Readme

Introduction:

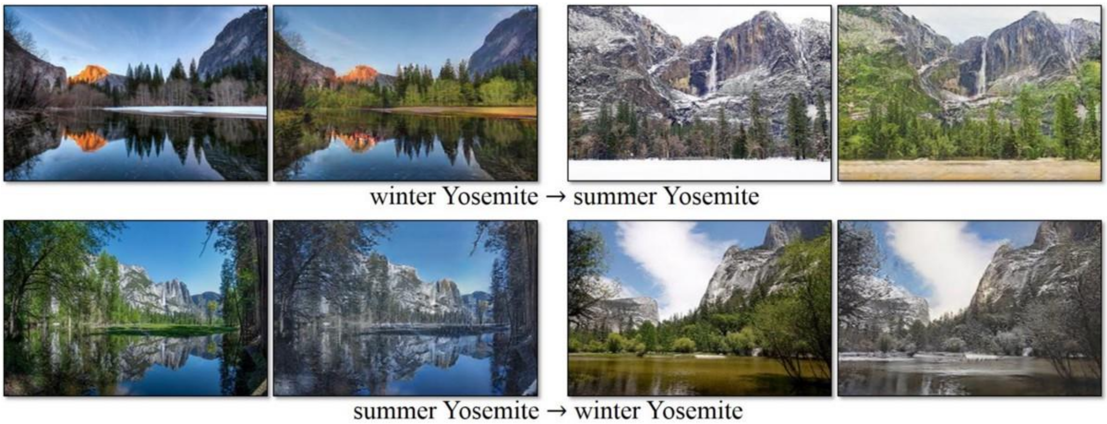
In Q3, we will introduce some methods to do data augmentation, to be more specific: data augmentation for image classification and data augmentation for text classification.

Methods:

1. Data Augmentation on Images

- **Conditional GAN**

Without going into gory detail, conditional GANs can transform an image from one domain to an image to another domain. This is a powerful neuron network for data augmentation. Below is an example of conditional GANs used to transform photographs of summer sceneries to winter sceneries.



The above method is robust, but computationally intensive. A cheaper alternative would be something called neural style transfer. It grabs the texture, ambiance, appearance of one image (aka, the "style") and mixes it with the content of another. Using this powerful technique, we produce an effect which is similar to that of our conditional GAN.

- **Traditional affine transformation**

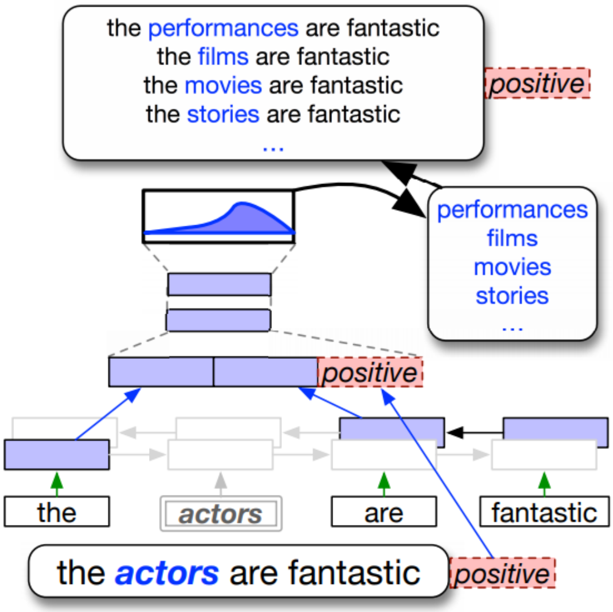
✧ Spatial geometry transformation:

✧ Pixel color conversion:

2. Data Augmentation for text

➢ Paradigmatic Relation

If we assume an invariance that sentences are natural even if the words in the sentences are replaced with other words with paradigmatic relations. We stochastically replace words with other words that are predicted by a bi-directional language model at the word positions. Words predicted according to a context are numerous but appropriate for the augmentation of the original words. [Sosuke,2018]



(Contextual augmentation with a bidirectional RNN language model, when a sentence "the actors are fantastic" is augmented by replacing only actors with words predicted based on the context.)

3. Data Augmentation on Audio

- Background noise

This perturbation adds noise to the input audio signal at a required signal to noise ratio (SNR). Currently, only white-noise is present but the code can easily be modified to add more types of noise. SNR is the ratio of root mean square(RMS) amplitude of the signal to RMS amplitude of noise, squared, which formed the basis of the implementation.

- Time stretching

This augmentation changes the time duration of the input audio signal. Depending upon the requirement, it can stretch or compress without changing the pitch of the audio. This is also implemented using Rubberband ([rub, 2012]), and ratios to be inputted can vary between -9.99 to +9.99.

- Equalization

It applies an equalization to the input audio signal given a center frequency, bandwidth and gain. With this transformation, the signal-level at and around a selected frequency can be increased or decreased, whilst (unlike band-pass and band-reject filters) that at all other frequencies is unchanged. This is implemented using the equalizer function of sox ([sox, 2013]).

References：

[1] Dongmei Han, Qigang Liu, Weiguo Fan: A new image classification method using CNN transfer learning and web data augmentation. Expert Syst. Appl. 95: 43-56 (2018)

[2] Odongo Steven Eyobu, Dong Seog Han: Feature Representation and Data Augmentation for Human Activity Classification Based on Wearable IMU Sensor Data Using a Deep LSTM Neural Network. Sensors 18(9): 2892 (2018)

[3] Yalong Bai, Kuiyuan Yang, Tao Mei, Wei-Ying Ma, Tiejun Zhao: Automatic Data Augmentation from Massive Web Images for Deep Visual Recognition. TOMCCAP 14(3): 69:1-69:20 (2018)

[4] Jose Carlos Villarreal Guerra, Zeba Khanam, Shoaib Ehsan, Rustam Stolkin, Klaus D. McDonald-Maier: Weather Classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of Convolutional Neural Networks. AHS 2018: 305-310

[5] Yezi Zhu, Marc Aoun, Marcel Krijn, Joaquin Vanschoren: Data Augmentation using Conditional Generative Adversarial Networks for Leaf Counting in Arabidopsis Plants. BMVC 2018: 324

[6] Hu Hu, Tian Tan, Yanmin Qian: Generative Adversarial Networks Based Data Augmentation for Noise Robust Speech Recognition. ICASSP 2018: 5044-5048

[7] Lei Xia, Jiancheng Lv, Yong Xu: A Data Augmentation Model Based on Variational Approach. ICONIP (2) 2018: 157-168

[8] Yutai Hou, Yijia Liu, Wanxiang Che, Ting Liu: Sequence-to-Sequence Data Augmentation for Dialogue Language Understanding. CoRR abs/1807.01554 (2018)

[9] Ghouthi Boukli Hacene, Vincent Gripon, Nicolas Farrugia, Matthieu Arzel, Michel Jézéquel: Transfer Incremental Learning using Data Augmentation. CoRR abs/1810.02020 (2018)

[10] Christopher Bowles, Roger N. Gunn, Alexander Hammers, Daniel Rueckert: GANsfer Learning: Combining labelled and unlabelled data for GAN based data augmentation. CoRR abs/1811.10669 (2018)