Scopus

Document details

→] Export　↓ Download　🖶 Print　✉ E-mail　📄 Save to PDF　☆ Add to List　More... ›

| Full Text |　View at Publisher

Viewpoint invariant semantic object and scene categorization with RGB-D sensors　(Article)

Zaki, H.F.M.[a] ✉,　Shafait, F.[b] ✉,　Mian, A.[c] ✉　👤

[a]Department of Mechatronics Engineering, International Islamic University Malaysia, Kuala Lumpur, 53100, Malaysia
[b]National University of Sciences and Technology, Islamabad, Pakistan
[c]School of Computer Science and Software Engineering, The University of Western Australia, Crawley, WA 6009, Australia

Abstract　⌄ View references (53)

Understanding the semantics of objects and scenes using multi-modal RGB-D sensors serves many robotics applications. Key challenges for accurate RGB-D image recognition are the scarcity of training data, variations due to viewpoint changes and the heterogeneous nature of the data. We address these problems and propose a generic deep learning framework based on a pre-trained convolutional neural network, as a feature extractor for both the colour and depth channels. We propose a rich multi-scale feature representation, referred to as convolutional hypercube pyramid (HP-CNN), that is able to encode discriminative information from the convolutional tensors at different levels of detail. We also present a technique to fuse the proposed HP-CNN with the activations of fully connected neurons based on an extreme learning machine classifier in a late fusion scheme which leads to a highly discriminative and compact representation. To further improve performance, we devise HP-CNN-T which is a view-invariant descriptor extracted from a multi-view 3D object pose (M3DOP) model. M3DOP is learned from over 140,000 RGB-D images that are synthetically generated by rendering CAD models from different viewpoints. Extensive evaluations on four RGB-D object and scene recognition datasets demonstrate that our HP-CNN and HP-CNN-T consistently outperforms state-of-the-art methods for several recognition tasks by a significant margin. © 2018, Springer Science+Business Media, LLC, part of Springer Nature.

SciVal Topic Prominence ⓘ

Topic:　Object recognition | Convolution | Convolutional sparse

Prominence percentile:　86.795　　　　ⓘ

Author keywords

(Multi-modal deep learning)　(Object　categorization)　(RGB-D image)　(Scene　recognition)

Indexed keywords

Engineering controlled terms:　(Computer aided design)　(Convolution)　(Image recognition)　(Neural networks)
(Object　recognition)　(Semantics)

Engineering uncontrolled terms　(Compact representation)　(Convolutional neural network)　(Extreme learning machine)
(Multi-modal)　(Object　categorization)　(RGB-D image)　(Scene　recognition)
(State-of-the-art methods)

Metrics　ⓘ

0　Citations in Scopus

0　Field-Weighted Citation Impact

❋

PlumX Metrics　⌄
Usage, Captures, Mentions, Social Media and Citations beyond Scopus.

Cited by 0 documents

Inform me when this document is cited in Scopus:

Set citation alert ›

Set citation feed ›

Related documents

Convolutional hypercube pyramid for accurate RGB-D object category and instance recognition

Zaki, H.F.M. , Shafait, F. , Mian, A.
(2016) Proceedings - IEEE International Conference on Robotics and Automation

Learning a deeply supervised multi-modal RGB-D embedding for semantic scene and object category recognition

Zaki, H.F.M. , Shafait, F. , Mian, A.
(2017) Robotics and Autonomous Systems

Localized Deep Extreme Learning Machines for Efficient RGB-D Object Recognition

Zaki, H.F.M. , Shafait, F. , Mian, A.
(2015) 2015 International Conference on Digital Image Computing: Techniques and Applications, DICTA 2015

References (53)

View in search results format ›

☐ All   Export   🖶 Print   ✉ E-mail   📄 Save to PDF   Create bibliography

☐ 1   Angeli, A., Filliat, D., Doncieux, S., Meyer, J.-A.

Fast and incremental method for loop-closure detection using bags of visual words

(2008) *IEEE Transactions on Robotics*, 24 (5), pp. 1027-1037. Cited 256 times.
doi: 10.1109/TRO.2008.2004514

View at Publisher

☐ 2   Asif, U., Bennamoun, M., Sohel, F.

Discriminative feature learning for efficient RGB-D object recognition

(2015) *IEEE International Conference on Intelligent Robots and Systems*, 2015-December, art. no. 7353385, pp. 272-279. Cited 8 times.
ISBN: 978-147999994-1
doi: 10.1109/IROS.2015.7353385

View at Publisher

☐ 3   Asif, U., Bennamoun, M., Sohel, F.

Efficient RGB-D object categorization using cascaded ensembles of randomized decision trees

(2015) *Proceedings - IEEE International Conference on Robotics and Automation*, 2015-June (June), art. no. 7139358, pp. 1295-1302. Cited 21 times.
doi: 10.1109/ICRA.2015.7139358

View at Publisher