

A program that captions your photos

November 16 2015

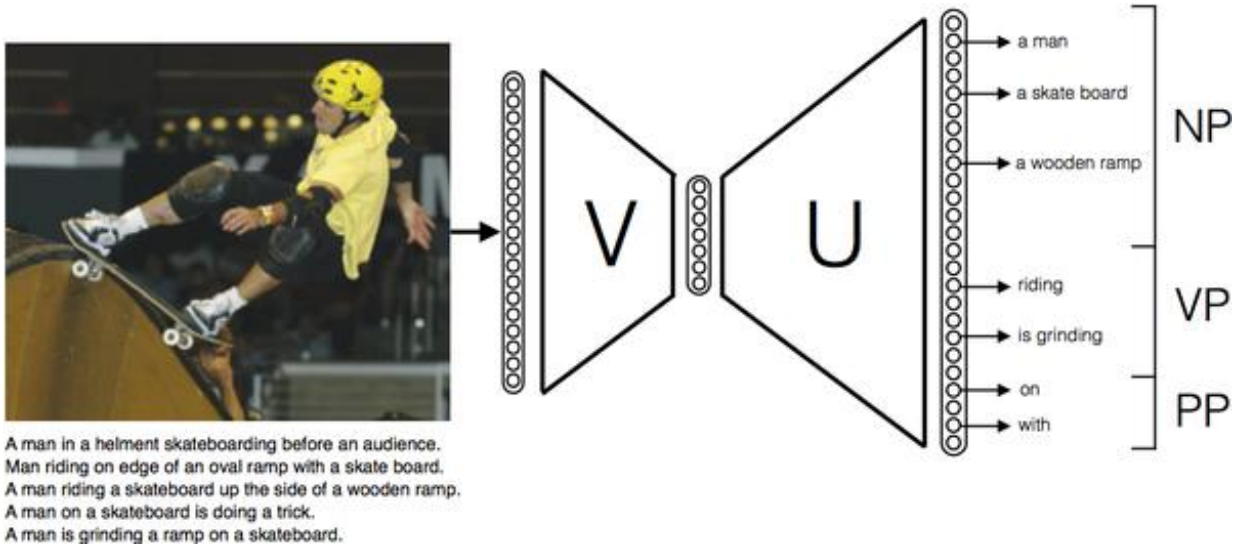


Two researchers at Idiap, a research institute in Martigny that is affiliated with EPFL, developed an algorithm that – unlike systems recently unveiled by Google and Microsoft – can describe an image without having to pull up captions that it has already learned. To do this, the researchers used a program capable of making vector representations of images and captions based on an analysis of caption syntax.

"When we give it a photo, the program compares the image vector to the vector of possible words and selects the most likely noun, verb and prepositional phrases," said Rémi Lebret, a PhD student specializing in Deep Learning at Idiap. This is how the system finds the most likely description for a photo of a man skateboarding, for example, even if it has never seen a similar photo previously. The computer breaks down the picture into elements ("a skateboard, a man, a ramp") and verbs that could describe the action ("riding") before captioning the picture.

Getting it right

This approach is unlike existing ones. "Those other systems propose the first word based on the photo and then use that word to predict subsequent ones," said Pedro Oliveira Pinheiro, the other Idiap researcher on this project. Those algorithm based on sequence labeling with recurrent neural networks can cause problems, however, because if it poorly predicts the start of the phrase, the entire caption will necessary be wrong. Those systems also have a longer learning curve, and they tend to recycle previously used captions.



The technology developed by Pinheiro and Lebet is simpler and works better. And it has piqued the interest of social media. The two researchers did a six-month research internship at Facebook, which is drawing on their work to develop its own model of automatic captions meant in part for the visually impaired. The two researchers believe that their algorithm could be improved in the future through the use of more complex language models and by linking it to larger databases.

More information: Phrase-based Image Captioning.
arxiv.org/abs/1502.03671

Provided by Ecole Polytechnique Federale de Lausanne

Citation: A program that captions your photos (2015, November 16) retrieved 26 April 2024 from <https://phys.org/news/2015-11-captions-photos.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.