A Deep Learning Model for Image Classification Using Augmented Datasets

**ABSTRACT**

Deep convolutional neural networks have performed remarkably well on many Computer Vision tasks. However, these networks are heavily reliant on big data to avoid overfitting. Overfitting refers to the phenomenon when a network learns a function with very high variance such as to perfectly model the training data. Unfortunately, many application domains do not have access to big data, such as medical image analysis. This paper focuses on Data Augmentation, a data-space solution to the problem of limited data. Data Augmentation encompasses a suite of techniques that enhance the size, quantity and quality of training datasets such that better Deep Learning models can be built using them.

Paragraph 1: Introduce machine learning how it has grown and where it applied - note all must be well referenced even in other sections

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.The primary purpose of machine learning is to create an automatic system that can continuously improve itself without human interference. The only thing it uses is collected data that is analyzed by it to find specific patterns impossible to spot by people. This way, the machine can learn new things only by analyzing existing information and drawing conclusions out of it.

How it has grown

its use and adaptation in many industries especially due to its improved accuracy and efficiency

It is applied in many services we use today:

Search engines-an algorithm of search engines like Google or Bing. They use algorithms to curate even better results each time we look for something on the internet. Its work is based on our behavior – it analyzes how we respond to the data it shows us and updates its knowledge accordingly.

Social media-Learning algorithms are used for better search results, machine learning to help people with vision impairments to use the page by describing the details of pictures.

Financial industry-Also, big financial corporations often use machine learning to analyze data they collect and extract conclusions that can be beneficial for their growth.

Customer care-In online customer support, chatbots are becoming more and more popular.

Health industry-Using algorithms to diagnose different types of cancer is much easier than doing it manually. The machine only needs a small piece of data – for example, slides with cases of infected cells – to compare it with the new record and determine if it actually has cancer or not.

Virtual assistants-Smart speakers like Amazon Echo, as well as personal assistants in our smartphones (Siri, Google Now), are using machine learning to provide better answers, based on our previous requests.

These are a few areas where machine learning is applied in today's world

Paragraph 11: Talk about the data set problem , some solutions used and why the need for augmentation

There are many problems experienced with non-augmented datasets:

The high cost of collection and labeling data, data scarcity,overfitting which is experienced when training models and the class imbalance issue when classifying data

Some of the solutions include:

Using slightly modified copies of already existing data by transforming them in various ways i.e shifts, zooms,flips and scaling etc. hence expanding the training dataset

Creating synthetic data which are transformed images that will act new data to train our model

Why the need for augmentation

* This is to provide a wider scope of data for those training machine learning models. The focus is on the most frequently mentioned problem in the field of machine learning, that is the lack of sufficient amount of the training data or uneven class balance within the datasets.
* Having a Training data set with variations of images to try and see if your model will be able to identify the image even after augmentation
* A technique to increase the diversity of your training set by applying random (but realistic) transformations, such as image rotation.
* It is useful to improve performance and outcomes of machine learning models by forming new and different examples to train datasets. If dataset in a machine learning model is rich and sufficient, the model performs better and more accurate
* Machine learning applications, especially in the deep learning domain continue to diversify and increase rapidly. Data augmentation techniques may be a good tool against challenges which the artificial intelligence world faces.
* For machine learning models, collecting and labeling data can be exhausting and costly processes. Transformations in datasets by using data augmentation techniques allow companies to reduce these operational costs.

Paragraph 111: state the solution and how you will do it, highlight other studies that have used the same and how our work is different

The system is fed non-augmented images and its accuracy of identification is determined, Then it is fed augmented data and tested to see its accuracy in identifying the data. So, to get more data, we need to make minor alterations to our existing training data. Here we are specifically talking about Image Data Augmentation. These alterations may include flipping the image horizontally, vertically, padding, cropping, rotating, scaling and few other translations to objects in the image

How is this different from existing systems?

Presented will be a method of data augmentation based on image style transfer. The method allows the generation of new images of high perceptual quality that combine the content of a base image with the appearance of other ones. The newly created images can be used to pre-train the given neural network in order to improve the training process efficiency.

Paragraph 1111: Give the significance or benefits from the solution

●Improving model prediction accuracy

● adding more training data into machine learning models

● preventing data scarcity for better models

● reducing data overfitting ( i.e. an error in statistics, it means a function corresponds too closely to a limited set of data points) and creating variability in data

● increasing generalization ability of the models

● helping resolve class imbalance issues in classification

● reducing costs of collecting and labeling data

Lastly a brief summary of how the rest of the paper is organized

* **LITERATURE REVIEW**

Present studies that have used the same techniques and the results and the gaps in each that we will fill

a.Machine Learning:

Add literature on machine learning, the different techniques and algorithms applied with focus on advantages of your selected method

LITERATURE REVIEW

Other approaches Currently developed image augmentation methods are not limited to the traditional and CNN-based methods. One of interesting approaches is a random erasing technique which is fast and relatively easy to implement yet giving good results in generalization ability of Convolutional Neural Networks. In the method one randomly paints a noise-filled rectangle in an image resulting in changing original pixel values. As the authors explained, “ expanding the dataset with images of various levels of occlusion reduces the risk of overfitting and makes the model more robust.”

THE METHOD SELECTED

Style transfer;

Style transfer for data augmentation High potential of style transfer idea still remains underestimated in the field of image augmentation. In this paper we would like to propose a fresh look at the style-transfer technique and to prove that with expert knowledge it can be applied in a variety of fields. Currently this method is widely used mostly to synthetize artistic style to create new visually pleasing artworks. In some recent works authors proposed the use of this method to change pictures from day to night , clear day to rain , or other landscape changes , however currently it is still applied in very limited fields in contrast to the wide existing possibilities.

We want to emphasize that the potential of this technique can find application also in many different fields. For example: mixing two images ended in creating a new one with more dense structure and different color palette

With expert knowledge the style transfer method can be applied in a variety of fields, especially in images from the same category but of a slightly different type. It is worth noticing that even with the data of lower quality, algorithms can perform better as long as useful information can be extracted. Artificially augmented datasets can be used to pre-train the CNN ( using transfer learning technique) in order to obtain the suitable initial conditions of the network weights and biases and then trained once again on the real dataset, to fine tune the searched parameters.

B. XXX prediction

Highlight why xx prediction is important and different non machine learning techniques that have been used tin the same area

C. Use of machine learning in XXX prediction

Review studies that have used machine learning on the same area add show what new we are doing

D. A summary of lit gap and how the research fills the gap

* **METHODOLOGY**

Give details of the tool and methods, the data sets, ML techniques and the ML steps other tools used and why

Tools;

Tensorflow is an open source library developed by Google primarily for deep learning applications. It is written in C++, Python and CUDA.

TensorFlow accepts data in the form of multi-dimensional arrays of higher dimensions called tensors. Multi-dimensional arrays are very handy in handling large amounts of data.

TensorFlow works on the basis of data flow graphs that have nodes and edges. As the execution mechanism is in the form of graphs, it is much easier to execute TensorFlow code in a distributed manner across a cluster of computers while using GPUs.

Platform;

The training was done on google colab, this is because it runs entirely on the cloud and it does not require setup. Google colab is well suited for machine learning and data analysis, since it is a hosted jupyter notebook service, it provides access to computing resources including Graphics Processing unit (GPUs.)

The platform supports many machine learning libraries which are necessary for this project, it allows writing and execution of python code. We are also able to import external datasets.

Methods;

We are using supervised learning which builds a model that makes predictions based on evidence in the presence of uncertainty. A supervised learning algorithm takes a known set of input data and known responses to the data (output) and trains a model to generate reasonable predictions for the response to new data. Use supervised learning if you have known data for the output you are trying to predict.

Supervised learning can be separated into two types of problems when data mining—classification and regression;

For this project we employed classification

Classification uses an algorithm to accurately assign test data into specific categories. It recognizes specific entities within the dataset and attempts to draw some conclusions on how those entities should be labeled or defined.

Under classification specifically we used neural networks which is a process training data by mimicking the interconnectivity of the human brain through layers of nodes. Each node is made up of inputs, weights, a bias (or threshold), and an output. If that output value exceeds a given threshold, it “fires” or activates the node, passing data to the next layer in the network. Neural networks learn this mapping function through supervised learning, adjusting based on the loss function through the process of gradient descent. When the cost function is at or near zero, we can be confident in the model’s accuracy to yield the correct answer.

This technique is better because;

* With the help of supervised learning, the model can predict the output on the basis of prior experiences.
* In supervised learning, we can have an exact idea about the classes of objects.
* Supervised learning models help us to solve various real-world problems such as fraud detection, spam filtering, etc.
* Neural Networks have the ability to learn by themselves and produce the output that is not limited to the input provided to them.
* The input is stored in its own networks instead of a database, hence the loss of data does not affect its working.
* These networks can learn from examples and apply them when a similar event arises, making them able to work through real-time events.
* Even if a neuron is not responding or a piece of information is missing, the network can detect the fault and still produce the output.
* They can perform multiple tasks in parallel without affecting the system performance.

Datasets;

Link to the dataset i used

<https://drive.google.com/drive/folders/1dadZbSsqYV7Mf3lmk5x-TWEeqL2nww1e?usp=sharing>

Description of the dataset;

This is a large dataset of images of flowers.The dataset contains approximately 3,670 images divided into five classes according to the type of flower: Daisies, Roses, Sunflowers,Tulips and Dandelions

This was my dataset of choice because flowers are recognizable objects and the augmentations on the images will also be clear to viewers

TensorFlow Datasets is a collection of datasets ready to use, with TensorFlow or other Python ML frameworks, such as Jax. All datasets are exposed as tf.data.Datasets , enabling easy-to-use and high-performance input pipelines. Which is what we used to carry out the augmentation in this project

In order to use a Dataset we need three steps:

* Importing Data. Create a Dataset instance from some data.
* Create an Iterator. By using the created dataset to make an Iterator instance to iterate through the dataset.
* Consuming Data. By using the created iterator we can get the elements from the dataset to feed the model.

Model training steps

* downloading the dataset this project uses the [tf\_flowers](https://www.tensorflow.org/datasets/catalog/tf_flowers) dataset from tensorflow for convenience. It is downloaded to our device
* Using keras preprocessing layers to resize and rescale the images to a consistent shape



above is a visualization of the image proportions that will be applied to all the images in the dataset. The pixels are in the[0, 1] range

* Data augmentation- using the keras preprocessing layers we create a few preprocessing layers approximately 5 and apply them repeatedly to the same image





We end up with this visualization having applied random augmentations ie change image brightness,change image contrast, add random hue, add saturation, alter the quality and cropping

* The preprocessing layers are now applied to our dataset, with this approach we create a dataset that yields batches of augmented images
* Now the model is trained with the prepared datasets

We train five epochs as an example and this is the outcome;

Epoch 1/5

92/92 [==============================] - 19s 96ms/step - loss: 1.2814 - accuracy: 0.4632 - val\_loss: 1.1246 - val\_accuracy: 0.5804

Epoch 2/5

92/92 [==============================] - 3s 26ms/step - loss: 1.0684 - accuracy: 0.5753 - val\_loss: 0.9848 - val\_accuracy: 0.6213

Epoch 3/5

92/92 [==============================] - 3s 28ms/step - loss: 0.9701 - accuracy: 0.6100 - val\_loss: 1.0106 - val\_accuracy: 0.6185

Epoch 4/5

92/92 [==============================] - 3s 26ms/step - loss: 0.9046 - accuracy: 0.6502 - val\_loss: 0.9455 - val\_accuracy: 0.6076

Epoch 5/5

92/92 [==============================] - 3s 27ms/step - loss: 0.8676 - accuracy: 0.6717 - val\_loss: 0.9243 - val\_accuracy: 0.6512

The result is;

12/12 [==============================] - 1s 14ms/step - loss: 0.8008 - accuracy: 0.6839

Accuracy 0.6839237213134766

**RESULTS AND Discussion**

Present and discuss results of training

* Found 3670 files belonging to five classes
* Using 2936 files for training
* Using 734 files for validation
* The batch size is 32 i.e the images are being passed through the model in batches of 32 at a time
* Epochs are 15 ie an epoch when we have input the entire dataset once, the entire dataset has passed through the network, the network has seen every single training example once that is an epoch ie one forward pass and one backward pass of the training examples through the network if i have 1000 examples and i use 1 epoch that means that data has been passed through the network once(which is not enough and will cause underfitting)

We have 3670 images divided into 32 batches this will result in 115 iterations to finish 1epoch, we have 15 epochs meaning we have done this 15 times

Images before augmentation = 3670

Images after augmentation= 55,050

The following graphs are a result of code written to visualize our training results



The plots show that training accuracy and validation accuracy are off by large margins, and the model has achieved only around 60% accuracy on the validation set.after inspection on what went wrong, overfitting is the main issue which is when there are a small number of training examples, the model sometimes learns from noises or unwanted details from training examples,to an extent that it negatively impacts the performance of the model on new examples.

We implement data augmentation ie random flip, random rotation and random zoom demonstrated below on the same image several times;



After applying the augmentation again there is less overfitting than before shown by the graphs below

There is an increase in training and validation accuracy compared to before, the losses have also reduced gradually indicating that the model has displayed improvement



* First train before augmentation
* Augmentation results
* Training results after augmentation
* Discussion of the results

To show the effects of image augmentation on the model, we will be training the same model multiple times with different image augmentation techniques and will compare the test accuracies of each model. We will be training 4 different models, 1 without augmentation and 3 models with different sets of image

augmentation techniques.

* Grayscale an image



* Creating custom dtd augmentation layers where both layers will randomly invert the colors of an image according to some probability



3.creation a random invert custom layer that inverts the original colors of the image



4. Neural style transfer is an optimization technique used to take two images—a content image and a style reference image (such as an artwork by a famous painter)—and blend them together so the output image looks like the content image, but “painted” in the style of the style reference image.

This is implemented by optimizing the output image to match the content statistics of the content image and the style statistics of the style reference image. These statistics are extracted from the images using a convolutional network.



The visualized result



* **Conclusion**

Conclusion showing what was done and the major findings summary

We can clearly conclude that by the implementation of image augmentation to our model the accuracy is increased, which is a major advantage of using augmentation techniques, this also reduces the overfitting making the model more robust to real world scenarios.

**VI. References**

(“Sci-Hub | Image Augmentation with Neural Style Transfer. ICT Innovations 2019. Big Data Processing and Mining, 212–224 | 10.1007/978-3-030-33110-8\_18”)

(“Sci-Hub | Data Augmentation for Improving Deep Learning in Image Classification Problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW) | 10.1109/IIPHDW.2018.8388338”)

(“Sci-Hub | a Survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1) | 10.1186/S40537-019-0197-0”)

Shorten, Connor. “Neural Style Transfer:Applications in Data Augmentation.” *Medium*, Towards DataScience, 8 Oct. 2018,towardsdatascience.com/neural-style-transfR-applications-data-augmentation-43d1dc1aecc. Accessed 15 Mar. 2022.

‌Sanchit Tanwar. “Effects of Image Augmentation on Model Performance - Analytics Vidhya - Medium.” Medium, Analytics Vidhya, 8 Sept. 2021, medium.com/analytics-vidhya/effects-of-image-augmentation-on-model-performance-3f8effbf54e5. Accessed 17 Mar. 2022.

“Google Colab - What Is Google Colab?” Tutorialspoint.com,2021, [www.tutorialspoint.com/google\_colab/what\_is\_google\_colab.htm](http://www.tutorialspoint.com/google_colab/what_is_google_colab.htm). Accessed 21 Mar. 2022.

‌<https://www.facebook.com/greycampus>. “Neural Networks: Advantages and Applications.” MarkTechPost, 18 Apr. 2019, [www.marktechpost.com/2019/04/18/introduction-to-neural-networks-advantages-and-applications/](http://www.marktechpost.com/2019/04/18/introduction-to-neural-networks-advantages-and-applications/). Accessed 21 Mar. 2022.

“Supervised Machine Learning - Javatpoint.” [Www.javatpoint.com](http://www.javatpoint.com/), 2021, [www.javatpoint.com/supervised-machine-learning](http://www.javatpoint.com/supervised-machine-learning). Accessed 21 Mar. 2022.

‌IBM Cloud Education. “What Is Supervised Learning?” Ibm.com, 19 Aug. 2020, [www.ibm.com/cloud/learn/supervised-learning](http://www.ibm.com/cloud/learn/supervised-learning). Accessed 21 Mar. 2022.

‌

‌

‌