, 2102, 2103, 2104) with the size S_m a determination is made whether S_m is an integer multiple of the downsampling ratio r_m and, if it is determined that S_m is not an integer multiple of the downsampling ratio r_m, the rescaling is applied to the input so that the size S_m is changed in the at least one dimension so that $\overline{S_m} = n \cdot r_m, n \in \mathbb{N}$.

~~3. The method according to claim 2, wherein, before applying (2321) the rescaling to the input (2101, 2102, 2103, 2104) with the size S_m a determination is made whether S_m is an integer multiple of the downsampling ratio r_m and, if it is determined that S_m is not an integer multiple of the downsampling ratio r_m, the rescaling is applied to the input so that the size S_m is changed in the at least one dimension so that $\overline{S_m} = n \cdot r_m, n \in \mathbb{N}$.~~

~~4. The method according to claim 2 or 3, wherein, if the size S_m of the input (2101, 2102, 2103, 2104) is an integer multiple of the downsampling ratio r_m of the downsampling~~

~~layer m (2110, 2120, 2130, 2140), no rescaling to a size $\overline{S_m} \neq S_m$ is applied to the input before processing the input by the downsampling layer m.~~

~~5. The method according to 3 or 4, wherein the determination whether S_m is an integer multiple of the downsampling ratio r_m comprises comparing the size S_m to an allowed input size of the downsampling layer m (2110, 2120, 2130, 2140).~~

~~6. The method according to 5, wherein the allowed input size of the downsampling layer m is calculated based on at least one of the downsampling ratio r_m and the size S_m.~~

~~7. The method according to claim 5 or 6, wherein the comparing comprises calculating a difference between S_m and the allowed input size of the downsampling layer m (2110, 2120, 2130, 2140).~~

~~8. The method according to any of claims 5 or 7, wherein the comparing comprises determining $ceil\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m$ or $floor\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m$.~~

~~9. The method according to claim 8, wherein $ceil\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m$ is determined and, if $ceil\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m \neq 0$, the rescaling is applied to the input with the size S_m.~~

~~10. The method according to claim 8, wherein $floor\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m$ is determined and, if $floor\left(\frac{S_m}{r_m}\right) \cdot r_m - S_m \neq 0$, the rescaling is applied to the input with the size S_m.~~

~~11. The method according to claim 10, wherein the size $\overline{S_m}$ is determined using at least one of the downsampling ratio r_m and the size S_m.~~

~~12.~~3. The method according to claim ~~11~~1 or 2, wherein the size $\overline{S_m}$ is determined using a function comprising at least one of ceil, int, floor.

~~13.~~4. The method according to claim ~~12~~3, wherein:

- the size $\overline{S_m}$ is determined using $floor\left(\frac{S_m}{r_m}\right) \cdot r_m = \overline{S_m}$; or
- the size $\overline{S_m}$ is determined using $ceil\left(\frac{S_m}{r_m}\right) \cdot r_m = \overline{S_m}$; or
- the size $\overline{S_m}$ is determined using $int\left(\frac{S_m}{r_m}\right) \cdot r_m = \overline{S_m}$.

~~14. The method according to any of claims 2 to 13, wherein the rescaling applied to an input (2101, 2102, 2103, 2104) of a downsampling layer m is independent from the downsampling ratios $r_k, k \neq m$ of other downsampling layers.~~

~~15.~~5. The method according to any of claims 1 to ~~14~~4, wherein the input (2101, 2102, 2103, 2104) to a downsampling layer m has a size S_m in the at least one dimension that has a value that is between a closest smaller integer multiple of the downsampling ratio r_m of the downsampling layer m (2110, 2120, 2130, 2140) and a closest larger integer multiple of the downsampling ratio r_m of the downsampling layer m (2110, 2120, 2130, 2140) and wherein, depending on a condition, the size S_m of the input is changed during the rescaling to either match the closest smaller integer multiple of the downsampling ratio r_m or to match the closest larger integer multiple of the downsampling ratio r_m.

~~16.~~6. The method according to any of claims 1 to ~~15~~5, wherein the input (2101, 2102, 2103, 2104) to a downsampling layer m has a size S_m in the at least one dimension that has a value that is not an integer multiple of the downsampling ratio r_m of the downsampling layer m (2110, 2120, 2130, 2140), wherein the size S_m of the input is changed during the rescaling to either match the closest smaller integer multiple of the downsampling ratio r_m or to match the closest larger integer multiple of the downsampling ratio r_m.

~~17. The method according to any of claims 1 to 16, wherein the input (2101, 2102, 2103, 2104) to a downsampling layer m has a size S_m in the at least one dimension, wherein $lr_m \leq S_m \leq r_m(l+1), l, m \in \mathbb{N}$ and r_m is the downsampling ratio of the downsampling layer m (2110, 2120, 2130, 2140) and wherein the size S_m is either rescaled to $\overline{S_m} = lr_m$ or to $\overline{S_m} = r_m(l+1)$ depending on a condition.~~

~~18. The method according to claim 1 to 17, wherein, if the size S_m of the input (2101, 2102, 2103, 2104) is closer to the closest smaller integer multiple of the downsampling ratio r_m of the downsampling layer m (2110, 2120, 2130, 2140) than to the closest larger integer multiple of the downsampling ratio r_m, the size S_m of the input is reduced to a size $\overline{S_m}$ that matches the closest smaller integer multiple of the downsampling ratio r_m.~~

~~19. The method according to claim 18, wherein reducing the size S_m of the input (2101, 2102, 2103, 2104) to the size $\overline{S_m}$ comprises cropping the input.~~

~~20.~~7. The method according to any of claims 1 to ~~19~~6, wherein, if the size S_m of the input (2101, 2102, 2103, 2104) is closer to the closest larger integer multiple of the downsampling ratio r_m of the downsampling layer m (2110, 2120, 2130, 2140) than to the closest smaller integer multiple of the downsampling ratio r_m, the size S_m of the input is increased to a size $\overline{S_m}$ that matches the closest larger integer multiple of the downsampling ratio r_m, wherein increasing the size S_m of the input (2101, 2102, 2103, 2104) to the size $\overline{S_m}$ comprises padding the input with the size S_m with zeros or with padding information obtained from the input with the size S_m..

~~21. The method according to claim 20, wherein increasing the size S_m of the input (2101, 2102, 2103, 2104) to the size $\overline{S_m}$ comprises padding the input with the size S_m with zeros or with padding information obtained from the input with the size S_m.~~

~~22.~~8. The method according to claim ~~21~~7, wherein the padding information obtained from the input with the size S_m is applied as redundant padding information to increase the size S_m of the input (2101, 2102, 2103, 2104) to the size $\overline{S_m}$ and wherein the padding with redundant padding information comprises at least one of reflection padding and repetition padding.

~~23. The method according to claim 22, wherein the padding with redundant padding information comprises at least one of reflection padding and repetition padding.~~

~~24. The method according to claim 22 or 23, wherein the padding information is or comprises at least one value of the input with the size S_m that is closest to a region in the input (2101, 2102, 2103, 2104) to which the redundant padding information is to be added.~~

~~25. The method according to claim 1 or 13, wherein the size S_m of the input (2101, 2102, 2103, 2104) to the downsampling layer m (2110, 2120, 2130, 2140) is increased to a size $\overline{S_m}$ that matches the closest larger integer multiple of the downsampling ratio r_m.~~

~~26. The method according to claim 17, wherein the condition makes use of $Min(|S - lr_m|, |S - r_m(l+1)|)$ and wherein the condition comprises, if Min throws $|S_m - lr_m|$, then the size S_m of the input is reduced to $\overline{S_m} = lr_m$ and, if Min throws $|S_m - (l+1)r_m|$, then the size S_m of the input is increased to $\overline{S_m} = (l+1)r_m$.~~

~~27. The method according to claim 26, wherein l is determined using at least one of the size S_m of the input to the downsampling layer m (2110, 2120, 2130, 2140) and the downsampling ratio r_m of the downsampling layer m.~~

~~28. The method according to claim 27, wherein l is determined by $l = floor\left(\frac{S_m}{r_m}\right)$ and/or wherein l+1 is determined by $l + 1 = ceil\left(\frac{S_m}{r_m}\right)$.~~

~~29. The method according to any of claims 1 to 28, wherein at least one of the downsampling layers (2110, 2120, 2130, 2140) applies a downsampling to the input (2101, 2102, 2103, 2104) in the two dimensions and the downsampling ratio in the first dimension is equal to the downsampling ratio in the second dimension.~~

~~30. The method according to any of claims 1 to 29, wherein the downsampling ratios of all downsampling layers (2110, 2120, 2130, 2140) are equal.~~

~~31.~~9. A method for decoding a bitstream representing a picture using a neural network, NN, wherein the NN comprises a plurality of layers comprising at least a first upsampling layer and a second upsampling layer that apply an upsampling to an input (2201, 2202, 2203, 2204) representing a matrix having a size T in at least one dimension, the method comprising processing the input by the first upsampling layer and providing an output of the first upsampling layer, wherein the output has a size $\overline{T_1} = T \cdot u_1, u_1 > 1, u_1 \in \mathbb{N}$, wherein u_1 is the upsampling ratio of the first upsampling layer; applying, before processing the output of the first upsampling layer by the second upsampling layer, a rescaling (2421) to the output of the first upsampling layer, wherein the rescaling comprises changing the size $\overline{T_1}$ of the output in the at least one dimension to a size $\widehat{T_1}$ in the at least one dimension based on information obtained; processing the rescaled output by the second upsampling layer and providing an output of the second upsampling layer, wherein the output has a size $\overline{T_2} = \widehat{T_1} \cdot u_2, u_2 > 1, u_2 \in \mathbb{N}$ in the at least one dimension, wherein u_2 is the upsampling ratio of the second upsampling layer; and providing, after processing the bitstream using the NN, a decoded picture as output of the NN.

~~32.~~10. The method according to claim ~~31~~9, wherein the upsampling layer comprises a transposed convolution or convolution layer.

~~33.~~11. The method according to claim ~~31~~ 9 or ~~32~~10, wherein the information comprises at least one of a target size of the decoded picture comprising at least one of a height H of the decoded picture and a width W of the decoded picture, the upsampling ratio u_1, the upsampling ratio u_2, a target output size $\overline{T_2}$ of the second upsampling layer, the size $\widehat{T_1}$.

34.12. The method according to any of claims 31 9 to 3311, wherein the information is obtained from at least one of: the bitstream, a second bitstream, information available at the decoder.

35. The method according to any of claims 31 to 34, wherein the rescaling (2421) for changing the size $\overline{T_1}$ in the at least one dimension to a size $\widehat{T_1}$ is determined based on a formula depending on $\overline{T_2}$ and u_2, wherein $\overline{T_2}$ is a target output size of the output of the second upsampling layer and u_2 is the upsampling ratio of the second upsampling layer.

36. The method according to any of claims 31 to 35, wherein the rescaling for changing the size $\overline{T_1}$ in the at least one dimension to a size $\widehat{T_1}$ is determined based on a formula depending on u_2 and N, where N is the total number of upsampling layers proceeding the first upsampling layer in processing order of the bitstream through the NN.

37. The method according to claim 31 or 36, wherein the formula is given by $\widehat{T_1} = ceil\left(\frac{T_{output}}{u_2^N}\right)$, wherein T_{output} is the target size of the output of the NN.

38.13. The method according to any one of claims 31 or 379 to 12, wherein the formula is given by $\widehat{T_1} = ceil\left(\frac{T_{output}}{u^N}\right)$, wherein T_{output} is the target size of the output of the NN and u is an upsampling ratio.

39. The method according to claim 38, wherein an indication for indicating the size T_{output} is included in the bitstream.

40. The method according to claim 35, wherein the formula is given by $\widehat{T_1} = ceil\left(\frac{\overline{T_2}}{u_2}\right)$ or wherein the formula is given by $\widehat{T_1} = floor\left(\frac{\overline{T_2}}{u_2}\right)$.

41. The method according to claim 40, wherein an indication is obtained from the bitstream that indicates which of the multiple predefined formulas are selected.

42. The method according to any of claims 31 to 41, wherein the method further comprises, before the rescaling (2412) the output with the size $\overline{T_1}$, determining whether the size $\overline{T_1}$ of the output matches the size $\widehat{T_1}$.

43. The method according to claim 42, wherein, if it is determined that the size $\overline{T_1}$ matches the size $\widehat{T_1}$, no rescaling changing the size $\overline{T_1}$ is applied.

~~44.~~14. The method according to any one of claims ~~42 or 43~~9 to 13, wherein the method further comprises determining whether $\overline{T_1}$ is larger than $\widehat{T_1}$ or whether $\overline{T_1}$ is smaller than $\widehat{T_1}$, and wherein if it is determined that $\overline{T_1}$ is larger than $\widehat{T_1}$, the rescaling comprises applying a cropping to the output with the size $\overline{T_1}$ such that the size $\overline{T_1}$ is reduced to the size $\widehat{T_1}$; and

if it is determined that $\overline{T_1}$ is smaller than $\widehat{T_1}$, the rescaling comprises applying a padding to the output with the size $\overline{T_1}$ such that the size $\overline{T_1}$ is increased to the size $\widehat{T_1}$, wherein the padding comprises padding the output with the size $\overline{T_1}$ with zeros or with padding information obtained from the output with the size $\overline{T_1}$, wherein the padding information obtained from the output with the size $\overline{T_1}$ is applied as redundant padding information to increase the size $\overline{T_1}$ of the output to the size $\widehat{T_1}$ and wherein the padding comprises reflection padding or repetition padding.

~~45. The method according to claim 44, wherein, if it is determined that $\overline{T_1}$ is larger than $\widehat{T_1}$, the rescaling comprises applying a cropping to the output with the size $\overline{T_1}$ such that the size $\overline{T_1}$ is reduced to the size $\widehat{T_1}$.~~

~~46. The method according to claim 44 or 45, wherein, if it is determined that $\overline{T_1}$ is smaller than $\widehat{T_1}$, the rescaling comprises applying a padding to the output with the size $\overline{T_1}$ such that the size $\overline{T_1}$ is increased to the size $\widehat{T_1}$.~~

~~47. The method according to claim 46, wherein the padding comprises padding the output with the size $\overline{T_1}$ with zeros or with padding information obtained from the output with the size $\overline{T_1}$.~~

~~48. The method according to claim 47, wherein the padding information obtained from the output with the size $\overline{T_1}$ is applied as redundant padding information to increase the size $\overline{T_1}$ of the output to the size $\widehat{T_1}$.~~

~~49. The method according to claim 48, wherein the padding comprises reflection padding or repetition padding.~~

~~50. The method according to claim 48 or 49, wherein the padding information is or comprises at least one value of the output with the size $\overline{T_1}$ that is closest to a region in the output to which the redundant padding information is to be added.~~

~~51. The method according to claim 44, wherein, if it is determined that $\overline{T_1}$ is not equal to $\widehat{T_1}$, the rescaling comprises applying an interpolation filter.~~

~~52. The method according to any of claims 31 to 51, wherein the information is provided in the bitstream or a further bitstream and comprises a downsampling ratio r_m of at least one downsampling layer of an encoder that encoded the bitstream, wherein the downsampling layer corresponds, in the order of processing the input, to the upsampling layer of the decoder.~~

~~53. The method according to any of claims 31 to 52, wherein at least one of the upsampling layers applies an upsampling in the two dimensions and the upsampling ratio in the first dimension is equal to the upsampling ratio in the second dimension.~~

~~54. The method according to any of claims 31 to 53, wherein the upsampling ratios of all upsampling layers are equal.~~

~~55.~~15. The method according to any of claims 1 to ~~54~~14, wherein the NN comprises, in the processing order of the bitstream through the NN, a further unit that applies a transformation to the input that does not change the size of the input in the at least one dimension, wherein the method comprises applying the rescaling after the processing of the input by the further unit and before processing the input by the following layer, if the rescaling results in an increase of the size of the input in the at least one dimension, and/or wherein the method comprises applying the rescaling before the processing of the input by the further unit, if the rescaling comprises a decrease of the size of the input in the at least one dimension, wherein the further unit is or comprises a batch normalizer and/or a rectified linear unit, ReLU.

~~56. The method according to claim 55, wherein the further unit is or comprises a batch normalizer and/or a rectified linear unit, ReLU.~~

~~57. The method according to any one of claims 1 to 56, wherein the bitstream comprises sub-bitstreams corresponding to distinct color channels of the picture and wherein the NN comprises sub-neural networks, sNN, that are each adapted to apply a method according to any of claims 1 to 51 to the sub-bitstream provided as input to the sNN.~~

~~58. The method according to any of claims 1 to 30, wherein, if the rescaling comprises increasing the size S to the size $\bar{S}$, the size $\bar{S}$ is given by $\bar{S} = Int\left(\frac{S+r-1}{r}\right)r$ and, if the rescaling comprises reducing the size S to the size $\bar{S}$, the size $\bar{S}$ is given by $\bar{S} = Int\left(\frac{S}{r}\right)r$.~~

~~59. The method according to any of claims 31 to 54, wherein, if the rescaling comprises increasing the size $\overline{T_i}$ to the size $\widehat{T_i}$, the size $\widehat{T_i}$ is given by $\widehat{T_i} = Int\left(\frac{\overline{T_i}+u_z-1}{u_z}\right)u_z$ and, if the rescaling comprises reducing the size $\overline{T_i}$ to the size $\widehat{T_i}$, the size $\widehat{T_i}$ is given by $\widehat{T_i} = Int\left(\frac{\overline{T_i}}{u_z}\right)u_z$.~~

~~60. An encoder (2100) for encoding a picture, wherein the encoder comprises a receiver for receiving a picture and one or more processors configured to implement a neural network, NN, the NN comprising, in a processing order of a picture through the NN, a plurality of layers (2110, 2120, 2130, 2140) comprising at least one downsampling layer that is adapted to apply a downsampling to an input (2101, 2102, 2103, 2104), and transmitter for outputting a bitstream, wherein the encoder is adapted to perform a method according to any of claims 1 to 30 or 55 to 59.~~

~~61.~~16. An encoder (2100) for encoding a picture, wherein the encoder comprises one or more processors for implementing a neural network, NN, wherein the one or more processors are adapted to perform a method according to any of claims 1 to ~~30~~8 or ~~55~~15 in combination with any of claims 1 to 8 ~~to 59~~.

~~62. An encoder (2100) for encoding a picture using a neural network, NN, wherein the encoder comprises one or more processors for implementing the NN, wherein the NN comprises a plurality of layers comprising at least one downsampling layer (2110, 2120, 2130, 2140) that applies a downsampling to an input (2101, 2102, 2103, 2104) representing a matrix having a size S in at least one dimension, and wherein the one or more processors are adapted to:~~

~~apply (2321), before processing (2311) the input with the at least one downsampling layer, a rescaling to the input, wherein the rescaling comprises changing the size S in the at least one dimension so that $\bar{S} = n \cdot r$, with $n \in \mathbb{N}$ and $r > 1, r \in \mathbb{N}$, where r is the downsampling ratio of the layer;~~

~~after the rescaling, process (2311) the input by the downsampling layer (2110, 2120, 2130, 2140) and provide (2312) an output with the size $S_i < S$; and~~

~~provide, after processing the picture using the NN, a bitstream as output of the NN.~~

~~63. A decoder (2200) for decoding a bitstream representing a picture, wherein the decoder comprises a receiver for receiving a bitstream and one or more processors configured to implement a neural network, NN, the NN comprising, in a processing order of a~~

~~bitstream through the NN, a plurality of layers (2210, 2220, 2230, 2240) comprising at least two upsampling layers that are adapted to apply upsampling to an input (2201, 2202, 2203, 2204), and a transmitter for outputting a decoded picture, wherein the decoder is adapted to perform any of the methods of claim 31 to 57 or 59.~~

~~64.~~17. A decoder (2200) for decoding a bitstream representing a picture, wherein the decoder comprises one or more processors for implementing a neural network, NN, wherein the one or more processors are adapted to perform a method according to any of claims ~~31~~9 to ~~57~~14 or ~~59~~15 in combination with any of claims 9 to 14.

~~65. A decoder (2200) for decoding a bitstream representing a picture, wherein the decoder comprises one or more processors for implementing a neural network, NN, wherein the NN comprises a plurality of layers comprising at least a first upsampling layer and a second upsampling layer that apply an upsampling to an input (2201, 2202, 2203, 2204) representing a matrix having a size T in at least one dimension, wherein the one or more processors are adapted to:~~

~~process the input by the first upsampling layer and provide an output of the first upsampling layer, wherein the output has a size $\overline{T_1} = T \cdot u_1, u_1 > 1, u_1 \in \mathbb{N}$, wherein u_1 is the upsampling ratio of the first upsampling layer;~~

~~apply, before processing the output of the first upsampling layer by the second upsampling layer, rescaling (2421) to the output of the first upsampling layer, wherein the rescaling comprises changing the size $\overline{T_1}$ of the output in the at least one dimension to a size $\widehat{T_1}$ in the at least one dimension based on information obtained;~~

~~process the rescaled output by the second upsampling layer and provide an output of the second upsampling layer, wherein the output has a size $\overline{T_2} = \widehat{T_1} \cdot u_2, u_2 > 1, u_2 \in \mathbb{N}$ in the at least one dimension, wherein u_2 is the upsampling ratio of the second upsampling layer; and~~

~~provide, after processing the bitstream using the NN, a decoded picture as output of the NN.~~

~~66. A computer-readable storage medium comprising computer executable instructions that, when executed on a computing system, cause the computing system to execute a method according to any of claims 1 to 59.~~