

Multimodal Machine Learning: A Survey and Taxonomy

Publisher: IEEE

Cite This

PDF

Tadas Baltrušaitis  ; Chaitanya Ahuja  ; Louis-Philippe Morency 

All Authors

1854

Cites in Papers

59333

Full Text Views



Need Full-Text

access to IEEE Xplore for your organization?

CONTACT IEEE TO SUBSCRIBE >

Abstract

Document Sections

1 Introduction

2 Applications: A Historical Perspective

3 Multimodal Representations

4 Translation

5 Alignment

Show Full Outline ▾

Authors

Figures

References

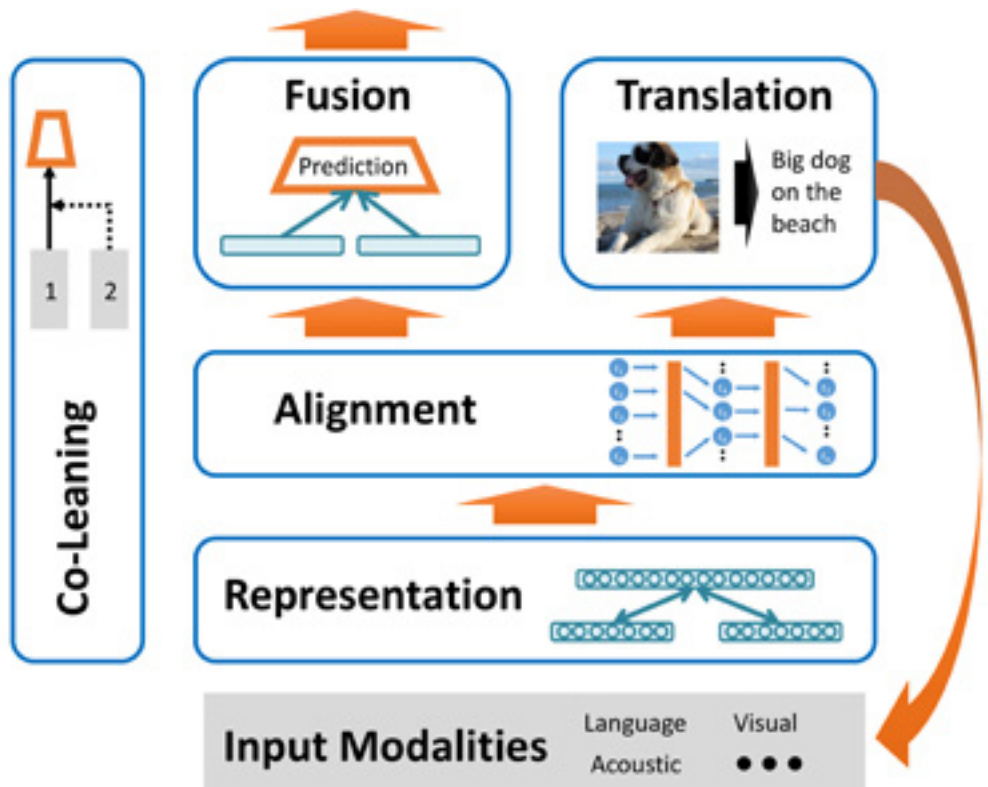
Citations

Keywords

Metrics

Abstract:

Our experience of the world is multimodal - we see objects, hear sounds, feel texture, smell odors, and taste flavors. Modality refers to the way in which something happens or is experienced and a research problem is characterized as multimodal when it includes multiple such modalities. In order for Artificial Intelligence to make progress in understanding the world around us, it needs to be able to interpret such multimodal signals together. Multimodal machine learning aims to build models that can process and relate information from multiple modalities. It is a vibrant multi-disciplinary field of increasing importance and with extraordinary potential. Instead of focusing on specific multimodal applications, this paper surveys the recent advances in multimodal machine learning itself and presents them in a common taxonomy. We go beyond the typical early and late fusion categorization and identify broader challenges that are faced by multimodal machine learning, namely: representation, translation, alignment, fusion, and co-learning. This new taxonomy will enable researchers to better understand the state of the field and identify directions for future research.



Published in: IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 41 , Issue: 2, 01 February 2019)

Page(s): 423 - 443

DOI: 10.1109/TPAMI.2018.2798607

Date of Publication: 25 January 2018 ?

Publisher: IEEE

▾ ISSN Information:

PubMed ID: 29994351

More Like This

Type-2 fuzzy hidden Markov models and their application to speech recognition

IEEE Transactions on Fuzzy Systems
Published: 2006

EMG-Based Speech Recognition Using Hidden Markov Models With Global Control Variables

IEEE Transactions on Biomedical Engineering
Published: 2008

Show More

Sign in to Continue Reading

Authors	▾
Figures	▾
References	▾
Citations	▾
Keywords	▾
Metrics	▾

IEEE Personal Account

CHANGE USERNAME/PASSWORD

Purchase Details

PAYMENT OPTIONS

VIEW PURCHASED DOCUMENTS

Profile Information

COMMUNICATIONS PREFERENCES

PROFESSION AND EDUCATION

TECHNICAL INTERESTS


Need Help?

US & CANADA: +1 800 678 4333


WORLDWIDE: +1 732 981 0060

CONTACT & SUPPORT

Follow

About IEEE Xplore | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | IEEE Ethics Reporting  | Sitemap | IEEE Privacy Policy

A public charity, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

 © Copyright 2025 IEEE - All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies.

Feedback