



Universiteit  
Leiden  
The Netherlands

## Learning from small samples

Kocaman, V.

### Citation

Kocaman, V. (2024, February 20). *Learning from small samples*. Retrieved from <https://hdl.handle.net/1887/3719613>

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/1887/3719613>

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

# Learning From Small Samples

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 20 februari 2024

klokke 13:45 uur

door

**Veysel Kocaman**

geboren te Adiyaman, Turkije

in 1981

**Promotor:**

Prof.dr. T.H.W. Bäck

**Co-promotor:**

Dr. O.M. Shir (Tel-Hai College, Israel)

**Promotiecommissie:**

Prof.dr. A. Laat

Prof.dr. J. Batenburg

Prof.dr. E. Hart (Edinburgh Napier University, UK)

Prof.dr. P. Kerschke (Technical University Dresden, Germany)

Dr. D.M. Pelt

Dr. N. van Stein

Copyright © 2024 Veysel Kocaman All Rights Reserved.

*This dissertation was made possible through the support of Leiden Institute of Advanced Computer Science (LIACS), Leiden University, which provided me with a part-time position, thereby facilitating my research and studies as an external PhD candidate. This arrangement was essential in enabling the completion of my doctoral studies without external grant funding.*

*No good can ever come from deviating from the path that you  
were destined to follow.  
—Robert Greene, *Mastery**

*With a heavy heart and profound respect, I dedicate this dissertation to my dear father, whose passing just a few months before the completion of my PhD journey has deeply affected me. The memory of his warm and reassuring smile is ever-present in my thoughts, providing comfort and inspiration in moments of reflection. His absence in the final stages of this academic pursuit is a profound sorrow. For six years, distance separated us, and I had always hoped he would be present to witness this milestone. His enduring love and the wisdom he shared continue to guide and inspire me. His smile, a symbol of joy and love, remains a vivid and cherished memory. As I reach this significant milestone, I do so with the image of his smiling face in my heart, hoping he is watching over me, feeling proud and fulfilled. His spirit has been a guiding light in the most challenging phases of this journey, making this accomplishment not just mine, but ours.*



# Contents

<b>1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Objectives	2
1.2.1 Research Questions	4
1.3 Outline of the Dissertation	4
1.4 Author's Contributions	5
1.5 Other Work by the Author	6
<b>2 Preliminaries</b>	<b>9</b>
2.1 Definition of Terms	9
<b>3 Dealing with Small Data in Machine Learning</b>	<b>15</b>
3.1 Introduction	16
3.2 Handling Small Data	19
3.2.1 Overfitting and Generalization	21

## Contents

---

<b>3.3 Approaches to tackle small data problems</b> . . . . .	26
<b>3.3.1 Data selection and preprocessing</b> . . . . .	27
<b>3.3.2 Incorporating domain, prior and context knowledge</b> . . . . .	28
<b>3.3.3 Picking the right approach</b> . . . . .	31
<b>3.3.4 Ensemble methods</b> . . . . .	32
<b>3.3.5 Transfer learning</b> . . . . .	33
<b>3.3.6 Parameter initialization</b> . . . . .	39
<b>3.3.7 Loss function reformulation</b> . . . . .	40
<b>3.3.8 Regularization techniques</b> . . . . .	41
<b>3.3.9 Data augmentation</b> . . . . .	43
<b>3.3.10 Synthetic data generation</b> . . . . .	46
<b>3.3.11 Problem reduction</b> . . . . .	48
<b>3.3.12 Optimization techniques</b> . . . . .	50
<b>3.3.13 Using physics-informed neural networks</b> . . . . .	54
<b>3.3.14 Unsupervised learning techniques</b> . . . . .	56
<b>3.3.15 Semi-supervised learning</b> . . . . .	56
<b>3.3.16 Self-supervised learning</b> . . . . .	60
<b>3.3.17 Zero-shot, one-shot and few-shot learning</b> . . . . .	61
<b>3.3.18 Meta learning</b> . . . . .	63
<b>3.3.19 Harnessing model uncertainty</b> . . . . .	64
<b>3.3.20 Active learning</b> . . . . .	66

3.3.21 Self-learning	71
3.3.22 Multi-task learning	71
3.3.23 Symbolic learning	73
3.3.24 Hierarchical learning	75
3.3.25 Knowledge distillation based learning	75
3.4 Dealing with imbalanced data	76
3.5 Anomaly Detection as a Small Data Problem	78
3.6 Conclusion	80
<b>4 The Role of Final Batch Normalization Layer</b>	<b>83</b>
4.1 Background	84
4.2 Related Work and Prior Art	87
4.3 Implementation Details and Experimental Results	89
4.3.1 Dataset Details	89
4.3.2 Adding a Final Batch Norm Layer Before the Output Layer	91
4.3.3 Experimentation Subject to Different Configurations	95
4.4 Discussion	101
4.4.1 Additional Experiments	106
4.5 Conclusions	106
<b>5 The Role of Self-Supervised Learning</b>	<b>111</b>
5.1 Self-Supervised Learning	111



## Contents

---

5.1.1 Motivation	111
5.1.2 Background	113
5.1.3 Contrastive Self-Supervised Learning	116
5.2 Preliminary: SSL applied to small subsets of Plant Village data	130
5.3 Salient Image Segmentation as a Surprising Player in SSL	132
5.3.1 Introduction: Salient Image Segmentation	132
5.3.2 Background	137
5.3.3 Related Work	142
5.3.4 Implementation, Setup & Results	144
5.3.5 Discussion	151
5.3.6 Conclusions	153
<b>6 Conclusions</b>	<b>163</b>
<b>7 Bibliography</b>	<b>169</b>
<b>English Summary</b>	<b>197</b>
<b>Nederlandse Samenvatting</b>	<b>199</b>
<b>Acknowledgements</b>	<b>201</b>
<b>About the Author</b>	<b>203</b>