Synthesis Lectures on Computer Vision

SYNTHESIS
COLLECTION OF TECHNOLOGY

Terrance Boult
Walter Scheirer *Editors*

A Unifying Framework for Formal Theories of Novelty

Discussions, Guidelines, and Examples for Artificial Intelligence

Springer

Synthesis Lectures on Computer Vision

This series publishes on topics pertaining to computer vision and pattern recognition. The scope follows the purview of premier computer science conferences, and includes the science of scene reconstruction, event detection, video tracking, object recognition, 3D pose estimation, learning, indexing, motion estimation, and image restoration. As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner. As a technological discipline, computer vision seeks to apply its theories and models for the construction of computer vision systems, such as those in self-driving cars/navigation systems, medical image analysis, and industrial robots.

Terrance Boult · Walter Scheirer
Editors

A Unifying Framework for Formal Theories of Novelty

Discussions, Guidelines, and Examples for Artificial Intelligence

Springer

Editors
Terrance Boult
Department of Computer Science
University of Colorado at Colorado Springs
Colorado Springs, CO, USA

Walter Scheirer
Department of Computer Science
and Engineering
University of Notre Dame
Notre Dame, IN, USA

ISSN 2153-1056 ISSN 2153-1064 (electronic)
Synthesis Lectures on Computer Vision
ISBN 978-3-031-33053-7 ISBN 978-3-031-33054-4 (eBook)
https://doi.org/10.1007/978-3-031-33054-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface: The Novelty Problem in AI

AI researchers these days are quick to tout the progress that has been made in the field over the past decade. From game playing to visual recognition, new capabilities are appearing all of the time for many different applications. And indeed, such achievements should be celebrated. However, some very useful AI capabilities remain out of reach. For instance, why aren't safe self-driving cars available in the market in 2023? A major limitation of today's AI systems has become apparent in the quest for autonomous systems that must operate in real environments: they cannot manage novelty in the environment they were designed for. That is to say, if something new appears, there is no capacity for an agent to detect, characterize, and learn how to handle it. Given the practically infinite number of ways an environment can configure itself, coupled with the routine appearance of new things within an environment, novelty can be a significant confound. This book is the first attempt to study novelty problems in a rigorous fashion through the use of a unifying framework for formal theories of novelty.

Ad hoc ways of addressing the novelty problem have proven to be insufficient. There exists a persistent belief that reinforcement learning is all that is needed for agents to manage novelty because any novelty can be learned over time. Similarly, there exists blind faith in the generalization properties of deep learning through invariant representation alone. Given the results found in this book for the simplest of AI domains, we can safely say that a more principled approach to novelty management is needed. Neither approach addresses the core detection problem at the classifier level, nor is there any capacity to characterize novelty, which can take on many different forms. On that latter point, what exactly does it mean for something to be novel? That isn't a question that can be answered using an off-the-shelf AI algorithm. The need for a theory matched to a specific domain can provide a better starting point for agent design.

A recent effort to address the novelty problem in AI has been the DARPA Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON) program. It established a research program to develop a set of engineering design principles for open-world learning in 2019 [1]. Throughout the four years of the program, a large consortium of academic, industry, and government researchers has collaborated on fundamental work looking into innovative strategies for effective open-world learning in both activity domains, e.g., interactive video games, and perceptual domains, e.g., datasets of images and videos. This book is the output of the program's Novelty Working Group, which was charged with developing viable theories for the study of novelty in AI. Each chapter was contributed by different participants in that working group.

This book is organized in the following manner. Chapter 1 is the focal point of the book. It introduces a unifying framework for creating theories of novelty that are matched to specific domains. This includes definitions on different types of novelty, as well as constructs for building agents that can detect, characterize, and manage novelty. This framework is general and can apply to *any* domain in which novelty appears. To justify this claim, each subsequent chapter provides an example domain for which a theory is developed and evaluated. These chapters include: (1) a task overview, (2) definitions of dissimilarity and regret operators, (3) definitions of measurements and observations, (4) a description of novelty types and examples, (5) a set of experiments validating predictions made by the developed theory, and (6) concluding remarks.

The domain-specific chapters cover a broad range of activity and perceptual domains. Chapter 2 starts things off as simple as possible with a study of novelty in the 2D CartPole activity domain. Chapter 3 extends the study of CartPole by examining a 3D version of the environment. Chapter 4 turns to the perceptual domain of image classification in computer vision. Chapter 5 discusses a related computer vision domain, handwriting recognition, which also contains elements of natural language processing. Chapter 6 pushes farther into the realm of natural language processing by studying contextual and semantic novelty in text. Chapter 7 comes back to activity domains with an examination of the game Monopoly. The book concludes in Chap. 8 by recapping what we have learned and suggesting the development of new theoretical directions that are interdisciplinary in nature.

Colorado Springs, USA Terrance Boult
Notre Dame, USA Walter Scheirer

Acknowledgments This research was sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Army Research Office (ARO) under multiple contracts/agreements including HR001120C0055, W911NF-20-2-0005, W911NF-20-2-0004, HQ0034-19-D-0001, and

W911NF2020009. The views contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of DARPA or ARO, or the US Government.

Reference

1. DARPA. Teaching AI systems to adapt to dynamic environments, (2019)

Contents

List of Figures