A Novel Actor Dual-Critic Model for Remote Sensing Image Captioning

Ruchika Chavhan 1, Biplab Banerjee 2, Xiao Xiang Zhu 3, Subhasis Chaudhuri 1

1Indian Institute of Technology, Bombay, India 2Signal Processing in Earth Observation, Technical University of Munich, Germany

Abstract

We introduce a novel actor dual-critic (ADC) method to generate captions for remote sensing images. The ADC method consists of an encoder-decoder LSTM and an additional critic network that is trained to predict the advantage factor for the critic. We demonstrate the efficiency of our method on two datasets: RSICD and UCMS-captions and perform interesting experiments to validate the functionality of the critic.

The Actor-Dual-Critic Setup

We introduce an additional critic to the A2C setup in the form of an Encoder-Decoder LSTM that jointly encodes the sentences and images and encourages prediction of semantically more precise sentences. The critic provides a significant boost in performance in the following ways:

• Sentence to image translation: This critic translates back into images to generate a novelty which closely resembles the features extracted from images, thus validating the contextual accuracy of the predicted captions.

• More exploration of environment: The policy successfully investigates the environment consisting of images and captions and gains more knowledge due to this critic’s extra upgrade step in the optimization of policy objective.

We demonstrate the efficiency of our method on two datasets: RSICD and UCMS-captions and perform interesting experiments to validate the functionality of the critic.

Results

An experiment demonstrating the validity of the critic

To validate if the critic alleviates high inter class similarity, we pass a different image from the same class with identical reference sentence as the test input.

References
