University of Groningen

Deep learning and hyperspectral imaging for unmanned aerial vehicles

Dijkstra, Klaas

DOI:
[10.33612/diss.131754011](https://doi.org/10.33612/diss.131754011)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Dijkstra, K. (2020). *Deep learning and hyperspectral imaging for unmanned aerial vehicles: Combining convolutional neural networks with traditional computer vision paradigms*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.131754011>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

university of
 groningen

Deep Learning and Hyperspectral Imaging for Unmanned Aerial Vehicles

Combining convolutional neural networks with
traditional computer vision paradigms

PhD Thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of
the Rector Magnificus Prof. C. Wijmenga,
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on
Tuesday 29 September 2020 at 11:00 hours

by

Klaas Dijkstra
born on 13 March 1982
in Leeuwarden

Supervisor

Prof. L.R.B. Schomaker

Co-supervisor

Dr. M.A. Wiering

Assessment committee

Prof. T.P. Breckon

Prof. R.N.J. Veldhuis

Prof. L.V.E. Koopmans

"I have not failed once. I've succeeded in proving 700 ways how not to build a light bulb."

Thomas A. Edison

Contents

1	Introduction	1
1.1	Research questions	4
1.2	Dissertation overview	7
2	Hyperspectral frequency selection	13
2.1	Introduction	15
2.2	Materials and methods	15
2.2.1	Hyperspectral normalization and sample selection	15
2.2.2	Hyperspectral frequency selection	16
2.2.3	Classifying hyperspectral image patches	17
2.2.4	Cascading classifiers	18
2.3	Experiments and results	18
2.4	Discussion and conclusion	21
3	Hyperspectral demosaicking and crosstalk correction	23
3.1	Convolutional neural networks	28
3.1.1	A basic single-layer neural network	29
3.1.2	Training the layers of the CNN	30
3.2	Sensor geometry and datasets	31
3.2.1	Calibration data	32
3.3	Similarity maximization	33
3.3.1	Normalization	35
3.3.2	Mosaic to cube	36
3.3.3	Downsampling	37
3.3.4	Upscaling	38
3.3.5	Demosaicking	39
3.3.6	Loss function	40

3.3.7	Structural similarity	40
3.3.8	Crosstalk correction	43
3.4	Experiments	45
3.4.1	The effects of crosstalk correction	45
3.4.2	Demosaicking	46
3.4.3	End-to-end trainable neural network	49
3.5	Results	51
3.5.1	Crosstalk correction function	52
3.5.2	Quantitative analysis	55
3.5.3	Visual analysis	59
3.5.4	Spectral analysis	67
3.6	Discussion and conclusion	70
4	CentroidNet	75
4.1	Datasets	79
4.1.1	Crops	79
4.1.2	Kaggle data science bowl 2018	80
4.2	CentroidNet	82
4.3	Experiments	87
4.4	Results	90
4.4.1	Comparison with the state-of-the art on the crops dataset	90
4.4.2	Testing on larger images	93
4.5	Discussion and conclusion	95
5	CentroidNetV2	97
5.1	Related work	100
5.1.1	Deep design patterns	102
5.2	Contributions and research questions	104
5.3	The CentroidNetV2 architecture	105
5.3.1	Backbones	107
5.3.2	Loss functions	109
5.3.3	Coders	112
5.4	Datasets	119
5.4.1	Aerial crops	119
5.4.2	Cell nuclei	120

5.4.3	Bacterial colonies	121
5.4.4	Tiling	122
5.5	Training and validation	123
5.5.1	Training	123
5.5.2	Validation	124
5.6	Experiments and results	126
5.6.1	Results on aerial crops	127
5.6.2	Results on cell nuclei	131
5.6.3	Results on bacterial colonies	134
5.7	Discussion and conclusion	135
5.7.1	Future work	139
6	Discussion and conclusion	141
6.1	Research questions	141
6.2	Computer vision and deep learning	143
6.3	Future work	146

Summary

Samenvatting

Acknowledgements

Glossary

∂P Differentiable Programming. 148

BL Bilinear Interpolation. 56, 61

CCD Charge Coupled Device. 25

CE Cross Entropy. 127

CFA Color Filter Array. 25

CNN Convolutional Neural Network. 3, 4, 7, 9, 11, 23, 26, 28, 30, 33, 39, 67, 68, 77, 83, 84, 85, 86, 87, 97, 99, 100, 101, 102, 103, 138, 139, 142, 144, 145, 147, 148,

CPU Central Processing Unit. 18, 20

FCN Fully Convolutional Network. 75, 77, 78, 103

FRCNN Faster Recurrent Convolutional Neural Network. 77

GPS Global Positioning System. 8

GPU Graphical Processing Unit. 4, 17, 18, 20, 30, 79, 81

HSISR Hyperspectral Single Image Super Resolution. 73, 147

IoU Intersection over Union. 88, 89, 90, 92, 93, 111, 124, 125, 127,

kNN k-Nearest Neighbor. 17, 19,

LCFT Liquid Crystal Tunable Filter. 15, 25, 141,

LDA Linear Discriminant Analysis. 16, 19, 21,

mAP Mean Average Precision. 78

mAR Mean Average Recall. 78

MCFA Multispectral Color Filter Array. 9, 21, 23, 32, 46, 73, 142, 144, 147,

- MLP** Multi Layer Perceptron. 17, 18, 20, 21,
- MRCNN** Mask Recurrent Convolutional Neural Network. 97, 100, 104, 119, 120, 122, 124, 126, 127, 128, 131, 133, 134, 135, 137, 138,
- MSE** Mean Squared Error. 30, 40, 87, 88, 103, 104, 109, 110, 127, 128, 133, 138,
- NIR** Near Infrared. 39
- PCA** Principal Component Analysis. 16, 19,
- ReLU** Rectified Linear Unit. 11, 17, 21, 29, 43, 83, 86
- RGB** Red Green Blue. 4, 8, 17, 25, 31, 44, 54, 55, 67, 68, 69, 73, 84, 107, 147
- RNN** Recurrent Neural Network. 100
- SGD** Stochastic Gradient Descent. 18, 30, 44
- SISR** Single Image Super Resolution. 26, 27, 72, 73
- SSD** Single Shot Detector. 77
- SSIM** Structural Similarity. 23, 27, 33, 38, 40, 41, 42, 43, 45, 50, 51, 55, 56, 57, 58, 59, 61, 63, 65, 67, 71, 72,
- SVM** Support Vector Machine. 17, 19, 20,
- TanH** Hyperbolic Tangent. 17
- UAV** Unmanned Aerial Vehicle. 4, 5, 6, 7, 8, 9, 10, 15, 16, 21, 23, 25, 31, 60, 72, 73, 75, 78, 94, 95, 141, 142, 143, 146, 147,
- VL** Vector Loss. 127,
- YOLOv2** You Only Look Once Version 2. 75, 77, 78, 87, 90, 91, 92, 95,
- YOLOv3** You Only Look Once Version 3. 97, 100, 103, 104, 119, 120, 122, 124, 126, 127, 128, 134, 135, 137, 138,

Mathematical notation

v	is a scalar
$\mathbf{v}$	is a vector
$\mathbf{v}_{(x)}$	is a vector identified by x
$\mathbf{v}^\top$	is the transpose of vector $\mathbf{v}$
$\mathbf{M}$	is a matrix (2 dimensional) or a tensor (more than 2 dimensional)
$\mathcal{S}$	is a set
$\mathcal{P}_{(x)}$	a set identified by x
$\mathbf{M}^{n \times m}$	is an $n \times m$ matrix
$\mathbf{M}_{(x)}$	is a matrix identified by x
$\mathbf{T}_{y,x,c}$	is the y^{th}, x^{th}, c^{th} element of a matrix or tensor
$\mathbf{T}_{t:b,l:r,f:b}$	is a slice of a tensor indicated by the intervals $[t..b)$, $[l, r)$ and $[f, b)$
$\mathbf{T}_{\mathcal{A}, \mathcal{B}, \mathcal{C}}$	is a slice of a tensor indicated by the sets $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$
$\mathcal{F}_i^{n \times m \times l}$	a set with i tensors of size $n \times m \times l$
$\phi(x)$	is the sigmoid function ($\frac{1}{1+e^{-x}}$)
$\psi(x)$	is the ReLU function ($\max(0, x)$)
$\text{func}(\cdot)$	is a function
$\otimes_x$	is a convolution operator with stride x
$\oslash_x$	is a transposed convolution operator with stride x
op	is an operator
$\{a, b, c\}$	is a set of scalars a, b, c
$[\mathbf{a}, \mathbf{b}, \mathbf{c}]$	is a matrix composed of vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$
$[\mathbf{A}, \mathbf{B}, \mathbf{C}]$	is a tensor composed of matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$
$[\mathbf{A} \mathbf{B} \mathbf{C}]$	is a concatenation of matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$

Dedicated to my parents...