



Universiteit
Leiden

The Netherlands

Using deep learning models in image processing

Xiong, Z.

Citation

Xiong, Z. (2026, May 19). *Using deep learning models in image processing*. Retrieved from <https://hdl.handle.net/1887/4303646>

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/1887/4303646>

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

Using Deep Learning Models in Image Processing

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 19 mei 2026

klokke 11:30 uur

door

Zhan Xiong

geboren te Chongqing, China

in 1988

Promotor:

Prof. dr. ir. F.J. Verbeek

Co-promotor:

Dr. X. Tang (Southwest University)

Promotiecommissie:

Prof. dr. T.H.W. Bäck

Prof. dr. M.M. Bonsangue

Prof. dr. M.S.K. Lew

Dr. D.M. Pelt

Dr. X. Jiang (Shanghai University)

Prof. dr. E.O. Postma (Tilburg University)

ISBN: 978-94-6537-559-5

Cover designed by Shaogui Xiong.

Printed by Ridderprint.

Copyright © 2026 Zhan Xiong. All rights reserved. No part of this thesis may be reproduced in any form or by any means without the permission of the author.

The research presented in this dissertation was financially supported by the China Scholarship Council (CSC). Grant No.201306990022.

Contents

1	Introduction	1
1.1	Background	1
1.2	Preliminaries	3
1.2.1	Data Sets for Computer Vision Studies	3
1.2.2	Importance of Feature Space Learning in Computer Vision . .	4
1.2.3	Introduction to Attention Mechanisms	5
1.3	Research Questions	6
1.4	Thesis Structure	6
1.5	Terminologies and Definitions	8
2	Survey	11
2.1	Introduction	13
2.2	Developments	16
2.2.1	Convolutional Neural Networks	17
2.2.2	Visual Transformers	31
2.2.3	Hybrid Networks	46
2.2.4	Development Trend Summary	52
2.3	Techniques	53
2.3.1	Neural Architecture Searching	53
2.3.2	Neural Architecture Compression	54
2.3.3	Self-supervised/unsupervised learning	55
2.3.4	Knowledge Distillation	57
2.3.5	Diffusion Models	57
2.3.6	Deep Unfolding Techniques	58
2.3.7	Section Summary	59
2.4	Techniques	60

Contents

2.4.1	Classification	60
2.4.2	Object Detection	69
2.4.3	Image Generation	74
2.4.4	Universal Models	79
2.4.5	Medical Data Analysis	82
2.5	Conclusions	84
2.5.1	Discussions	84
2.5.2	Conclusion	87
3	Semantic Segmentation	89
3.1	Introduction	91
3.2	Related Work	95
3.2.1	Variational Bayesian methods	95
3.2.2	Clustering Algorithms	96
3.2.3	Attention Mechanism	98
3.3	Variational Gaussian Mixture Embedded Attention	99
3.3.1	Problem Formulation and Notation	100
3.3.2	Variational Bayesian Inference with Geometric Regularization	100
3.3.3	Cross-Attention as Concept Distillation	104
3.4	Implementation Details	106
3.4.1	Online Bayesian Updates for Batch Processing	106
3.4.2	Adaptions to Deep Learning	108
3.5	Experiment	111
3.5.1	Dataset and Metrics	111
3.5.2	Performance Analysis	113
3.5.3	Ablation Study	121
3.5.4	Limitations	126
3.6	Conclusions	128
4	Weakly Supervised Semantic Segmentation	131
4.1	Introduction	133
4.2	Related Work	136
4.3	Method	137
4.3.1	Probabilistic Framework	138
4.3.2	Implementation	141
4.3.3	Loss Function	143
4.4	Experiments and Results	144

4.4.1	Experimental Setup	144
4.4.2	Results and Analysis	144
4.4.3	Ablation Study	147
4.5	Conclusion	151
5	Instance Segmentation	153
5.1	Introduction	155
5.2	Related Work	160
5.3	Methods	162
5.3.1	Preliminaries	162
5.3.2	Dense Instance Segmentation Model	163
5.3.3	Lesion classifier	166
5.3.4	Implementation Details	167
5.4	Results	171
5.4.1	Datasets	171
5.4.2	Experiments	173
5.4.3	Complexity Assessment of Model Training	179
5.4.4	Ablation Study	180
5.5	Conclusion and Future Work	186
6	Conclusions and Discussion	187
6.1	Main Findings	187
6.2	Discussion	190
6.2.1	Algorithmic Perspective	190
6.2.2	Theoretical Perspective	191
6.3	Future Work	191
	Bibliography	193
	Appendix A Supplementary Materials in Chapter 5	271
	Summary	273
	Nederlandse samenvatting	275
	Curriculum Vitae	279
	Author Publication List	281

Contents

Acknowledgments	283
------------------------	------------