



Universiteit
Leiden
The Netherlands

On the optimization of imaging pipelines

Schoonhoven, R.A.

Citation

Schoonhoven, R. A. (2024, June 11). *On the optimization of imaging pipelines*. Retrieved from <https://hdl.handle.net/1887/3762676>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3762676>

Note: To cite this publication please use the final published version (if applicable).

On the optimization of imaging pipelines

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 11 juni 2024
klokke 13:45 uur

door

Richard Arnoud Schoonhoven
geboren te Utrecht, Nederland
in 1995

Promotor:

Prof.dr. K. J. Batenburg

Co-promotores:

Dr. D. M. Pelt

Dr. B. J. C. van Werkhoven

Promotiecommissie:

Prof.dr. T. H. W. Bäck

Prof.dr. R. V. van Nieuwpoort

Prof.dr. A. Plaat

Prof.dr. ir. M. Staring

Dr. L. Cao

Dr. N. Viganò (European Synchrotron Radiation Facility)

The research presented in this dissertation was carried out at the Centrum Wetenschap & Informatica (CWI) in Amsterdam.

Financial support was provided by the Netherlands Organisation for Scientific Research (NWO) in the framework of the NWA-ORC Call (file number NWA.1160.18.316). This work is also financially supported by the Netherlands Organization for Scientific Research (NWO), project number 639.073.506 and 016.Veni.192.235. This work has made use of the experimental systems on the Dutch national e-infrastructure with the support of the SURF Cooperative. The DAS-6 cluster is funded through NWO-M and Open Competition (617.001.204) grants.

Ridderprint, the Netherlands

© 2024 Richard A. Schoonhoven,

Contents

1	Introduction	1
1.1	X-ray CT pipelines	2
1.2	Deep learning for imaging	6
1.3	Graphics Processing Units	8
1.4	Research questions	11
2	Real-time segmentation for tomographic imaging	19
2.1	Introduction	19
2.2	Method	21
2.3	Results and Discussion	24
2.4	Conclusions	31
3	Auto-differentiation for CT workflows	33
3.1	Introduction	33
3.2	Related work	34
3.3	Methodology	36
3.4	Case studies	38
3.5	Discussion	55
3.6	Conclusion	56
4	LEAN: graph-based pruning for convolutional neural networks by extracting longest chains	59
4.1	Introduction	59
4.2	Related work	60
4.3	Preliminaries	61
4.4	Method	63
4.5	Experimental setup	67
4.6	Results	69
4.7	Conclusion	71
5	Benchmarking optimization kernels for auto-tuning GPU kernels	73
5.1	Introduction	73
5.2	Related work	75
5.3	Method: Optimization problem	77
5.4	Implementation	83

5.5	Experimental Setup	84
5.6	Results: Benchmarking optimization algorithms on runtime	87
5.7	Quantifying GPU tuning difficulty	96
5.8	Conclusion	103
6	Going green: optimizing GPUs for energy efficiency through model-steered auto-tuning	105
6.1	Introduction	105
6.2	Related Work	106
6.3	Methodology	108
6.4	Experimental setup	111
6.5	Experimental results	114
6.6	Conclusions	127
7	Conclusion and outlook	129
7.1	Contributions and limitations	129
7.2	Outlook	131
	Bibliography	133
	List of publications	155
	Samenvatting	157
	Summary	159
	Curriculum Vitae	161
	Acknowledgments	163
A	Appendices	165
A.1	Appendix: (LEAN) graph-based pruning for convolutional neural networks by extracting longest chains	166
A.2	Appendix: Benchmarking optimization algorithms for auto-tuning GPU kernels	168