



Universiteit  
Leiden

The Netherlands

## Multi modal representation learning and cross-modal semantic matching

Wang, X.

### Citation

Wang, X. (2022, June 24). *Multi modal representation learning and cross-modal semantic matching*. Retrieved from <https://hdl.handle.net/1887/3391031>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3391031>

**Note:** To cite this publication please use the final published version (if applicable).

## Stellingen behorende bij het proefschrift

### Multi Modal Representation Learning and Cross-Modal Semantic Matching

Xue Wang

1. The use of multimodal data drives the merging of vision and language in the learning of the representation. The synergy of supervision by vision reduces the bias in the textual representation embeddings. This thesis Chapter 2
2. Object instances marked with a coarse label can benefit from a fine-grained semantic map. Using such map will extend the object categories and at the same time it will support the learning of more complex object detectors. This thesis Chapter 3
3. The success of the process of curriculum learning is in the analysis of the distribution of data while ranking the data by complexity. This thesis Chapter 3
4. The cross-modal semantic space addresses the complexity distribution of cross-modal data. The kernel non-linear transformation model can measure the similarity between the features of different modalities. This thesis Chapter 4
5. An improvement of the efficiency and accuracy of phrase grounding for cross-modal semantic matching will enhance the modelling of complex interactions between visual components as well as between visual and textual components. This thesis Chapter 5
6. The contrastive-based weakly-supervised referring expression grounding model localizes the target object by the extracting discriminative attribute triple from language expression. This thesis Chapter 5
7. The aim of AI is to mimic the human way to learn and extract information for the purpose of advancing automated systems. AI could efficiently fuse information from different modalities.
8. Each problem embodies a possibility in finding a challenging and meaningful solution to existing or new cases.
9. Always take a moment to admire all the beautiful things you take for granted.
10. The route to a PhD is full of obstacles and unforeseen circumstances. When in doubt, you should realize that at the end of the trajectory, in retrospect, you can appreciate what you have accomplished and learned.