

## Accepted Manuscript

Learning Visual and Textual Representations for Multimodal Matching and Classification

Yu Liu, Li Liu, Yanming Guo, Michael S. Lew

PII: S0031-3203(18)30233-4  
DOI: [10.1016/j.patcog.2018.07.001](https://doi.org/10.1016/j.patcog.2018.07.001)  
Reference: PR 6593



To appear in: *Pattern Recognition*

Received date: 29 August 2017  
Revised date: 22 May 2018  
Accepted date: 1 July 2018

Please cite this article as: Yu Liu, Li Liu, Yanming Guo, Michael S. Lew, Learning Visual and Textual Representations for Multimodal Matching and Classification, *Pattern Recognition* (2018), doi: [10.1016/j.patcog.2018.07.001](https://doi.org/10.1016/j.patcog.2018.07.001)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Highlights**

- A unified network for image-text matching and classification;
- Seamlessly incorporating the matching and classification components;
- A multi-stage training algorithm by combining the matching and classification loss;
- Comprehensive study on the effectiveness of the proposed approach;
- Comparisons on four well-known multimodal benchmarks;

ACCEPTED MANUSCRIPT

Download English Version:

<https://daneshyari.com/en/article/6938679>

Download Persian Version:

<https://daneshyari.com/article/6938679>

[Daneshyari.com](https://daneshyari.com)