

Data Augmentation and Neural Networks to Improve Image Classification

Prepared for: DATS6203 Machine Learning II

Prepared by: Group 3, Joseph Francis, Srilatha Lakka

Introduction:

In machine learning, image classification is a process to analyze the extracted image features and organize them into categories by using neural networks. In recent years, the neural network techniques have improved image classification accuracy quickly, such as AlexNet can achieve 15.3% image classification error rates¹. These techniques are practically applied in many fields, such as artificial intelligence, medical diagnostics, E-commerce, gaming, or automotive industries. The convolutional neural network (CNN) is one of the most popular deep neural network (DNN) learning algorithms which can perform classification tasks directly from images. CNN models can produce the state-of-the-art classification predictions with directly learned features. A CNN model generally contains more than one convolutional layer for specialized linear operations and includes local or global pooling layers to perform nonlinear down sampling. After learning features from many layers, a fully connected layer outputs the probabilities for each class to be predicted. However, model overfitting and poor performance are common problems in applying neural network techniques.

Problem Statement:

Model overfitting and poor performance are common problems in applying neural network techniques because some of the high frequency features may not be useful in classification. Approaches to bring intra-class differences down and retain sensitivity to the inter-class variations are important to maximize model accuracy and minimize the loss function.

Project Description:

Data augmentation is a common technique to overcome the lack of large, annotated databases, a usual situation when applying deep learning to medical imaging problems. Nevertheless, there is no consensus on which transformations to apply for a particular field. This work aims at identifying the effect of different transformations on polyp segmentation using deep learning.

Research Overview:

Data augmentation for images consists of increasing the amount and diversity of training cases based on the available images in the database through the application of image transformations such as translation or flipping of the original image [1]. Different computational libraries have been created to perform these transformation functions [2, 3]. However, the selection of the most suitable strategy remains a trial-and-error process that depends on the experience, imagination and time of the researcher [4]. There are several studies analyzing the effect of data augmentation for image classification tasks [5,6,7,8], but this field is not fully explored for semantic segmentation yet.

Related Material and Background Supportive:

Data augmentation techniques are often used to regularize models which work with images in neural networks and other learning algorithms. With the labelled original training dataset, synthetic images can be created by various transformations to the original images. Keras ImageDataGenerators¹¹ is the tool used for generating more training data from the original data to avoid model overfitting. It is conducted online by looping over in small batches during each optimizer iteration. There are some graphic parameters (e.g. rotation, shift, flip, add Gaussian noises) to help generate artificial images.

There are various data augmentation techniques: (1) flipping images horizontally or vertically; (2) rotating images at some degrees; (3) rescaling outward or inward; (4) randomly cropping; (5) translating by width and height shifts; (6) whitening, (7) shearing, (8) zooming and (9) adding 10% Gaussian noises to prevent model over-fitting and enhance learning capability.

Dataset and Features:

The dataset contains 4323 images of flowers from Kaggle. The data collection is based on the data Flickr, google images, yandex images. It includes 5 types of flowers (5 classes or 5 labels) for this classification problem.

The pictures are divided into five classes: chamomile, tulip, rose, sunflower, dandelion.(image 1). Photos are not high resolution, about 320x240 pixels. Photos are not reduced to a single size; they have different proportions and resolution.(below image 2 and 3)

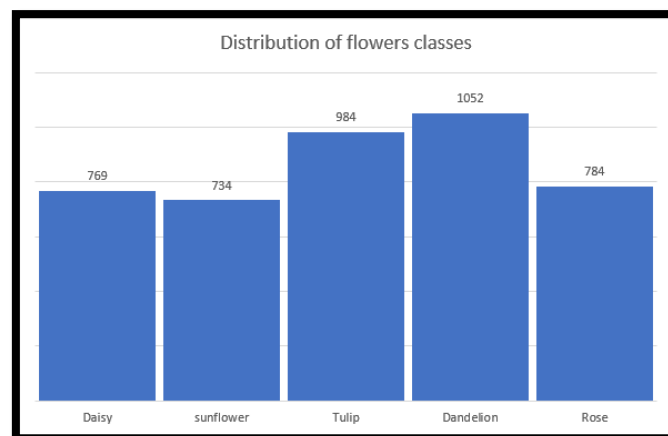


Image1: Distribution of different classes



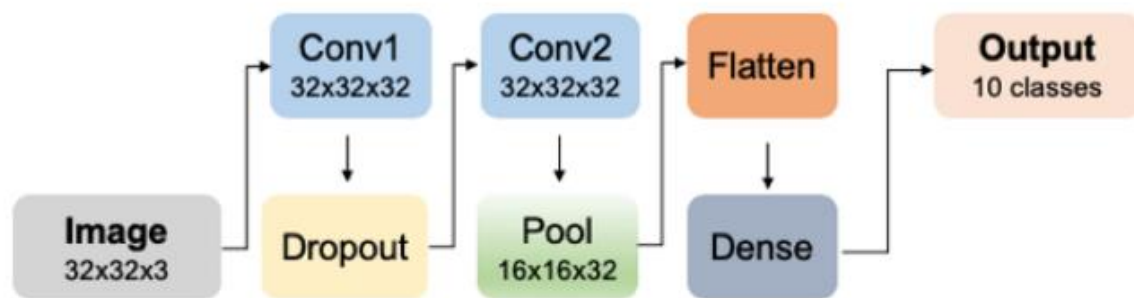
Image 2



Image 3

Initial Analysis and Procedure for Project:

We start with a basic CNN model with two convolutional layers, one max-pooling layer, and one classifier layer for image classification. Secondly, the standard CNN models are built with six convolutional layers and one dense layer (a VGG16-like model). Before each convolutional layer, a batch normalization layer is applied. The dropout layers are also added between convolutional layers. The final fully connected Softmax layer produces a probability distribution over ten predicted output classes. Thirdly, VGG16 models were applied with optimized architectures, hyper-parameter tuning, and data augmentation techniques.



Basic CNN model M0 architecture

The model performance will be evaluated with accuracy and loss function for the training, validation and test datasets in the absence or presence of applying multiple data augmentation techniques.

Project Schedule:

March 15th – March 29th - POC on Datasets

March 30th – April 13th - Deep-Learning network selection and network design section

April 14th – April 20th - Evaluation section and Improvement

April 21st – April 27th - Summary and Presentation section

References:

1. Cubuk ED, Zoph B, Mane D, Vasudevan V, Le QV (2018) AutoAugment: learning augmentation policies from data. arXiv. <https://arxiv.org/abs/1805.09501>. Accessed 5 Jan 2020
2. Jung AB, Wada K, Crall J, Tanaka S, Graving J, Yadav S, Banerjee J, Vecsei G, Kraft A, Borovec J, Vallentin C, Zhydenko S, Pfeiffer K, Cook B, Fernández I, Chi-Hung W, Ayala-Acevedo A, Meudec R, Laporte M (2019) ImgAug. <https://github.com/aleju/imgaug>. Accessed 5 Jan 2020
3. Casado-García Á, Domínguez C, García-Domínguez M, Heras J, Mata E, Pascual V (2019) CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. BMC Bioinform 20:1–14. <https://doi.org/10.1186/s12859-019-2931-1>
4. Okafor E, Schomaker L, Wiering MA (2018) An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals augmentation in classifying images of animals. J Inf Telecommun. <https://doi.org/10.1080/24751839.2018.1479932>
5. Shijie J, Ping W, Peiyi J, Siping H (2017) Research on data augmentation for image classification based on convolution neural networks. In: 2017 Chinese automation congress (CAC). pp 4165–4170
6. Wang J, Perez L (2017) The effectiveness of data augmentation in image classification using deep learning. <https://arxiv.org/abs/1712.04621>
7. Zhang C, Tavanapong W, Wong J, de Groen PC, Oh JH (2017) Real data augmentation for medical image classification. In: Intravascular imaging and computer assisted stenting, and large-scale annotation of biomedical data and expert label synthesis. pp 67–76. https://doi.org/https://doi.org/10.1007/978-3-319-67534-3_8
8. Ma R, Tao P, Tang H (2019) Optimizing data augmentation for semantic segmentation on small-scale dataset. In: Proceedings of the 2nd international conference on control and computer vision—ICCCV 2019. pp 77–81