6th Workshop on Vision and Language (VL 2017)

Valencia, Spain
4 April 2017

ISBN: 978-1-5108-3873-4
# Table of Contents

*The BURCHAK corpus: a Challenge Data Set for Interactive Learning of Visually Grounded Word Meanings*
  Yanchao Yu, Arash Eshghi, Gregory Mills and Oliver Lemon ........................................ 1

*The Use of Object Labels and Spatial Prepositions as Keywords in a Web-Retrieval-Based Image Caption Generation System*
  Brandon Birmingham and Adrian Muscat ................................................................. 11

*Learning to Recognize Animals by Watching Documentaries: Using Subtitles as Weak Supervision*
  Aparna Nurani Venkitasubramanian, Tinne Tuytelaars and Marie-Francine Moens .......... 21

*Human Evaluation of Multi-modal Neural Machine Translation: A Case-Study on E-Commerce Listing Titles*
  Iacer Calixto, Daniel Stein, Evgeny Matusov, Sheila Castilho and Andy Way ............... 31

*The BreakingNews Dataset*
  Arnau Ramisa, Fei Yan, Francesc Moreno-Noguer and Krystian Mikolajczyk ................. 38

*Automatic identification of head movements in video-recorded conversations: can words help?*
  Patrizia Paggio, Costanza Navarretta and Bart Jongejan ........................................... 40

*Multi-Modal Fashion Product Retrieval*
  Antonio Rubio Romano, LongLong Yu, Edgar Simo-Serra and Francesc Moreno-Noguer ...... 43