Version	Folder	Date	Owner	Methodology	Overall MSF	Commit Hash	Takeaways
• 8181011	. Jiuei	Date		- ResNet50 (Places)	O TOTAL MOL	Committee and the committee an	· · · · · · · · · · · · · · · · · · ·
v1.0	v1.0	27/05/2025	Leo Kling	- Hybrid Training: train and eval on synth, train and eval on real - synth dataset: leo(1,,10), jojo(1,,6, 9)	0.1075831652	46f9c2c7b4ff8340b6d4a95a700b99debe69415e	
v1.1	v1.1	28/05/2025	Leo Kling	- same architecture as v1.1 - synth dataset: leo(1,,10), jojo(1,,5, 9), milan (1,, 4), markus (1,, 10), david (1,, 3)	0.0934	46f9c2c7b4ff8340b6d4a95a700b99debe69415e	
v1.2	v1.2	28/05/2025	Leo Kling	- same architecture as v1.1 - synth dataset: add jojo (6, 7, 8)	0.1144	46f9c2c7b4ff8340b6d4a95a700b99debe69415e	
			-	- same architecture as v1.1 - split synth train data into hybrid and "Image Source Method (ISM) and Ray-Radiosity (RR)" rooms	hybrid: 0.1128242090344429 non-hybrid: 0.1078229770064354		- hybrid simulation does not yield better results - 4.4% difference in mse, prob. in margin of error
	v1.3 v1.4	30/05/2025	Jonathan Kron Jonathan Kron	 train thrice (only hybrid, only non hybrid, combined) add more parameters to data augmentation (incl. random perspective, more agressive crop and rotation) 	combined: 0.11138167232275009 0.09731775522	41fdaf7d1c72031651795d9d32e3b43780942acf 077e802f8a92f1580e4ea6d2aa89b99b71d4b08a	meaning no noticable difference - slight improvement to v1.3
v1.5	v1.5	31/05/2025	Jonathan Kron	improved architecture, trained on combined synth data ensamble of 5 models ensamble of 5 models einigle model size: 100 083 KB ensemble of 5 size: 327 244 KB (over 250MB soft limit)	- en 1: 0.0923 - en 2: 0.0923 - en 3: 0.0930 - combined: 0.0925	39b007cc06b8c2d092ef530aaae55f5798a03d39	- improved architecture shows little improvement to v1.1 - no improvement by using an ensemble
			Leo Kling	- gradient accumulation (batch size 128 = 2°64) - higher initial I.R with OneCydeLR - better early stopping with min_delta - gradient clipping - gradient clipping - gradient size in the		b75014b8cd119d3aa4b498ea570d4c0fbfc59f12	- only slightly worse - a lot smaller (cs. 50MB)
			_	- same as v1.6			
			Leo Kling	- resnet50 - same as v1.6		4319b2706e11e280331d4a0ab1babe0a3a6fe070	
v1.8	v1.8	01/06/2025	Leo Kling	- progressive resolution (224px -> 320px -> 448px) - same architecture as v1.5 (without ensemble)		53aa6171a8535aff33166823a33038094d85a1aa	- resolution of 224px seems fine? - real and synth at the same time makes the loss jump a lot
v1.9	v1.9	05/06/2025	Leo Kling	- mix train and real data during training - no pretraining	0.2803 (best: 0.157)	be1c8ad48384de0b5c18062092df5da1eec12f34	- introduces a lot of inconcistencies/instability
-		05/06/2025	David Klein	- no perdalining - tried U-Net architecture - calculated mean/std for our dataset		5ea2370fd20c7df53b9703eeccb14d242abb2a07	-takes about 12min per epoch -model is too large to run on my pc
v1.10	v1.10	06/06/2025	Leo Kling	- same architecture as v1.5 (without ensemble)	0.0919	8aa2af3bd8c614c1d110609259bc06f4d8367160	- idk why it is slightly better than v1.5, same architecture - eval wise this does not make a lot of sense, basically always predicts the same RT
				- same architecture as v1.5 (without ensemble) - different pretrained backbones - DINO/2 - OpenCLIP (VIT-B-32) - OpenCLIP (VIT-L-14) - OpenCLIP (VIT-L-14)			there is a lot going on here, got the following output during training: dinov2: val loss: 0.1789, Val R-2.03867 cip : Val Loss: 0.1789, Val R-2.03312 cip : (Val Loss: 0.1799, Val R-2.03352 efficientnet_b4 : Val Loss: 0.1799, Val R-2.03555 efficientnet_b4 : Val Loss: 0.1799, Val R-2.03551 cip_convnext : Val Loss: 0.1796, Val R-2.03710
v1.11	v1.11	06/06/2025	Leo Kling	- EfficientNet (b4) - same architecture as v1.4	?	f3feb558c5a8117cf0b994712657d6936bb13691	Can't replicate those numbers on test-set sadly. Evaluation results have a lot of variability (from r2 values close to 0 to -35)
v1.12		06/06/2025	Jonathan Kron	- different pretained backbones - resnetSU places365 - efficienting LH 4 - commax, base - densenet169	- resnet50_places365: 0.0973177552223 - efficientnet_b4: 0.10257571190595627 - convnext_base: 0.10787791758775711 - densenet169: 0.12339576333761215	2 87750343499e2372c4f854b52b6da3b875a4306c	- cannot beat resnet 50
v1.13	v1.13	09/06/2025	Philipp Wendt	resnext50 32x4d	0.1233	500037d50222fba6213c3d0ba802fac766bd3629	- not better than resnet50-places365 (haven toked into changes to the model by leo and jonathan yet, will update)
v1.14	v1.14	06/10/2025	Milan Jezovsek	- freshly branched from main and then adjusted the architecture based on claude suggestion - therefore its based on v1.0 it hink - the model allegedly has spatial attention now ??? sure daddy claude, whatever u say.	"overall": { "mse": 0.09930435568094254, "mses": 0.3151259362238255, "mae": 0.21606934070587158, "2": -0.27137336134910583 }	d55c6ac42e650d8b20a376ccd9fb77d84ae3f444	- maybe combinable with other improvements ? - will test soon
v1.15	v1.15	11.06.2025	Milan Jezovsek	-based on v1.14 -spatial attention all the way -only implemented half of it before	"overall": { "mse": 0.0975494384765625, "mses": 0.3123290548068855, "mse": 0.22261279821395874, "2": -0.0375484935939312 }	93ffa9268dc871353328e1fe0f498ac1bf40b798	- same as before
				-U-Net with improved bottleneck and early dropout	"overall": { "mse": 0.09463762491941452, "mse": 0.3076322884864567, "mae": 0.20528370141983032, "r2": -0.24498794972896576		-best /2 i was able to achieve so far
v1.16	v1.16	12.06.25	David Klein	-added function to save best synthetic model instead of the last epoch	}	2fb32587093b8acd1db66b99457fd825f91b80e6	-not great not terrible
v1.10.1	v1.10.1	29.06.25	David Klein	-Architecture from v1.10 (ResNet50 places365) -Did a normal run + one run using real data only and one run using synth data only.		25396c9f9db022a4874d5e8634ce9f0902af360c	-Same params, real data only ran way longer until early stopping hit -Real data gives better results than both combinedSynth data on it's own gives terrible results -Combined data produces the same result for almost all images -Real data only produces more varied results on the test dataReal data only sevents to somewhat recognize rooms or the characteristics, images from the same room produce similar outcomes, even if they are not accurate to the actual RT80 values
v1.4.1		28.06.25	Philipp Wendt	- based on 1.4 - freeze backbone - train on excl. synth - val on excl. synth - test on real - later early slopping - more regulation - gradient olipping	"overall": { "mse": 1.0948086977005005, "mse": 1.0463310550556545, "mae": 0.7197692394256592, "r2": -0.26217177510261536 }	e70f24496180807da420f66e2484d009cf314470	- bad - worst in low freqs