



Universiteit
Leiden
The Netherlands

Exploring deep learning for data-efficient and generalizable visual recognition

He, K.

Citation

He, K. (2026, September 22). *Exploring deep learning for data-efficient and generalizable visual recognition*. Retrieved from <https://hdl.handle.net/1887/4310276>

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/1887/4310276>

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

Exploring Deep Learning for Data-Efficient and Generalizable Visual Recognition

Kai He

Copyright © 2026 Kai He, All Rights Reserved

ISBN: 978-90-9043099-7

Printed by UFB Grafisch centrum, The Netherlands

Cover design: Kai He

Exploring Deep Learning for Data-Efficient and Generalizable Visual Recognition

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 22 september 2026
klokke 13.00 uur

door

Kai He

geboren te Yangzhou, China
in 1995

Promotor: Prof.dr. M.S.K. Lew
Co-promotor: Dr. E.M. Bakker

Promotiecommissie: Prof.dr. T.H.W. Bäck
Prof.dr. M.M. Bonsangue
Prof.dr. A. Plaat
Prof.dr. M.R. Spruit
Prof.dr. S. Satoh (National Institute of Informatics, Tokyo)
Dr. K.J. Wolstencroft (Amsterdam UMC, Amsterdam)



Kai He was financially supported through the China Scholarship Council (CSC) to participate in the PhD programme of Leiden University. Grant number 202106870027.

The research in this thesis was performed at the LIACS MediaLab, Leiden University, The Netherlands, and we would like to thank the NVIDIA Corporation for the donation of GPU cards.

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Background and Related Work	5
1.2.1	Metric Learning	5
1.2.2	Feature Enrichment	6
1.2.3	Knowledge Transfer	8
1.3	Thesis Outline and Research Questions	10
1.4	Main Contributions	13
1.4.1	Task Contributions	13
1.4.2	Technical Contributions	14
2	Few-shot & meta-learning methods for image understanding: a survey	17
2.1	Introduction	18
2.2	The framework of few-shot image classification	19
2.2.1	Notation and definitions	19
2.2.2	Datasets	20
2.2.3	Evaluation process of few-shot image classification	21
2.3	Paradigms of meta-learning for few-shot image classification	22
2.3.1	Metric-based meta-learning	22
2.3.2	Model-based meta-learning	26
2.3.3	Optimization-based meta-learning	29
2.3.4	Other methods	32
2.4	Major Challenges and future directions	33
2.4.1	Limitations and challenges	36
2.4.2	Future directions	36
2.5	Chapter Conclusions	37
3	ProtoSimNet: Exploiting Prototype Similarity for Few-shot Learning	39
3.1	Introduction	40
3.2	Related Work	41
3.3	Proposed Method	44

3.3.1	Problem Formulation	44
3.3.2	Cosine Sharpening Transformation	44
3.3.3	Elastic Regularization	45
3.4	Experiments	47
3.4.1	Datasets and Implementation details	47
3.4.2	Evaluations	47
3.4.3	Ablative Analysis	48
3.5	Chapter Conclusions	53
4	Attention-guided Feature Pyramid Network for few-shot learning	55
4.1	Introduction	56
4.2	Related Work	60
4.2.1	Few-Shot Learning Paradigm	60
4.2.2	Attention Mechanisms for Feature Enrichment	60
4.3	Proposed Methods	61
4.3.1	Multi-Scale Feature Extraction	62
4.3.2	Feature Pyramid Construction	62
4.3.3	Simple Parameter-Free Attention Module (SimAM)	62
4.3.4	Cross-Layer Spatial Attention Module (CSAM)	63
4.3.5	Classification	63
4.4	Experiments	64
4.4.1	Benchmarks and Implementation Details	64
4.4.2	Experimental results	64
4.4.3	Ablation Study	65
4.5	Chapter Conclusions	66
5	Adaptive Multi-scale Feature Fusion Network for Few-shot Learning	69
5.1	Introduction	70
5.2	Related Work	71
5.3	Adaptive multi-scale feature fusion	72
5.3.1	Problem formulation and notions	72
5.3.2	Feature enrich module	72
5.3.3	Feature fusion module	73
5.4	Experiments	74
5.4.1	Datasets, implementation details and evaluation protocol	75
5.4.2	Experimental results and ablative analysis	76
5.5	Chapter Conclusions	77
6	Dual Selective Knowledge Transfer for Few-Shot Classification	79
6.1	Introduction	80
6.2	Related Work	82
6.3	Dual Selective Knowledge Transfer	83
6.3.1	Problem Formulation	83

6.3.2	Multi-task Learning	84
6.3.3	Dual Selective Knowledge Distillation	84
6.4	Experiments	86
6.4.1	Datasets and Implementation details	86
6.4.2	Evaluations	87
6.4.3	Ablative Analysis	89
6.5	Chapter Conclusions	93
7	Conclusions	95
7.1	Main Findings	95
7.2	Limitations	96
7.2.1	Algorithmic perspective	97
7.2.2	Practical perspective	97
7.3	Future Research Directions	98
7.3.1	Metric Learning	98
7.3.2	Feature Enrichment	98
7.3.3	Knowledge Transfer	99
	Bibliography	101
	List of Abbreviations	119
	English Summary	121
	Nederlandse Samenvatting	123
	Curriculum Vitae	127
	Publication List	129