PREDICTING SOLAR FLARES WITH NEURAL NETWORKS

VIKALP UPADHYAY

VIKALP.UPADHYAY@UTDALLAS.EDU, 07/ 24/ 2019

1.EXECUTIVE SUMMARY

Sporadic events can be hard to predict. Solar flares are one of such events. Solar flares can produce an enormously energetic stream of protons. If encountered, it can be fatal to spacecrafts, astronauts and even some parts of earth. Therefore, it is necessary to develop a reliable system to predict them accurately.

The job of predicting solar flares is easier said than done as they have often been compared to predicting the weather. Currently, they are predicted by experts and partially automated systems using image feature extraction. I reckon that CNNs can ease this task. First, they are cheaper and more accurate. Furthermore, they can even offer insights into the underlying physical systems.

My interest towards this rather uncanny phenomenon has drawn me to build a neural network model for peak flux prediction of solar flares. In this project, I have built two models. One is baseline and the other utilizes neural networks.

By running these models on image dataset of magnetogram sensor (onboard Solar Dynamics Observatory satellite), I evaluated the performance of these models. Furthermore, I have calculated MAE corresponding to each model. As a final step, I have discussed scopes of further improvement in this model.

2.CONTENTS

3.INTRODUCTION

Within a decade or so, Convolutional Neural Networks (CNNs) have achieved comparable performance to, and in some cases even surpassed, manual and conventional feature extraction methods in the fields of image recognition and natural language processing.

Many CNN architectures have been successfully adapted and employed in other domains, such as self-driving cars, the prediction of atomic and molecular interactions, natural language processing and we even find them on a small scale in our phones. CNNs promise to be adaptable to many more domains.
One application domain which could benefit from CNNs is the prediction of solar flares. Additionally, the analysis of the trained networks can enable domain experts to discover previously unknown properties of the underlying systems.

A major challenge in applying machine learning techniques to this problem is preprocessing the dataset used to train any potential model. One of the primary challenges to preprocessing such a dataset is the fact that the amount of relevant data is relatively small. In this case, relevant data refers to samples of large to extreme solar flares. This also leads to a significant bias within the small dataset.

4.SOLAR FLARE PREDICTION

a. SPACE WEATHER

Space weather refers to various temporary or permanent conditions and events within our solar system. Some of which can have an adverse effect on human systems on earth or within space. Moreover, for some of these events, it would be scientifically beneficial to be able to study them with precise instruments for their entire duration. Ideally instruments would be targeted at these events, even before they occur.[1]

b. SOLAR FLARES

Solar flares are a source of space weather events of particular interest. They are a sudden moment of significantly increased brightness on the sun. Usually, solar flares occur in close proximity of a sunspot group and often, but not always, coincide with a coronal mass ejection. They produce extremely high energy particle streams, which can damage spacecraft and astronauts through radiation. Additionally, very powerful solar flares, accompanied by massive coronal mass ejections, can fully disable satellites and terrestrial power, navigation and communication grids for extended periods of time. These and various other risks posed by solar flares are of growing concern, as humanity is expanding its sphere of influence and habitation in the 21st century. Therefore, the fast, cheap and precise prediction of these events is increasingly important. Such predictions would allow us to aim observation instruments at these events in advance, to study them in greater detail. Eventually they could also enable possible countermeasures or evasive maneuvers.[2]

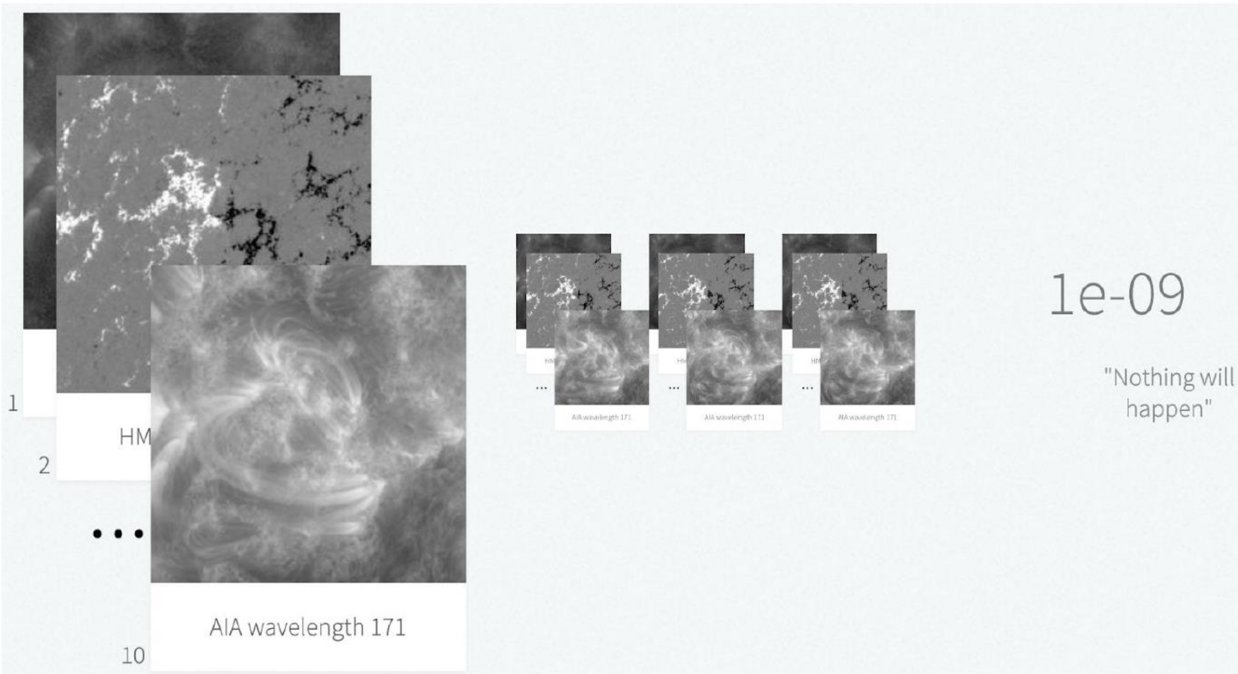
c. SOLAR FLARES PREDICTION

Currently, most solar flare predictions are almost exclusively done manually by experts, through the analysis of sunspots and their characteristics. However, such manual methods are relatively expensive, slow and imprecise. Recently, there have been efforts to develop partially and fully automated systems. These systems feature high complexity and require extensive domain knowledge. Partially automated systems employ conventional feature extraction and image analysis methods in an initial step. In a second step an expert then uses these features and the original images to make a prediction. Alternatively, neural networks have been applied to take these features and make a final prediction. These techniques are comparable to the methods employed by automated terrestrial weather prediction systems and early image recognition systems.

5.DATA ANALYSIS

The dataset is fairly complex.

a. DESCRIPTION



The dataset consists of about 8,300 training samples and about 900 test samples. Each sample contains ten different
image types at four different points in time before the targeted 24-hour prediction period.
Each set of images is already cropped to a small region of interest on the sun.

The ten different image types are made by the AIA (Atmospheric Imaging Assembly) and HMI (Helioseismic and Magnetic Imager) detectors onboard the Solar Dynamics Observatory satellite mission [3], targeting various wavelengths. The four points in time are approximately twelve hours, five hours, one and a half hours and finally ten minutes before the previously mentioned 24-hour prediction period. The labels of the samples, and therefore, the numbers that should be predicted by a machine learning model, are the peak flux values on the sun within the 24-hour prediction period. Flux, in this case, refers to the rate of flow of high energy particles ejected from the sun. Solar flares are classified using the letters A, B, C, M and X, in accordance to their peak X-ray flux. X-rays having wavelengths of 100 to 800 picometre are considered. The peak flux is measured in watts per square meter (W/m^2), by the GOES spacecraft.

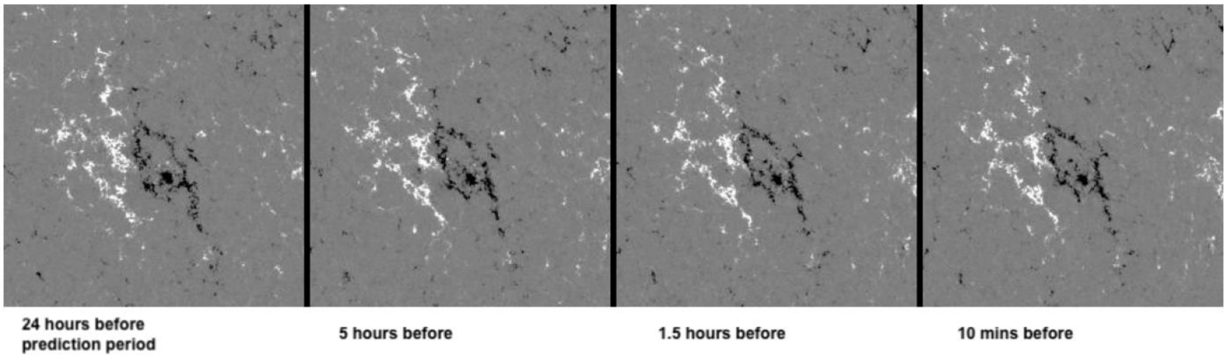
Classification	**Peak flux range** at 100–800 picometre (watts/square metre)
A	< 1e-7
B	1e-7 - 1e-6
C	1e-6 - 1e-5
M	1e-5 - 1e-4
X	> 1e-4

b. DATA IMBALANCE

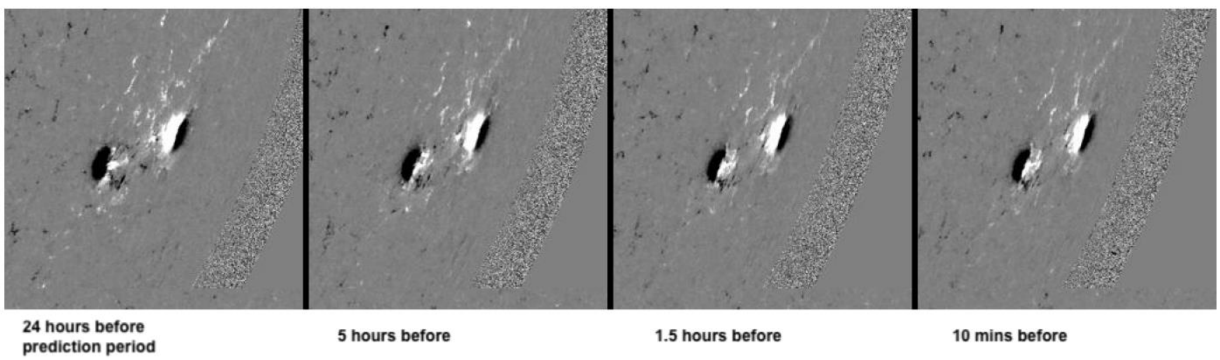
Only 6.5% of the samples are in the classes M and X, while class A makes up almost 44% percent of the samples on its own. Such a biased dataset makes this machine learning problem significantly harder. It is also notable, that the samples stem from very few regions on the sun.

Classification	# of Samples	% of Samples	# of Different Regions on Sun	% of Different Regions on Sun
A	3652	43.8 %	450	29.5 %
B	1210	14.5 %	106	7.0 %
C	2936	35.2 %	407	26.7 %
M	500	6.0 %	111	7.3 %
X	39	0.5 %	17	1.6 %

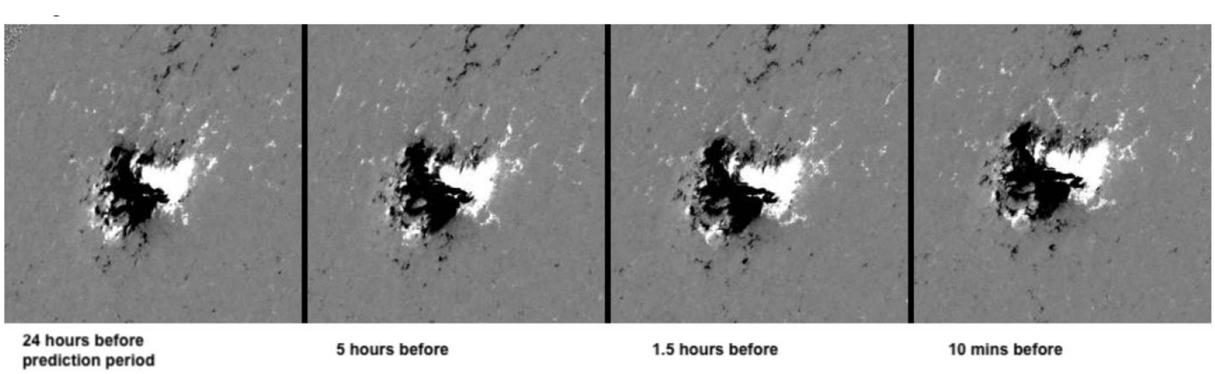
Traditionally, only the magnetogram image type has been used to predict solar flares. Therefore, I am going to implement my model using just magnetogram images.



Magnetogram images for a solar flare with low flux.



Magnetogram images for a solar flare with medium flux.



Magnetogram images for a solar flare with high flux.

It is notable that the spots for a high flux flare are bigger and much more complex. Moreover, its edges have become irregular.

6.MACHINE LEARNING

a. PREPROCESSING AND AUGMENTATION

Missing values have to be removed from the data. This makes the dataset smaller than desired, but again, that is almost always the case.

When datasets are smaller than desired, which is almost always the case since more training data increases almost any model's performance, data augmentation is employed. This commonly done with images, and can involve random cropping, zooming, rotations and changes to the brightness.

b. CONVOLUTIONAL NEURAL NETWORKS

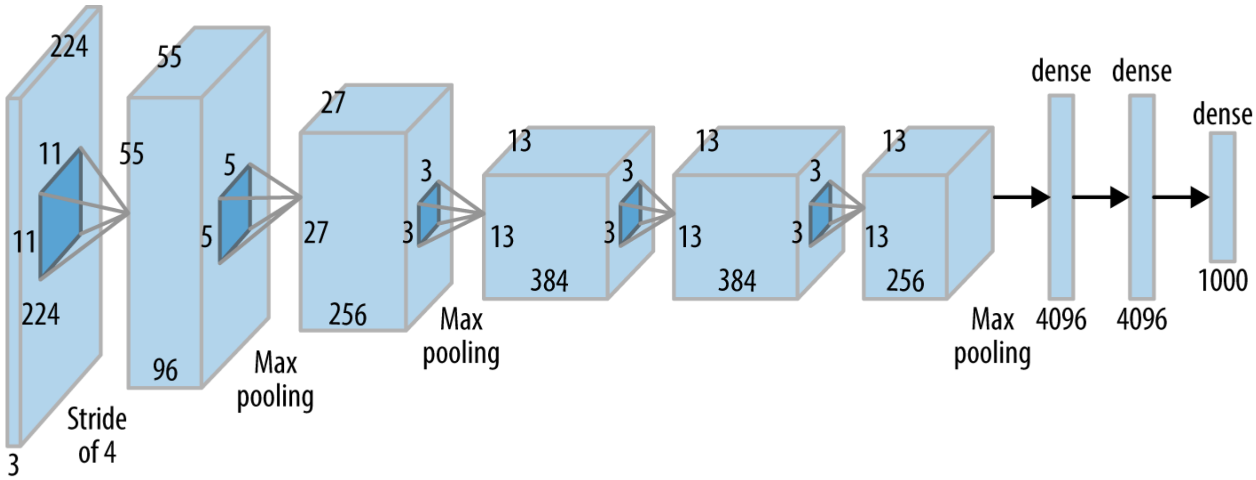
Conventional neural networks do not scale well to larger images, due to them requiring a weight for each pixel in each channel for each neuron on the first layer. They also do not make use of the three-dimensional nature of images with multiple channels.

Convolution layers use rectangular image filters, that is matrices of weights, to extract a single value from a region in the input data. This is done using the well-known convolution operation in mathematics. These filters are moved across the input data in order to compute the value of each neuron of a single dimension in the next layer. Additional filters result in additional dimensions in the output, while additional dimensions in the input lead to filters with greater depth. Pooling layers are similarly constructed but use different operations to extract a value from a given region of the input data. The goal is to reduce the height and width of the next layer, so the filters are moved more than one pixel to calculate the value for the next neuron, thus resulting in fewer values per dimension than in the input.

For each neuron, after summing up its inputs and adding its bias, the value is run through an activation function. The results of that function are then forwarded to the next layer. I have used ReLU as activation function as it is not only faster to compute, but also tackles the vanishing gradient problem more effectively.

c. ARCHITECTURE

The architecture I used for this neural network is similar to Alexnet. AlexNet model first applies a 11x11 convolution layer to the input with a stride of five. This indicates that the eleven by eleven filter of is moved by four pixels in the input image for each pixel in the output image. The output of the first convolution layer is thus about one quarter of the size of the input. The next layer is a max pooling layer that further reduces the size, followed by another convolutional layer and so on. Altogether, the Alexnet model has five convolution layers followed by three conventional fully connected layers. There are around 60 million parameters in total. [4]



d. EVALUATION OF MODELS

I have considered MAE (Mean absolute error) as the evaluation metric. Along with MAE, I have included True Skill Statistic and Heidke Skill Score as additional evaluations, very similar to what we use to evaluate weather predictions.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}$$

7.MY MODELS

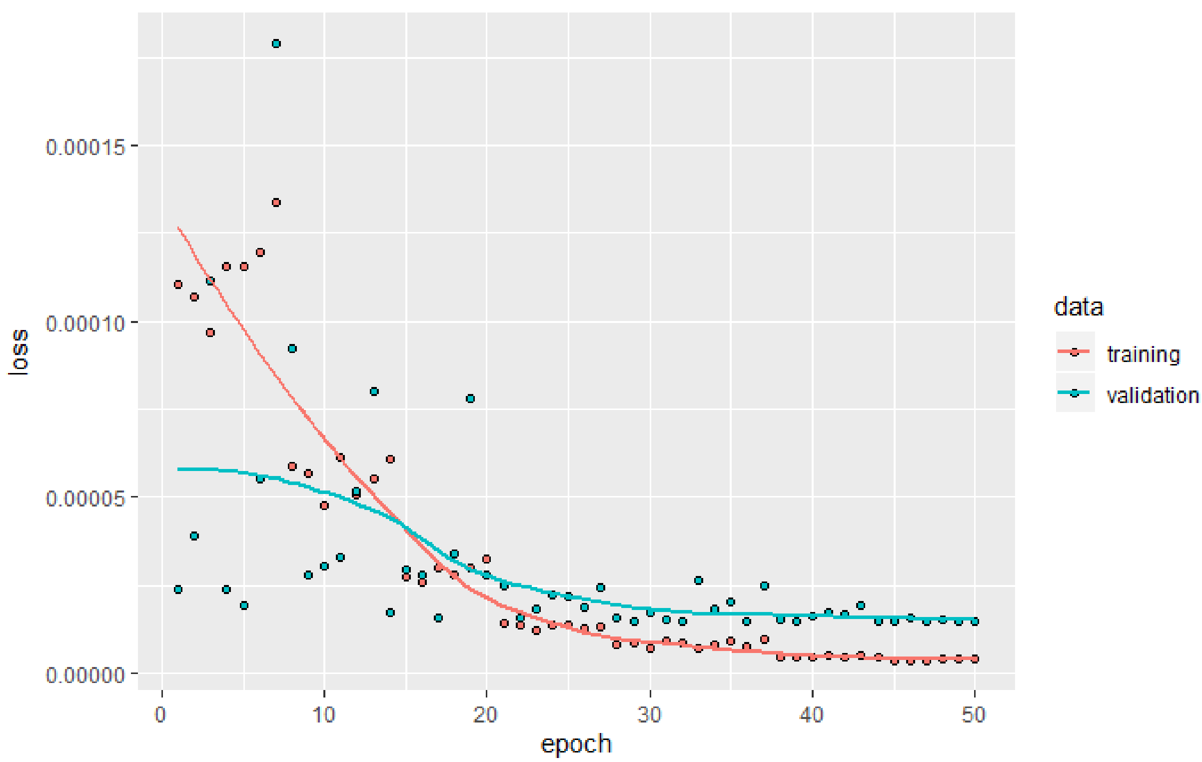
a. BASELINE MODEL

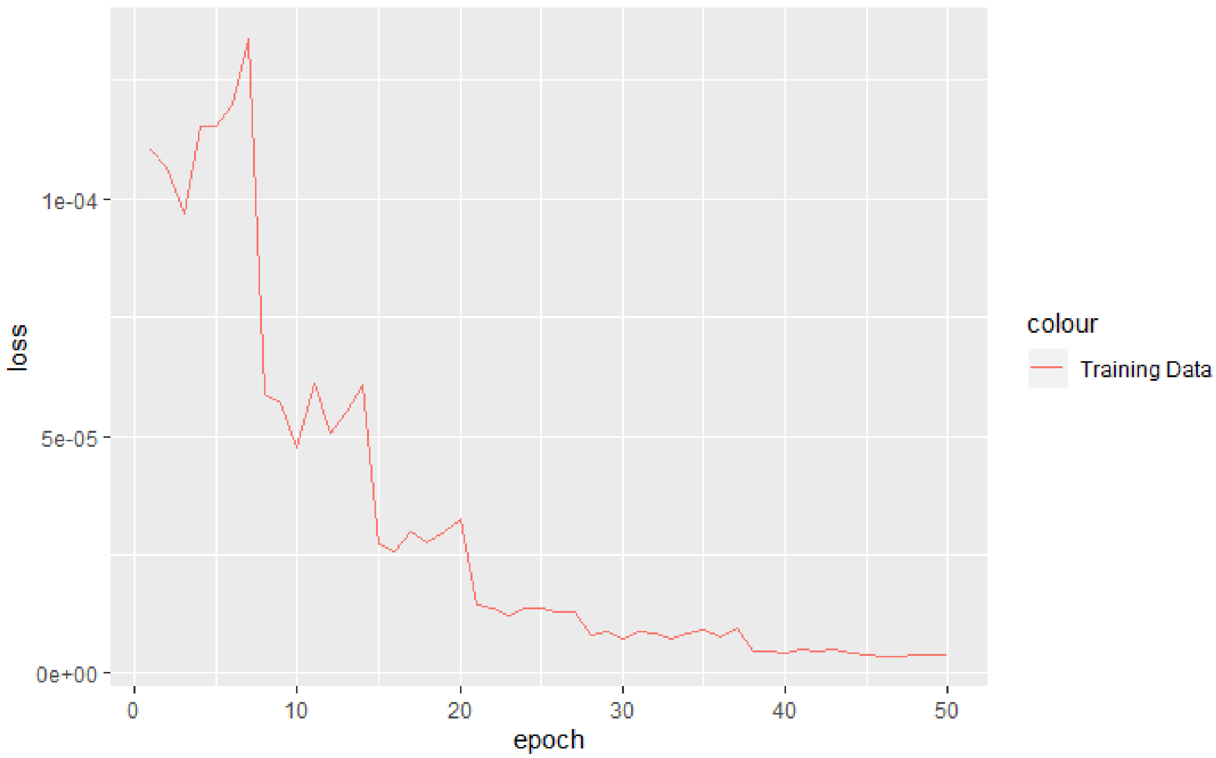
First model I have used is a baseline model. This model makes a static prediction independent of any input it is provided with. The optimal static value for the MAE metric is the median of the goes flux values in the training data. I have set the static prediction peak flux at 5.3e-07. Using this value, I scored a MAE of 1.50363043420528e-05.

b. COMPLEX NEURAL NETWORK

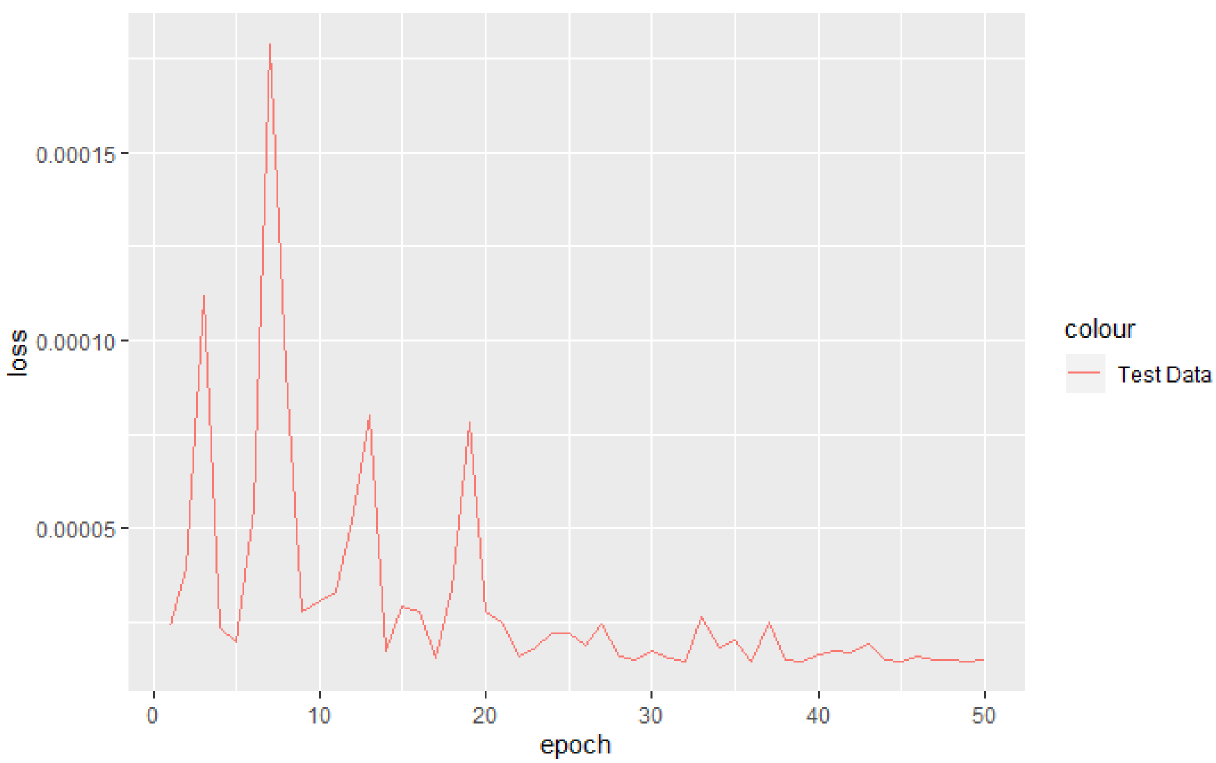
My neural network model has several dense, convolutional and pooling layers like Alexnet. This model required extensive data preprocessing into a 4-dimensional array. Additionally, to monitor the learning rate, I have used callback_reduce_lr_on_plateau. In order to fit data to keras fit_generator, I had to use keras image_data_generator.

Output





1e-04
5e-05
0e+00
loss
0
10
20
30
40
50
epoch
colour
Training Data



0.00015
0.00010
0.00005
loss
0
10
20
30
40
50
epoch
colour
Test Data

8.FURTHER SCOPES

a. REGRESSION BY CLASSIFICATION

Turning regression problems into classification problems and then applying leading machine learning classification techniques results in excellent predictions, oftentimes beating identical models with a direct regression output.

b. DATA AUGMENTATION

A solution to the problem of not having enough data would be employing additional data augmentation methods, to get more training out of the same base data. One concrete augmentation method would be making minor adjustments to the brightness of the images.[5]

9.REFERENCES

1. [1] “Space weather,” *Wikipedia.*
2. [2] “Solar flare,” *Wikipedia.*
3. [3] “Solar Dynamics Observatory,” *Wikipedia*.
4. [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,”
5. [5] “Solar cycle,” *Wikipedia*.