

# Cambridge Elements

Elements in Quantitative and Computational Methods  
for the Social Sciences

edited by  
R. Michael Alvarez  
*California Institute of Technology*  
Nathaniel Beck  
*New York University*

## IMAGES AS DATA FOR SOCIAL SCIENCE RESEARCH

An Introduction to Convolutional Neural Nets for  
Image Classification

Nora Webb Williams  
*University of Georgia*  
Andreu Casas  
*New York University*  
John D. Wilkerson  
*University of Washington*



CAMBRIDGE  
UNIVERSITY PRESS

Cambridge University Press  
978-1-108-81685-4 — Images as Data for Social Science Research  
Nora Webb Williams, Andreu Casas, John D. Wilkerson  
Frontmatter  
[More Information](#)

---

## CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of  
education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781108816854](http://www.cambridge.org/9781108816854)

DOI: 10.1017/9781108860741

© Nora Webb Williams, Andreu Casas, and John D. Wilkerson 2020

This publication is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without the written  
permission of Cambridge University Press.

First published 2020

*A catalogue record for this publication is available from the British Library.*

ISBN 978-1-108-81685-4 Paperback

ISSN 2398-4023 (online)

ISSN 2514-3794 (print)

Cambridge University Press has no responsibility for the persistence or accuracy of  
URLs for external or third-party internet websites referred to in this publication  
and does not guarantee that any content on such websites is, or will remain,  
accurate or appropriate.

Cambridge University Press  
978-1-108-81685-4 — Images as Data for Social Science Research  
Nora Webb Williams, Andreu Casas, John D. Wilkerson  
Frontmatter  
[More Information](#)

## Images as Data for Social Science Research

An Introduction to Convolutional Neural Nets for Image Classification

Elements in Quantitative and Computational Methods for the Social Sciences

DOI: 10.1017/9781108860741  
First published online: July 2020

Nora Webb Williams  
*University of Illinois Urbana-Champaign*  
Andreu Casas  
*Vrije Universiteit Amsterdam*  
John D. Wilkerson  
*University of Washington*

**Author for correspondence:** Nora Webb Williams, [nww3@illinois.edu](mailto:nww3@illinois.edu)

**Abstract:** Images play a crucial role in shaping and reflecting political life. Digitization has vastly increased the presence of such images in daily life, creating valuable new research opportunities for social scientists. We show how recent innovations in computer vision methods can substantially lower the costs of using images as data. We introduce readers to the deep learning algorithms commonly used for object recognition, facial recognition, and visual sentiment analysis. We then provide guidance and specific instructions for scholars interested in using these methods in their own research.

**Keywords:** political methodology, images as data, machine learning, convolutional neural nets

© Nora Webb Williams, Andreu Casas, and John D. Wilkerson 2020

ISBNs: 9781108816854 (PB), 9781108860741 (OC)  
ISSNs: 2398-4023 (online), 2514-3794 (print)

Contents

1	Introduction	1
2	Prerequisites for Computer Vision Methods and Tutorials	11
3	Introduction to CNNs for Social Scientists	14
4	Overview of Fine-Tuning a CNN Classifier for Images	24
5	Political Science Working Example: Images Related to a Black Lives Matter Protest	38
6	The Promise and Limits of Autotaggers	41
7	Application: Fine-Tuning an Open-Source CNN	50
8	Legal and Ethical Concerns in Using Images as Data	63
9	Conclusion	68
10	Data Availability Statement	69
	References	70