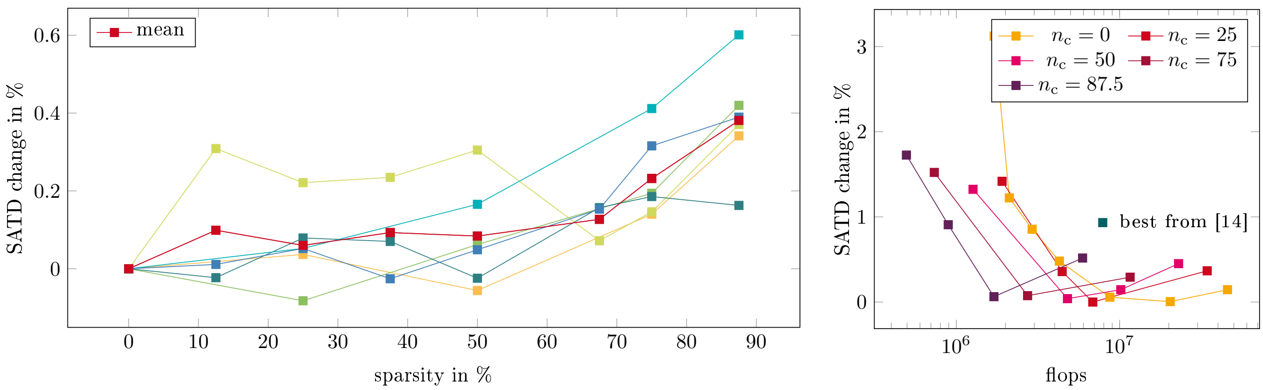
Additions to the Architecture Pretests

1 Further Result Visualisations

There are two types of visualisations used to visualize the effect of different hyperparameter changes on the performance and complexity of the prediction networks.
The first ones are lower bounds as used in the paper as well. These show the piecewise linear lower bound on of the results with a certain parameter. In some cases, if using pruning or deconvolutions was only tried with one of the parameter values, these results were excluded as they are largely biasing the comparison.
For a second type of visualisation, subsets that have exactly the same settings except for the parameter under consideration were found and the performance change compared to the result with a chosen reference parameter value was plotted. Each subset is indicated by a different colour.
In both cases conclusions can only be drawn from the visualised results under carefull consideration of the different numbers of tests for the different parameter values.

1.1 Sparsity

Figure 1b and 1a analyse the results with regards to the target sparsity. It can be clearly seen in figure 1a that low sparsity values barely decrease the performance. In some cases, the performance even increases slightly compared to the version not using pruning. Only above 62.5% the average performance decrease over the different subsets starts to become more noticeably. However, there is still one subset, where even at 87.5% the performance decrease is barely above the variation among the lower sparsity values. This subset is the one using deconvolutional layers.
Figure 1b shows how significant the resulting complexity reduction is compared to the performance decrease.



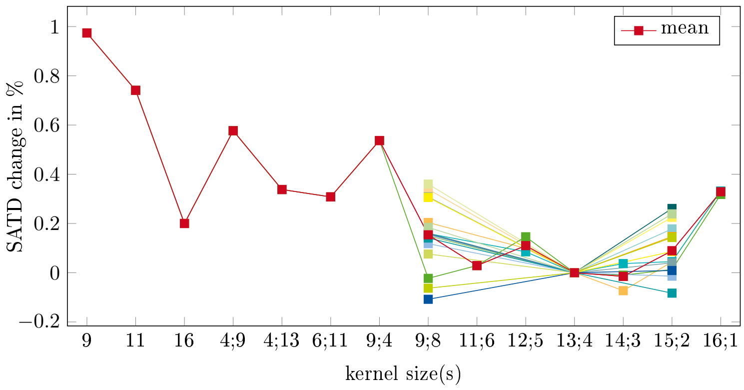
(a) Subsets of results with otherwise same settings over different sparsity values.

(b) Lower bounds on the results with different sparsity values.

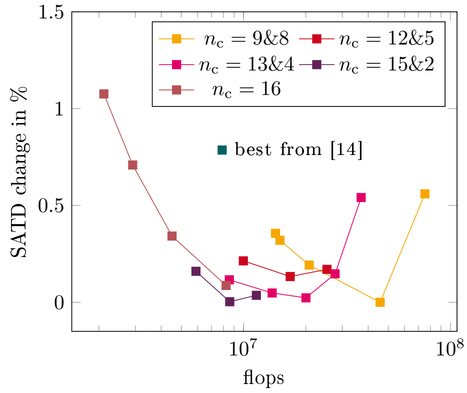
Figure 1: Visualisations of the results of different target sparsity values.

1.2 Convolution design

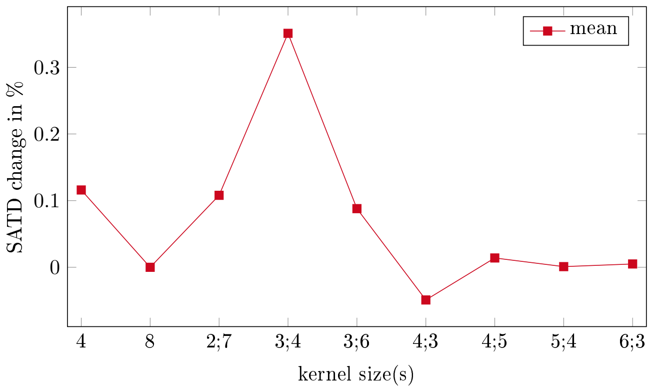
The use of different kernel sizes is evaluated in figure 2. One obvious result visible here is, that kernel combinations that do not add up to the number of refference lines +1 are performing poorly, althoug they have a higher complexity than similar combinations that do add up. It can further be seen, that a second convolutional layer increases the performance the performance slightly especially when using comparably large reference areas. Choosing the first kernel large than the second one decreases the complexity. As long as the first kernel is smaller than 15, this does not decrease the performance, but even slightly increases it for most cases. Only when using only 4 reference lines a single convolutional layer performs best.



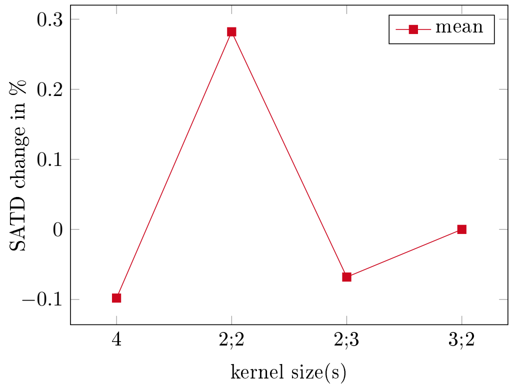
(a) Different subsets of results that use 16 reference lines and vary in no setting but the kernel size.



(b) Lower bounds on the results with 16 reference lines, no pruning, no deconvolution and different kernel combinations.



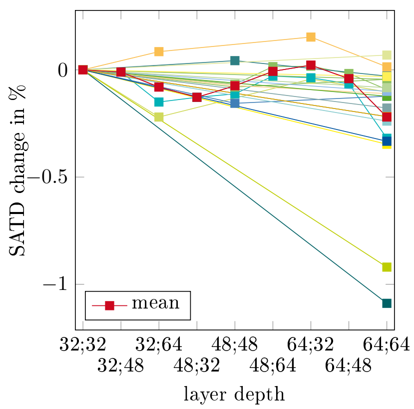
(c) The subset of results that uses 8 reference lines and varies in no setting but the kernel size



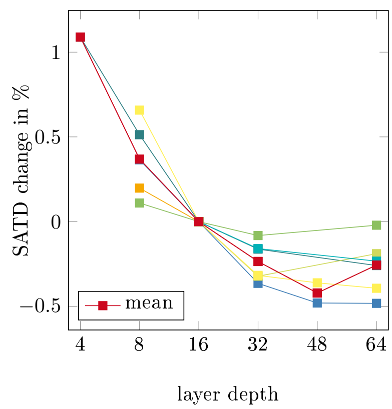
(d) The subset of results that uses 4 reference lines and varies in no setting but the kernel size

Figure 2: Visualisations of the prediction accuracy change, when varying the kernel size(s) of the convolutional layers.

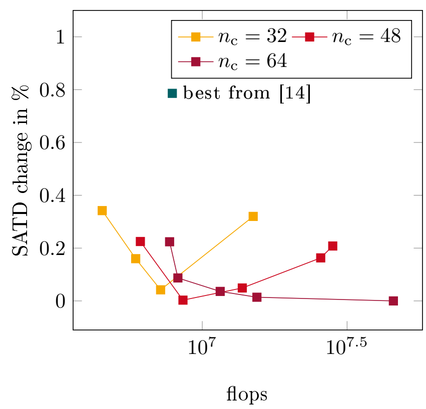
It is quite clear from figure 3, that espesially when using only a single reference line, the performenc increases with the layer depth up to the highest tested value of 64. However, the additional gain steadily decreases. For two convolutional layers the effect is less clear, but still there is a gain when increasing the size from 32; 32 to 64; 64. The achieved complexity reduction is noticeable but small compared to the change caused by other parameters.



(a) Different subsets of results that use two conv. layers and vary in no setting but the depth of these layers.



(b) Different subsets of results that use one conv. layer and vary in no setting but the depth of these layers.

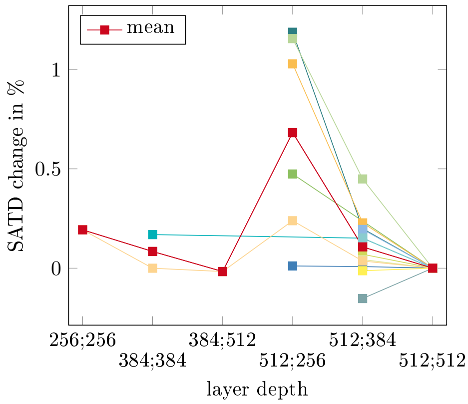


(c) Lower bounds on the results without pruning and deconvolution for different average layer depth values.

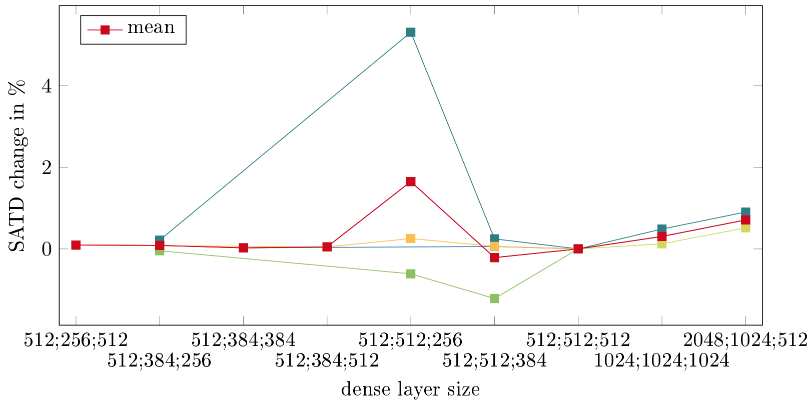
Figure 3: Visualisations of the prediction accuracy change, when varying the depth of the convolutional layers.

1.3 Dense Layer Design

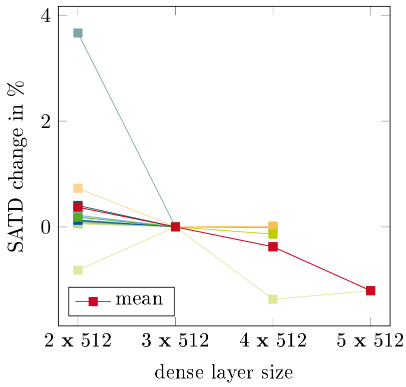
The number and size of the used dense layers was evaluated in detail in figure 4. As can be seen in figure 4a and 4b the most noticeable effect when changing the layer sizes is that there are some outliers that performe considerably worse than the reference. Especially, when using at least one layer of size 256, which is equal to the number of pixels to predict, the results are very varying. This ranges from performing slightly better than the reference up to more than 5% worse. Thus, it is likely, that a layer size close to the number of pixels is possible, but much more difficult to learn than slightly larger sizes. It is also noteworthy, that an increase of the layer size above 512 seems to neither improve the prediction quality nor stabilise the learning process further.



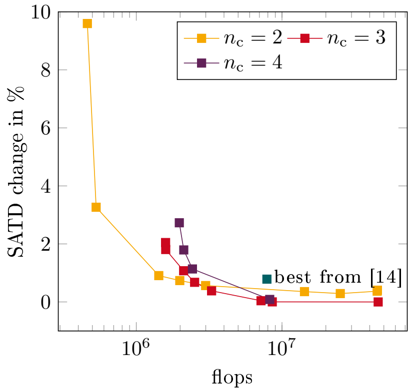
(a) Different subsets of results that use two dense layers and vary in no setting but the size of these layers.



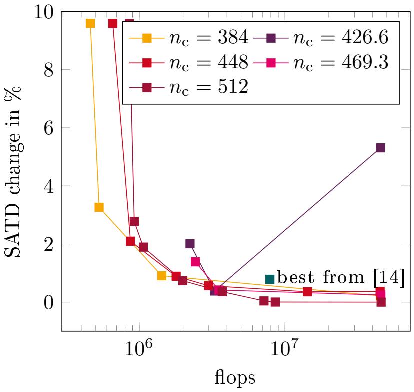
(b) Different subsets of results that use three dense layers and vary in no setting but the size of these layers.



(c) Different subsets of results that use a size of 512 for all dense layers and vary in no setting but the number of used dense layers.



(d) Lower bounds on the results without pruning and deconvolution for different numbers of dense layers.



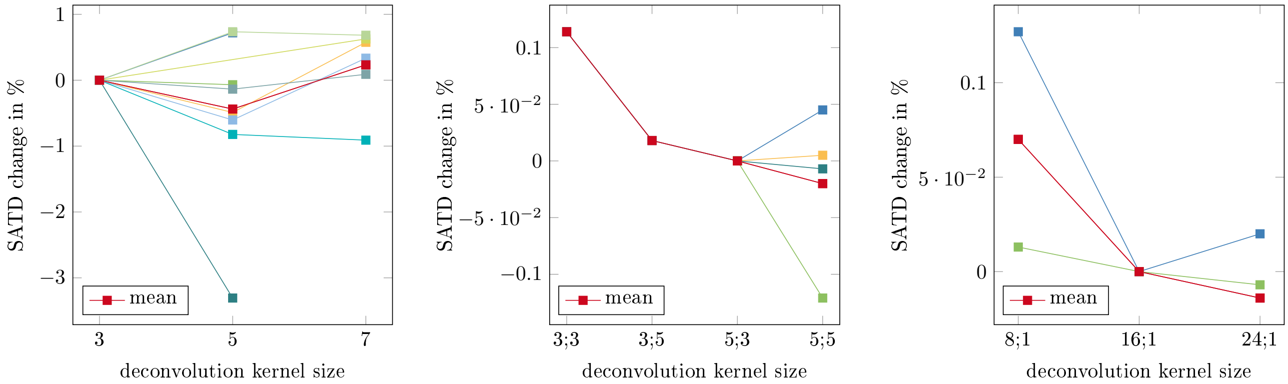
(e) Lower bounds on the results without pruning and deconvolution for different average dense layer sizes.

Figure 4: Visualisations of the prediction accuracy change, when varying the number and size of the dense layers.

A clearer prediction improvement can be observed when increasing the number of used layers as shown in figure 4c. Any additionally used layer steadily increased average performance with very few outliers, but also adds to the complexity as can be seen in figure 4d. Compared to adding an additional Layer the complexity change from using a different layer size is rather small. The lower bounds additionally show, that noticeable variation that could be seen in the different subset results is not correlated significantly correlated to the network complexity, as in that case the lower bounds would diverge.

1.4 Deconvolution

A final regarded aspect of the network design are the deconvolutional layers. While it is already shown in the paper, that their usage is beneficial, figure 5 illustrates the effect of different kernel sizes and layer depth. While any change to the kernel or the layer size seems to have a rather small effect on the quality when using two deconvolutional layer, the resulting changes are far larger when using only one. While on average a kernel size of 5 performs best for this case, it can not be concluded for sure, as the individual results again show significant variations. [h!]



(a) Different subsets of results that use one deconvolutional layer and vary in no setting but the kernel size of this layer.

(b) Different subsets of results that use two deconvolutional layers and vary in no setting but the kernel sizes of these layer.

(c) Different subsets of results that use two deconvolutional layer and vary in no setting but the depth of the first deconvolutional layer.

Figure 5: Visualisations of the prediction accuracy change, when varying the depth and kernel size of the deconvolutional layers.

1.5 Blocksize and Channel Type Coparisson

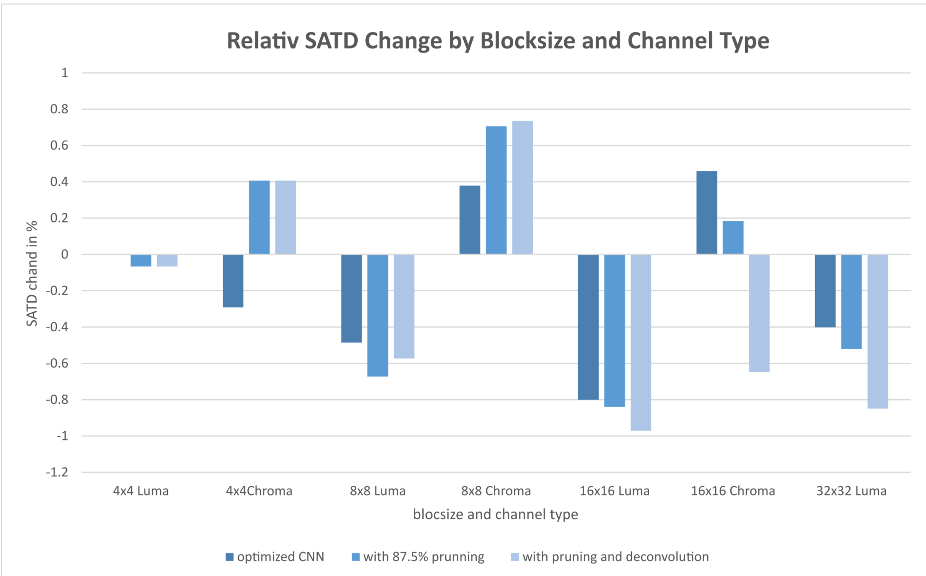


Figure 6: Relative SATD change on the valitadtion set between the different networks used in the final HM tests by blocksoze and channel type. The version using the new training material was used as refference.

Figure 6 shows the SATD chane between the different versions used in the final HM test. The 16x16 luma version here shows the same behaviour as in the pretests with additional gains for each of the consecutive versions compared to version using the partially coded training material. However, the conclusions from the pretests seem to be not easily transferable to other block sizes. For the 32x32 case the overall gain is already lower, while for the 8x8 luma version with a deconvolutional layer even decreases the gain compared to the pruned one. Even more problematic however is the transfer to the chroma prediction. For these networks the same type of architecture changes often causes a increas in the resulting validation SATD.

2 List of all Results

The tables 1 - 5 contain the result of all architecture variations that where tried during the pretest. For all of these the block size was fixed to $b = 16$. Also the test was limited to the luma channel prediction.
Table 6 shows the settings and results of the networks used in the final HM tests.

Table 1: Pretest results part 1

L_{SATD} abs.	flops abs.	settings n_{c}	$\mathbf{s}_{\mathrm{c}}$	$\mathbf{k}_{c}$	n_{f}	$\mathbf{s}_{\mathrm{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\mathrm{d}}$	$\mathbf{k}_{d}$	s_{par}	L_{SATD} change	flops change
5.6008	457 058	1	16	8	2	512,512	8	0	0	0	875	2.20	-93.32
6.0150	461 824	0	0	0	2	512,256	1	0	0	0	0	9.76	-93.25
5.5749	495 458	0	0	0	4	4 x 512	16	0	0	0	875	1.73	-92.76
5.6675	530 432	0	0	0	2	512,256	2	0	0	0	0	3.42	-92.25
5.5600	566 306	1	32	8	2	512,512	8	0	0	0	875	1.45	-91.73
6.0148	658 688	0	0	0	2	512,384	1	0	0	0	0	9.75	-90.38
5.6587	673 792	0	0	0	2	512,256	4	0	0	0	0	3.25	-90.16
5.5681	683 820	1	16	8	2	512,512	8	0	0	0	750	1.60	-90.01
5.5637	734 096	0	0	0	4	4 x 512	16	0	0	0	0.75	1.52	-89.28
6.0144	855 552	0	0	0	2	512,512	1	0	0	0	0	9.75	-87.50
5.6034	870 656	0	0	0	2	512,384	4	0	0	0	0	2.25	-87.28
5.5301	893 544	2	32,32	15,2	3	3 x 512	16	0	0	0	875	0.91	-86.95
5.6407	924 160	0	0	0	2	512,512	2	0	0	0	0	2.93	-86.50
5.5976	953 312	1	16	4	2	512,512	8	0	0	0	875	2.14	-86.08
5.5440	972 028	1	32	8	2	512,512	8	0	0	0	750	1.16	-85.80
5.6325	985 088	0	0	0	2	512,256	8	0	0	0	0	2.78	-85.61
5.5922	1 067 520	0	0	0	2	512,512	4	0	0	0	0	2.04	-84.41
5.5419	1 101 038	1	16	8	2	512,512	8	0	0	0	500	1.12	-83.92
5.5879	1 181 952	0	0	0	2	512,384	8	0	0	0	0	1.96	-82.74
5.5177	1 214 800	2	32,32	13,4	3	3 x 512	16	0	0	0	875	0.68	-82.26
5.5528	1 266 944	0	0	0	4	4 x 512	16	0	0	0	0.5	1.32	-81.50
5.5751	1 378 816	0	0	0	2	512,512	8	0	0	0	0	1.73	-79.86
5.5381	1 427 488	1	16	8	2	384,384	8	0	0	0	0	1.05	-79.15
5.5346	1 558 564	1	16	8	2	512,512	8	0	0	0	250	0.99	-77.24
5.6007	1 589 384	1	4	4	3	3 x 512	4	0	0	0	0	2.20	-76.79
5.5197	1 590 420	2	32,32	15,2	3	3 x 512	16	0	0	0	750	0.72	-76.77
5.5874	1 592 832	0	0	0	3	3 x 512	4	0	0	0	0	1.95	-76.74
9.1329	1 595 424	1	16	8	2	512,256	8	0	0	0	0	66.65	-76.70
5.5117	1 600 418	2	64,64	15,2	3	3 x 512	16	0	0	0	875	0.57	-76.63
5.4838	1 701 116	2	32,32	13,4	3	3 x 512	16	2	24,1	5,3	875	0.06	-75.15
5.6515	1 705 984	0	0	0	2	512,256	16	0	0	0	0	3.12	-75.08
5.5300	1 707 234	1	32	8	2	512,512	8	0	0	0	500	0.91	-75.07
5.7533	1 736 224	1	16	8	2	512,512	8	1	1	3	0	4.98	-74.64
5.5630	1 752 608	1	16	8	2	512,512	8	1	1	5	0	1.51	-74.40
5.5371	1 792 288	1	16	8	2	512,384	8	0	0	0	0	1.04	-73.82
5.5635	1 805 040	1	16	4	2	512,512	8	0	0	0	750	1.52	-73.64
5.6111	1 851 984	1	8	2	3	3 x 512	2	0	0	0	0	2.39	-72.95
5.5688	1 864 464	1	8	4	3	3 x 512	4	0	0	0	0	1.61	-72.77
5.6120	1 902 848	0	0	0	2	512,384	16	0	0	0	0	2.40	-72.21
5.5627	1 904 128	0	0	0	3	3 x 512	8	0	0	0	0	1.50	-72.19
5.5580	1 908 088	0	0	0	4	4 x 512	16	0	0	0	0.25	1.42	-72.13
5.5618	1 914 384	1	8	8	3	3 x 512	8	0	0	0	0	1.49	-72.04
5.6380	1 974 784	0	0	0	4	4 x 512	2	0	0	0	0	2.88	-71.16
5.5287	1 989 152	1	16	8	2	512,512	8	0	0	0	0	0.88	-70.95
5.6304	2 034 432	0	0	0	3	512,384,256	16	0	0	0	0	2.74	-70.29
5.5869	2 099 712	0	0	0	2	512,512	16	0	0	0	0	1.94	-69.33
5.5474	2 114 064	1	8	16	3	3 x 512	16	0	0	0	0	1.22	-69.12
5.5867	2 118 144	0	0	0	4	4 x 512	4	0	0	0	0	1.94	-69.06
5.5985	2 231 296	0	0	0	3	512,512,256	16	0	0	0	0	2.16	-67.41
5.5053	2 256 580	2	32,32	13,4	3	3 x 512	16	0	0	0	750	0.46	-67.04
5.5498	2 261 536	1	16	8	3	3 x 512	8	1	1	3	0	1.27	-66.97
5.5227	2 277 920	1	16	8	3	3 x 512	8	1	1	5	0	0.77	-66.73
5.5818	2 302 496	1	16	8	3	3 x 512	8	1	1	7	0	1.85	-66.37
5.6049	2 389 664	1	16	2	3	3 x 512	2	0	0	0	0	2.27	-65.10
5.5403	2 414 624	1	16	4	3	3 x 512	4	0	0	0	0	1.10	-64.73
5.5644	2 428 160	0	0	0	3	512,512,384	16	0	0	0	0	1.53	-64.54

Table 2: Pretest results part 2

L_{SATD} abs.	flops abs.	settings n_{c}	$\mathbf{s}_{\text{c}}$	$\mathbf{k}_c$	n_{f}	$\mathbf{s}_{\text{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\text{d}}$	$\mathbf{k}_d$	s_{par}	L_{SATD} change	flops change
5.5635	2 429 440	0	0	0	4	4 x 512	8	0	0	0	0	1.52	-64.52
5.5222	2 459 202	1	32	8	2	512,512	8	0	0	0	250	0.76	-64.08
5.5254	2 514 464	1	16	8	3	3 x 512	8	0	0	0	0	0.82	-63.28
5.6329	2 625 024	0	0	0	3	3 x 512	16	0	0	0	0	2.79	-61.66
5.4985	2 672 024	2	64,64	13,4	3	3 x 512	16	0	0	0	875	0.33	-60.97
5.5400	2 732 064	2	32,16	3,2	3	3 x 512	4	0	0	0	0	1.09	-60.10
5.4844	2 732 570	2	32,32	13,4	3	3 x 512	16	2	24,1	5,3	750	0.08	-60.09
9.1336	2 795 584	1	32	8	2	512,256	8	0	0	0	0	66.66	-59.17
5.5272	2 913 824	1	16	16	3	3 x 512	16	0	0	0	0	0.86	-57.44
9.1267	2 936 384	1	32	8	2	512,512	8	1	1	3	0	66.54	-57.11
5.5863	2 952 768	1	32	8	2	512,512	8	1	1	5	0	1.93	-56.87
5.5076	2 991 570	2	64,64	15,2	3	3 x 512	16	0	0	0	750	0.50	-56.31
5.5193	2 992 448	1	32	8	2	512,384	8	0	0	0	0	0.71	-56.29
5.5061	3 070 298	2	32,32	15,2	3	3 x 512	16	0	0	0	500	0.47	-55.16
5.5560	3 150 336	0	0	0	4	4 x 512	16	0	0	0	0	1.38	-53.99
5.5200	3 189 312	1	32	8	2	512,512	8	0	0	0	0	0.72	-53.42
5.5362	3 262 016	1	32	4	3	3 x 512	4	1	1	3	0	1.02	-52.36
5.5760	3 278 400	1	32	4	3	3 x 512	4	1	1	5	0	1.75	-52.12
5.5093	3 288 128	1	32	8	3	512,384,384	8	0	0	0	0	0.53	-51.98
5.5484	3 432 494	1	16	4	2	512,512	8	0	0	0	500	1.24	-49.87
5.5169	3 461 696	1	32	8	3	3 x 512	8	1	1	3	0	0.67	-49.44
5.6003	3 465 024	1	32	2	3	3 x 512	2	0	0	0	0	2.19	-49.39
9.1283	3 478 080	1	32	8	3	3 x 512	8	1	1	5	0	66.57	-49.20
5.5515	3 502 656	1	32	8	3	3 x 512	8	1	1	7	0	1.30	-48.84
5.5313	3 514 944	1	32	4	3	3 x 512	4	0	0	0	0	0.93	-48.66
5.5112	3 517 760	1	32	8	3	512,512,384	8	0	0	0	0	0.56	-48.62
5.5510	3 533 312	0	0	0	4	4 x 512	16	2	24,1	5,3	0	1.29	-48.40
5.5652	3 675 648	0	0	0	5	5 x 512	16	0	0	0	0	1.55	-46.32
5.5461	3 709 072	1	8	4	2	512,512	8	0	0	0	0	1.20	-45.83
5.5079	3 714 624	1	32	8	3	3 x 512	8	0	0	0	0	0.50	-45.75
5.5365	3 756 864	2	16,32	3,2	3	3 x 512	4	0	0	0	0	1.03	-45.13
5.5270	4 065 344	2	32,32	3,2	3	3 x 512	4	0	0	0	0	0.85	-40.62
5.5401	4 260 416	1	32	16	3	3 x 512	16	1	1	3	0	1.09	-37.78
5.5067	4 276 800	1	32	16	3	3 x 512	16	1	1	5	0	0.48	-37.54
5.5586	4 301 376	1	32	16	3	3 x 512	16	1	1	7	0	1.43	-37.18
5.4980	4 344 344	2	32,32	13,4	3	3 x 512	16	0	0	0	500	0.32	-36.55
5.4987	4 432 040	2	64,64	15,2	3	3 x 512	16	0	0	0	625	0.34	-35.27
5.4999	4 450 174	2	32,32	15,2	3	3 x 512	16	0	0	0	250	0.36	-35.00
5.5071	4 513 344	1	32	16	3	3 x 512	16	0	0	0	0	0.49	-34.08
5.5132	4 650 624	2	32,32	16,1	3	3 x 512	16	0	0	0	0	0.60	-32.08
5.6034	4 661 856	1	48	8	3	3 x 512	8	1	1	3	0	2.25	-31.91
5.5573	4 678 240	1	48	8	3	3 x 512	8	1	1	5	0	1.41	-31.67
5.5270	4 682 304	2	64,32	3,2	3	3 x 512	4	0	0	0	0	0.85	-31.61
5.5524	4 702 816	1	48	8	3	3 x 512	8	1	1	7	0	1.32	-31.31
5.4825	4 794 996	2	32,32	13,4	3	3 x 512	16	2	24,1	5,3	500	0.04	-29.97
5.5055	4 914 784	1	48	8	3	3 x 512	8	0	0	0	0	0.46	-28.22
5.5399	5 040 802	1	16	4	2	512,512	8	0	0	0	250	1.09	-26.38
5.6134	5 090 432	1	64	2	2	512,512	2	0	0	0	0	2.43	-25.65
5.4997	5 159 364	2	64,64	13,4	3	3 x 512	16	0	0	0	750	0.35	-24.65
9.1336	5 195 904	1	64	8	2	512,256	8	0	0	0	0	66.66	-24.11
5.5266	5 392 768	1	64	8	2	512,384	8	0	0	0	0	0.85	-21.24
5.5355	5 462 656	1	64	4	3	3 x 512	4	1	1	3	0	1.01	-20.22
5.5317	5 479 040	1	64	4	3	3 x 512	4	1	1	5	0	0.94	-19.98
5.5159	5 589 632	1	64	8	2	512,512	8	0	0	0	0	0.65	-18.36
5.6061	5 615 744	1	64	2	3	3 x 512	2	0	0	0	0	2.29	-17.98
5.5260	5 715 584	1	64	4	3	3 x 512	4	0	0	0	0	0.83	-16.52
5.4867	5 826 378	2	32,32	13,4	3	3 x 512	16	2	24,1	5,3	375	0.12	-14.90

Table 3: Pretest results part 3

L_{SATD} abs.	flops abs.	n_{c}	$\mathbf{s}_{\mathrm{c}}$	$\mathbf{k}_c$	n_{f}	$\mathbf{s}_{\mathrm{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\mathrm{d}}$	$\mathbf{k}_d$	s_{par}	L_{SATD} change	flops change
		settings											
9.1293	5 859 936	1	48	16	3	3 x 512	16	1	1	3	0	66.58	-14.41
5.4987	5 876 320	1	48	16	3	3 x 512	16	1	1	5	0	0.34	-14.17
5.4970	5 890 112	2	32,32	15,2	3	3 x 512	16	0	0	0	0	0.31	-13.97
5.5539	5 900 896	1	48	16	3	3 x 512	16	1	1	7	0	1.34	-13.82
5.4930	5 902 628	2	64,64	15,2	3	3 x 512	16	0	0	0	500	0.23	-13.79
5.5086	5 932 204	2	64,64	9,8	3	3 x 512	16	0	0	0	875	0.52	-13.36
5.5006	6 112 864	1	48	16	3	3 x 512	16	0	0	0	0	0.37	-10.72
5.5038	6 114 944	1	64	8	3	3 x 512	8	0	0	0	0	0.43	-10.69
5.5092	6 418 624	2	48,48	16,1	3	3 x 512	16	0	0	0	0	0.53	-6.25
5.4901	6 469 864	2	32,32	13,4	3	3 x 512	16	0	0	0	250	0.18	-5.51
5.5352	6 629 152	1	16	4	2	512,512	8	0	0	0	0	1.00	-3.18
5.5310	6 731 904	2	32,64	3,2	3	3 x 512	4	0	0	0	0	0.93	-1.68
5.4803	**6 846 860**	**2**	**32,32**	**13,4**	**3**	**3 x 512**	**16**	**2**	**24,1**	**5,3**	**250**	**0.00**	**0.00**
9.1337	6 962 304	2	64,64	3,2	2	512,256	4	0	0	0	0	66.66	1.69
5.5414	7 159 168	2	64,64	3,2	2	512,384	4	0	0	0	0	1.11	4.56
5.4906	7 180 672	2	32,32	14,3	3	3 x 512	16	0	0	0	0	0.19	4.88
5.4889	7 303 122	2	64,64	15,2	3	3 x 512	16	0	0	0	375	0.16	6.66
5.5375	7 356 032	2	64,64	3,2	2	512,512	4	0	0	0	0	1.04	7.44
5.5110	7 459 456	1	64	16	3	3 x 512	16	1	1	3	0	0.56	8.95
5.5514	7 475 840	1	64	16	3	3 x 512	16	1	1	5	0	1.30	9.19
5.5485	7 500 416	1	64	16	3	3 x 512	16	1	1	7	0	1.24	9.55
5.4982	7 612 274	2	64,64	13,4	3	3 x 512	16	0	0	0	625	0.33	11.18
5.5005	7 712 384	1	64	16	3	3 x 512	16	0	0	0	0	0.37	12.64
5.5314	7 881 344	2	64,64	3,2	3	3 x 512	4	0	0	0	0	0.93	15.11
5.4858	7 886 562	2	32,32	13,4	3	3 x 512	16	2	24,1	5,3	125	0.10	15.19
5.4931	8 237 696	1	64	16	4	4 x 512	16	0	0	0	0	0.23	20.31
5.5070	8 253 184	2	64,64	16,1	3	3 x 512	16	0	0	0	0	0.49	20.54
5.5438	8 284 736	2	32,32	13,4	3	3 x 512	16	1	1	3	0	1.16	21.00
5.5363	8 301 120	2	32,32	13,4	3	3 x 512	16	1	1	5	0	1.02	21.24
5.4946	8 537 664	2	32,32	13,4	3	3 x 512	16	0	0	0	0	0.26	24.69
5.5041	8 557 184	2	64,64	7,2	3	3 x 512	8	0	0	0	0	0.43	24.98
5.4885	8 577 376	2	48,48	15,2	3	3 x 512	16	0	0	0	0	0.15	25.27
5.4931	8 699 808	2	64,64	15,2	3	3 x 512	16	0	0	0	250	0.23	27.06
5.4834	8 705 600	2	32,32	13,4	3	3 x 512	16	2	16,1	5,3	0	0.06	27.15
5.4859	8 738 368	2	32,32	13,4	3	3 x 512	16	2	16,1	5,5	0	0.10	27.63
5.5000	9 977 984	2	32,32	12,5	3	3 x 512	16	0	0	0	0	0.36	45.73
5.4909	10 125 888	2	64,64	15,2	3	3 x 512	16	0	0	0	125	0.19	47.89
5.4882	10 158 048	2	64,64	13,4	3	3 x 512	16	0	0	0	500	0.14	48.36
5.5277	10 503 424	2	64,64	2,3	3	3 x 512	4	0	0	0	0	0.86	53.40
5.4930	11 012 416	2	48,48	14,3	3	3 x 512	16	0	0	0	0	0.23	60.84
5.5038	11 326 720	2	64,64	6,3	3	3 x 512	8	0	0	0	0	0.43	65.43
5.4903	11 530 880	2	64,64	15,2	3	3 x 512	16	0	0	0	0	0.18	68.41
5.4963	11 596 952	2	64,64	9,8	3	3 x 512	16	0	0	0	750	0.29	69.38
5.4934	12 613 998	2	64,64	13,4	3	3 x 512	16	0	0	0	375	0.24	84.23
5.4930	12 801 664	2	32,64	13,4	3	3 x 512	16	0	0	0	0	0.23	86.97
5.4955	13 626 944	2	64,32	13,4	3	3 x 512	16	0	0	0	0	0.28	99.02
5.4909	13 746 784	2	48,48	13,4	3	3 x 512	16	0	0	0	0	0.19	100.78
5.5078	14 268 736	2	32,32	9,8	2	512,384	16	0	0	0	0	0.50	108.40
5.5470	14 435 968	2	64,64	2,2	3	3 x 512	4	0	0	0	0	1.22	110.84
5.5163	14 465 600	2	32,32	9,8	2	512,512	16	0	0	0	0	0.66	111.27
5.5083	14 482 048	2	64,64	5,4	2	512,512	8	0	0	0	0	0.51	111.51
5.5058	14 990 912	2	32,32	9,8	3	3 x 512	16	0	0	0	0	0.47	118.95
5.5045	15 007 360	2	64,64	5,4	3	3 x 512	8	0	0	0	0	0.44	119.19
5.4939	15 092 886	2	64,64	13,4	3	3 x 512	16	0	0	0	250	0.25	120.44
5.4890	15 443 200	2	64,64	14,3	3	3 x 512	16	0	0	0	0	0.16	125.55
5.4954	16 411 264	2	48,64	13,4	3	3 x 512	16	0	0	0	0	0.28	139.69
5.4956	16 805 824	2	48,48	12,5	3	3 x 512	16	0	0	0	0	0.28	145.45

Table 4: Pretest results part 4

L_{SATD} abs.	flops abs.	n_{c}	$\mathbf{s}_{\text{c}}$	$\mathbf{k}_c$	n_{f}	settings $\mathbf{s}_{\text{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\text{d}}$	$\mathbf{k}_d$	s_{par}	L_{SATD} change	flops change
5.4937	16 823 904	2	64,48	13,4	3	3 x 512	16	0	0	0	0	0.24	145.72
5.4922	17 117 878	2	64,64	9,8	3	3 x 512	16	0	0	0	625	0.22	150.01
5.4883	17 556 872	2	64,64	13,4	3	3 x 512	16	0	0	0	125	0.15	156.42
5.5011	19 645 184	2	64,64	4,5	3	3 x 512	8	0	0	0	0	0.38	186.92
5.4865	19 696 768	2	64,64	13,4	3	512,256,512	16	2	16,1	5,5	0	0.11	187.68
5.4952	19 754 112	2	64,64	13,4	3	512,512,256	16	2	16,1	5,5	0	0.27	188.51
5.4835	19 767 936	2	64,64	13,4	3	3 x 512	16	1	1	3	0	0.06	188.72
9.1333	19 775 104	2	64,64	13,4	3	3 x 512	16	1	1	4	0	66.66	188.82
5.4854	19 784 320	2	64,64	13,4	3	3 x 512	16	1	1	5	0	0.09	188.95
5.4919	19 795 584	2	64,64	13,4	3	3 x 512	16	1	1	6	0	0.21	189.12
5.4884	19 808 896	2	64,64	13,4	3	3 x 512	16	1	1	7	0	0.15	189.31
5.4880	19 926 656	2	64,64	13,4	3	3 x 512	16	2	16,1	3,3	0	0.14	191.03
5.4840	19 959 168	2	64,64	13,4	3	512,384,512	16	2	16,1	5,5	0	0.07	191.51
5.4827	19 959 424	2	64,64	13,4	3	3 x 512	16	2	16,1	3,5	0	0.04	191.51
5.4887	19 973 760	2	64,64	13,4	3	3 x 512	16	2	8,1	5,3	0	0.15	191.72
5.4847	19 987 840	2	64,64	13,4	3	512,512,384	16	2	16,1	5,5	0	0.08	191.93
5.4820	19 990 144	2	64,64	13,4	3	3 x 512	16	2	8,1	5,5	0	0.03	191.96
5.4896	20 020 864	2	64,64	13,4	3	3 x 512	16	0	0	0	0	0.17	192.41
5.4817	20 188 800	2	64,64	13,4	3	3 x 512	16	2	16,1	5,3	0	0.03	194.86
5.4813	20 221 568	2	64,64	13,4	3	3 x 512	16	2	16,1	5,5	0	0.02	195.34
5.5053	20 317 792	2	32,48	9,8	3	3 x 512	16	0	0	0	0	0.46	196.75
5.4806	20 403 840	2	64,64	13,4	3	3 x 512	16	2	24,1	5,3	0	0.01	198.00
5.4809	20 452 992	2	64,64	13,4	3	3 x 512	16	2	24,1	5,5	0	0.01	198.72
5.4988	20 762 176	2	48,32	9,8	3	3 x 512	16	0	0	0	0	0.34	203.24
5.5050	22 949 038	2	64,64	9,8	3	3 x 512	16	0	0	0	500	0.45	235.18
5.5125	24 922 496	2	32,64	9,8	2	512,384	16	0	0	0	0	0.59	264.00
5.5041	25 119 360	2	32,64	9,8	2	512,512	16	0	0	0	0	0.43	266.87
5.5086	25 289 344	2	64,64	3,6	3	3 x 512	8	0	0	0	0	0.52	269.36
5.4976	25 297 664	2	64,64	12,5	3	3 x 512	16	0	0	0	0	0.32	269.48
5.4972	25 644 672	2	32,64	9,8	3	3 x 512	16	0	0	0	0	0.31	274.55
5.5163	25 811 264	2	64,32	9,8	2	512,384	16	0	0	0	0	0.66	276.98
5.5139	26 008 128	2	64,32	9,8	2	512,512	16	0	0	0	0	0.61	279.85
5.5038	26 533 440	2	64,32	9,8	3	3 x 512	16	0	0	0	0	0.43	287.53
5.4964	27 691 648	2	64,64	13,4	3	3 x 1024	16	0	0	0	0	0.29	304.44
5.4997	28 218 976	2	48,48	9,8	3	3 x 512	16	0	0	0	0	0.35	312.14
5.5012	28 619 028	2	64,64	9,8	3	3 x 512	16	0	0	0	375	0.38	317.99
5.5199	30 082 816	2	64,64	4,3	3	3 x 512	8	0	0	0	0	0.72	339.37
5.4912	31 310 464	2	64,64	11,6	3	3 x 512	16	0	0	0	0	0.20	357.30
5.5097	31 992 064	2	64,64	2,7	3	3 x 512	8	0	0	0	0	0.54	367.25
5.5004	34 341 020	2	64,64	9,8	3	3 x 512	16	0	0	0	250	0.37	401.56
5.5302	35 398 144	1	64	11	3	3 x 512	16	0	0	0	0	0.91	417.00
5.5042	35 675 776	2	48,64	9,8	3	3 x 512	16	0	0	0	0	0.44	421.05
5.5023	36 120 160	2	64,48	9,8	3	3 x 512	16	0	0	0	0	0.40	427.54
5.5179	36 998 784	2	64,64	13,4	3	2048,1024,512	16	0	0	0	0	0.69	440.38
5.5052	40 037 010	2	64,64	9,8	3	3 x 512	16	0	0	0	125	0.45	484.75
5.5231	41 395 840	2	64,64	3,4	3	3 x 512	8	0	0	0	0	0.78	504.60
5.5212	42 526 336	2	64,64	9,8	2	256,256	16	0	0	0	0	0.75	521.11
5.5106	43 821 184	2	64,64	9,8	2	384,384	16	0	0	0	0	0.55	540.02
5.5096	43 985 280	2	64,64	9,8	2	384,512	16	0	0	0	0	0.54	542.42
5.5238	44 787 840	2	64,64	9,8	2	512,256	16	0	0	0	0	0.79	554.14
5.6432	44 886 272	2	64,64	9,8	2	512,320	16	0	0	0	0	2.97	555.57
5.5086	44 984 704	2	64,64	9,8	2	512,384	16	0	0	0	0	0.52	557.01
5.5430	45 108 736	1	64	9	3	3 x 512	16	0	0	0	0	1.14	558.82
5.5001	45 116 288	2	64,64	9,8	3	512,384,256	16	0	0	0	0	0.36	558.93
5.5106	45 181 568	2	64,64	9,8	2	512,512	16	0	0	0	0	0.55	559.89
5.7798	45 313 152	2	64,64	9,8	3	512,512,256	16	0	0	0	0	5.47	561.81

Table 5: Pretest results part 5

L_{SATD}	flops					settings						L_{SATD}	flops
abs.	abs.	n_{c}	$\mathbf{s}_{\text{c}}$	$\mathbf{k}_c$	n_{f}	$\mathbf{s}_{\text{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\text{d}}$	$\mathbf{k}_d$	s_{par}	change	change
5.5018	45 510 016	2	64,64	9,8	3	512,512,384	16	0	0	0	0	0.39	564.68
5.4883	45 706 880	2	64,64	9,8	3	3 x 512	16	0	0	0	0	0.15	567.56
5.5150	53 377 664	2	64,64	9,8	3	3 x 1024	16	0	0	0	0	0.63	679.59
5.5378	62 684 800	2	64,64	9,8	3	2048,1024,512	16	0	0	0	0	1.05	815.53
5.5065	73 919 744	2	64,64	6,11	3	3 x 512	16	0	0	0	0	0.48	979.62
5.5190	75 233 920	2	64,64	9,4	3	3 x 512	16	0	0	0	0	0.71	998.81
5.5081	97 749 760	2	64,64	4,13	3	3 x 512	16	0	0	0	0	0.51	1327.66
5.5212	255 072 000	2	64,64	4,9	3	3 x 512	16	0	0	0	0	0.75	3625.39

Table 6: SATD results of the finally used networks

L_{SATD}	flops						settings						
abs.	abs.	n_{c}	$\mathbf{s}_{\text{c}}$	$\mathbf{k}_c$	n_{f}	$\mathbf{s}_{\text{f}}$	n_{r}	n_{d}	$\mathbf{s}_{\text{d}}$	$\mathbf{k}_d$	s_{par}	b	luma
3.8478	2 787 488	2	64,64	3,2	3	3 x 512	4	0	0	0	0	4	1
3.2029	6 246 464	2	64,64	3,2	3	3 x 512	4	0	0	0	0	4	0
4.7931	4 452 608	2	64,64	3,2	3	3 x 512	4	0	0	0	0	8	1
3.6191	11 198 720	2	64,64	3,2	3	3 x 512	4	0	0	0	0	8	0
5.5192	7 881 344	2	64,64	3,2	3	3 x 512	4	0	0	0	0	16	1
4.0738	21 300 224	2	64,64	3,2	3	3 x 512	4	0	0	0	0	16	0
6.0421	15 132 800	2	64,64	3,2	3	3 x 512	4	0	0	0	0	32	1
3.8478	2787488	2	64,64	3,2	3	3 x 512	4	0	0	0	0	4	1
3.1935	4 627 520	2	64,64	3,2	3	3 x 512	4	0	0	0	0	4	0
4.7698	4 800 768	2	64,64	7,2	3	3 x 512	8	0	0	0	0	8	1
3.6328	12 503 296	2	64,64	7,2	3	3 x 512	8	0	0	0	0	8	0
5.4751	20 020 864	2	64,64	13,4	3	3 x 512	16	0	0	0	0	16	1
4.0925	69 722 624	2	64,64	13,4	3	3 x 512	16	0	0	0	0	16	0
6.0178	165 385 344	2	64,64	25,8	3	3 x 512	32	0	0	0	0	32	1
3.8453	356 560	2	64,64	3,2	3	3 x 512	4	0	0	0	87.5	4	1
3.2159	1 258 824	2	64,64	3,2	3	3 x 512	4	0	0	0	87.5	4	0
4.7608	653 516	2	64,64	7,2	3	3 x 512	8	0	0	0	87.5	8	1
3.6446	4 030 818	2	64,64	7,2	3	3 x 512	8	0	0	0	87.5	8	0
5.4729	2 688 822	2	64,64	13,4	3	3 x 512	16	0	0	0	87.5	16	1
4.0813	12 980 638	2	64,64	13,4	3	3 x 512	16	0	0	0	87.5	16	0
6.0106	21 344 220	2	64,64	25,8	3	3 x 512	32	0	0	0	87.5	32	1
3.8453	356 560	2	64,64	3,2	3	3 x 512	4	0	0	0	87.5	4	1
3.2159	1 258 824	2	64,64	3,2	3	3 x 512	4	0	0	0	87.5	4	0
4.7656	632 256	2	64,64	7,2	3	3 x 512	8	1	1	5	87.5	8	1
3.6457	3 978 856	2	64,64	7,2	3	3 x 512	8	1	1	5	87.5	8	0
5.4657	3 155 862	2	64,64	13,4	3	3 x 512	16	2	24,1	5,3	87.5	16	1
4.0474	13 217 832	2	64,64	13,4	3	3 x 512	16	2	24,1	5,3	87.5	16	0
5.9908	22 092 812	2	64,64	25,8	3	3 x 512	32	3	24,24,1	5,3,3	87.5	32	1