









## REFERENCES

- [1] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Trans. Assoc. Comput. Linguist.* 5 (6 2017), 135–146.
- [2] Michael J Cortese and April Fugett. 2004. Imageability ratings for 3,000 monosyllabic words. *Beh. Res. Methods Instrum. Comput.* 36, 3 (8 2004), 384–387.
- [3] Marc A. Kastner, Ichiro Ide, Frank Nack, Yasutomo Kawanishi, Takatsugu Hirayama, Daisuke Deguchi, and Hiroshi Murase. 2020. Estimating the imageability of words by mining visual characteristics from crawled image data. *Multimed. Tools Appl.* (2 2020). <https://doi.org/10.1007/s11042-019-08571-4>
- [4] Nikola Ljubešić, Darja Fišer, and Anita Peti-Stantić. 2018. Predicting concreteness and imageability of words within and across languages via word embeddings. In *Proc. 56th Annu. Meet. Assoc. Comput. Linguist.* 217–222.
- [5] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *Comput. Res. Repos.*, arXiv: 1301.3781.
- [6] Tomas Mikolov, Edouard Grave, Piotr Bojanowski, Christian Puhersch, and Armand Joulin. 2018. Advances in Pre-Training Distributed Word Representations. In *Proc. Int. Conf. on Language Resources and Evaluation 2018.* 52–55.
- [7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proc. 26th Int. Conf. Neural Inf. Process. Syst.*, Vol. 2. 3111–3119.
- [8] Allan Paivio. 1990. *Mental representations: A dual coding approach.* Vol. 9. Oxford University Press, Oxford, England.
- [9] Allan Paivio, John C. Yuille, and Stephen A. Madigan. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *J. Exp. Psychol.* 76, 1 (1 1968), 1–25.
- [10] Joseph Redmon and Ali Farhadi. 2017. YOLO9000: Better, faster, stronger. In *Proc. 2017 IEEE Conf. Comput. Vis. Pattern Recognit.* 6517–6525.
- [11] Jamie Reilly and Jacob Kean. 2007. Formal distinctiveness of high- and low-imageability nouns: Analyses and theoretical implications. *J. Cogn. Sci.* 31, 1 (2 2007), 157–168.
- [12] Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: The new data in multimedia research. *Comm. ACM* 59, 2 (2 2016), 64–73.
- [13] Mingda Zhang, Rebecca Hwa, and Adriana Kovashka. 2018. Equal but not the same: Understanding the implicit relationship between persuasive images and text. In *Proc. 2018 Br. Mach. Vis. Conf.* 14p.