

SUNSPOT DETECTION USING DEEP LEARNING TECHNIQUES

Pieter Swanepoel, Günther Richard Drevin, Roelf Du Toit Strauss and Petrus Johannes Steyn
North-West University Potchefstroom, South Africa

ABSTRACT

Accurate and automated monitoring of space weather has become increasingly important in our modern-day world, due to the reliance on critical infrastructure such as wireless communication, power grids, satellite communication, and aviation systems. Solar events such as solar flares and coronal mass ejections can disrupt or damage these systems and sunspots are used as markers to determine solar activity through a manual process, thus automated detection and tracking of sunspots form part of the prediction of solar events and prediction of space weather. In this paper the use of deep learning techniques for automated sunspot detection will be investigated.

KEYWORDS

Sunspots, Space Weather, Deep Learning, Object Detection

1. INTRODUCTION

Sunspots are dark regions characterized by a dark core called the umbra and a filamentary outer layer named the penumbra that appears in the photosphere of the Sun which starts at the surface and extends about 500km and can be observed in the visible parts of the continuous radiation spectrum. Sunspots appear darker due to the difference in temperature compared to the surrounding photosphere. Umbral diameters may vary from 10^4 km to 20^4 km (Wilkinson, 2012). Penumbra diameters vary between 10^4 km and 15^4 km. Sunspots start out as small dark spots referred to as pores, whereas only certain pores develop into full Sunspot regions. These regions may then grow larger and darker and even break away to form groups of sunspots. Sunspots are a phenomenon due to the buildup of magnetic field pressure in the core of the Sun. Due to the pressure then becoming larger than that of the outer layers, the spot starts rising to the outer layers and becomes visible on the surface. The plasma processes that cause the phenomena is due to the dense hot plasma that is confined within the tangled magnetic fields that are rising from the interior part of the Sun, and these phenomena contain large amounts of free magnetic energy which is thought of as the main source of transient phenomena. Re-configurations of these magnetic fields may be accompanied by clouds of plasma that are ejected into interplanetary space, also referred to as a coronal mass ejection. When these eruptions occur at smaller scale, electrons are released that can collide with denser matter at the surface of the Sun which produces a burst of electromagnetic radiation which is particularly noticeable in the EUV and X-ray wavelengths called solar flares (Hanslmeier, 2004). As the Sun rotates these sunspots or sunspot groups travel along the Sun from east to west and may shear and change shape. One theory suggests sunspots are formed by the convergence of magnetic elements and pores which need a driving force. This force is theorized to be the individual fragments that form sunspots and are part of a larger flux tube in the convection zone as opposed to a coalescence of smaller flux tubes (Solanki, 2003). Sunspots not only have visual characteristics, but the magnetic structure of a Sunspot can also provide information regarding the solar magnetic field of the Sun. The polarity of a Sunspot is directly related to the solar cycle, where 22 years is a full cycle. This is due to the polarity of the Sun that changes with every 11 years (Wilkinson, 2012). Sunspot numbers were found to vary according to a 11-year cycle, whereas the relative Sunspot number goes through a maximum and minimum known as the solar maximum and the solar minimum (Solanki, 2003). As a rule, it is regarded that the Sunspot cycle is closely related to the magnetic structure of the Sun, due to how Sunspot numbers vary as the magnetic structure of the Sun changes. At the beginning of a solar cycle, most sunspots will form at 25 degrees north or south of the solar equator, whereas, later in the cycle, they will form closer to the equator. This formation of sunspots only in those regions and

latitudes indicate that they are only forming in activity belts of the Sun. Because sunspots are only in very active regions leads to the understanding that sunspots are the heart of the active regions with a great deal of magnetic activity surrounding the sunspot (Solanki, 2003). Figure 1 shows there exists a clear correlation between solar events and the number of sunspots on the solar surface. When the Sunspot cycle goes through a maximum, there is a peak in solar events.

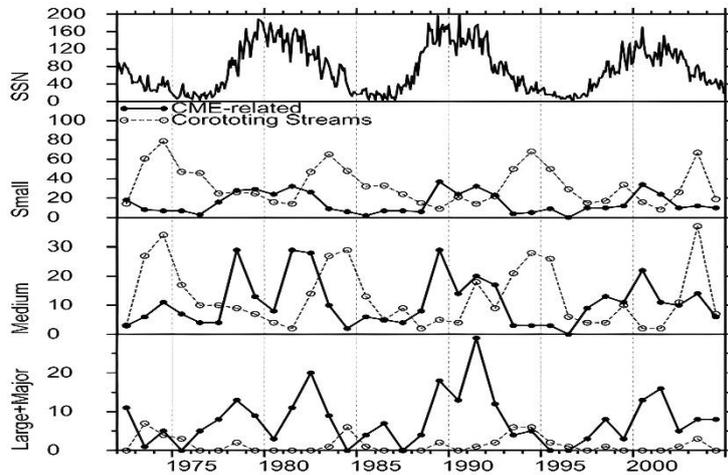


Figure 1. Sunspot numbers with associated small, medium, and large solar events

Solar events occur in multiple stages and start with the emission of energetic X-ray photons which is more commonly called a solar flare and affects various systems and processes. Following the X-ray flash, there is a release of solar energetic particles which then propagate into interplanetary Space (Hanslmeier, 2004, Hapgood, 2017). This is then followed by a coronal mass ejection which is a continuous release of particles and solar plasma but is much slower and takes longer to reach Earth (Hanslmeier, 2004). Solar flares and coronal mass ejections both contribute to space weather and influence Earth in very different ways. In general, there are two types of solar flares: the first type is a simple loop flares which is not usually associated with coronal mass ejections due to the flare being very small, usually brightens up, and remains unchanged because it only consists of a single magnetic flux tube. The other type of flare, referred to as a two-ribbon flare, is generally associated with major class types and can be much more dramatic with a longer duration, which is associated with coronal mass ejections. The flux of high energy particles produced by two ribbon flares enhances solar wind, which compresses the magnetosphere and increases the magnetic field near the surface of the earth (Hanslmeier, 2004). Once the solar wind plasma enters the magnetosphere it is stored in the tail of the magnetosphere and can cause a re-connection event which releases an explosion of stored energy that causes what is known as a sub-storm. Major solar events can create a sudden large influx of energy into the magnetosphere which can then set off an ensemble of sub-storm, termed geomagnetic storms (Hapgood, 2017). Being aware of solar events, will allow for proper mitigation plans to be in place, and avoid the consequences of these events. This can be achieved through automated monitoring and notifications systems, to allow proper measures to be in place. The primary aim of this study is to train and test state of the art object detection models to detect and monitor sunspots. Related Research surrounding event detection using images alongside text-based reports have been performed to classify events with an 88% accuracy (Fang, 2019). This was done using a convolutional neural network. Similar areas of study include brain metastases detection due to the appearance being like the blob structure of sunspots. This study used SSD for metastases with an 84% accuracy (Zhou, 2019). Other Studies include Condition classification using a CNN and a support vector machine that achieved a 99% accuracy (Channabasava & Biabl, 2022).

2. METHOD

The Data set used to train the models consists of 815 images of which the training data was separated into 570 for training, 164 for validation and 82 for testing. The images were obtained from the Royal Observatory of Belgium, with their corresponding arch drawings on which manual sunspot identification was done. These arch drawings were used in the labeling as the ground truth source. To further enhance the dataset and have more training data the initial 570 training images underwent a series of pre-processing steps to not only enhance, but also augment the images. The enhancement was needed as some images are very partially obscured when there are clouds in the sky. The images span across the start of 2008 to 2019, of which there are images which were taken that had no sunspots which were discarded. Object detection models from different areas were implemented falling into either single stage or two stage detection, which have various tradeoffs regarding speed and accuracy under different circumstances. The models trained for this paper were faster RCNN, YOLO (You Only Look Once), SSD (Single Shot Detector), and EfficientDet using a NVIDIA V100. Thus, comparing which of the object detection models have the best precision in detecting sunspots on white light images.

2.1 Regional Proposal (Two Stage) vs Regression Classification (One Stage)

High level overview of two stage and one stage detection:

- Region proposal is a two-step process which was designed to extract regions of interest and was inspired partially by the attentional mechanism of the human brain (Zhao, 2019). This mechanism functions by scanning the entire scenario coarsely, and then focuses on the regions that are of interest. a convolutional neural network is used as a backbone to extract features, which allows predictions of bounding boxes directly from locations of the topmost feature map after obtaining the relevant confidences of the underlying objects (Zhao, 2019).
- The regression classification detection framework is based on global regression classification, which maps straight from the pixels to bounding box coordinates and class probabilities. Single stage detectors were adapted and designed to use deep neural networks with regression final layers instead of soft-max activations as the final layers which is capable of generating object binary masks (Zhao, 2019). The final step is to draw the bounding boxes from the masks, by re scaling the masks to the input image and estimating the coordinates.

2.2 Training Process

High level overview of modelling and training:

1. Build a config and training pipeline file in framework format
2. Process and format data into correct format for the model
3. Iterate with training at different batch sizes and number of epochs
4. Adjust hyper parameters
5. Train the model until no meaningful precision increase occurs without overfitting
6. Test and validate the model against the testing images
7. Change the model back bone and or architecture
8. Iterate the process again.

2.3 Faster RCNN (Two Stage)

The Faster RCNN model was trained using Detectron 2, TensorFlow 2 and PyTorch. All the frameworks use different training formats and input formats, and this was done to note if there is any meaningful difference between a model between the frameworks. Different backbones were also used throughout training to establish if a deeper backbone that has more training parameters would yield better precision or only add more training time and yield no increase in average precision.

Overview of Faster RCNN training process:

- The region proposal network receives an image of the sun as input from a convolutional feature extractor which outputs a feature map

- The feature map is then fed into the region proposal network which applies a sliding window protocol with a small network.
- The sliding window will then map the features to a lower dimensional or more recognizable feature.
- The lower dimensional features are then fed into two sibling fully connected layers which are a box-regression layer and a box-classification layer to perform detection and classification.

Figure 2 illustrates a detection by one of the Faster RCNN models, specifically showing the capability of the model to detect sunspots that are just going out of sight on the outer edge of the solar limb.

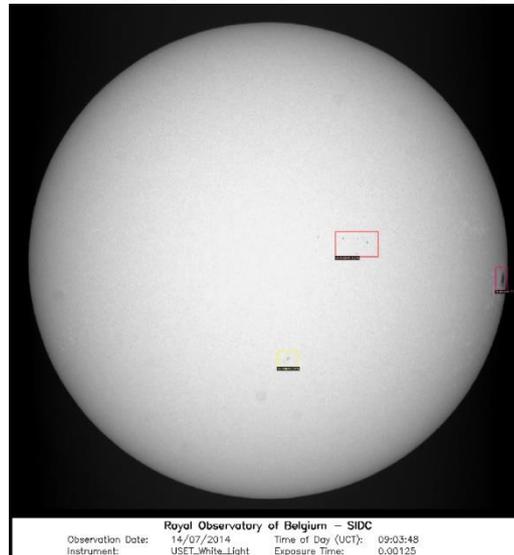


Figure 2. Faster RCNN sunspot detection

2.4 YOLO

The YOLO model was trained using PyTorch and TensorFlow 2, as Detectron 2 does not support any YOLO architectures. Different formats of data structures were used for the two frameworks and all the architectures of yolo available including the darknet architecture were trained and iterated with using different hyperparameters and architectures.

Overview of YOLO training process:

- The input image is divided into a grid, where if an object falls into the center of a grid cell that cell is responsible for detecting that object.
- These grid cells each predict several bounding boxes and confidence scores for the corresponding boxes, which are then used to evaluate the intersection over union using the ground truth
- If a box does indeed contain a object that value will be assigned to the box or zero otherwise.
- Each of the grid cells have corresponding bounding boxes with predictions, labelled and the confidence where the center is of the box, which are relative to the bounds of the grid cell.
- The height and the width are predicted relative to the whole image.
- Confidence is predicted using the ground truth box and the intersection over union between the predicted and ground truth box.
- Each of the grid cells also predicts several conditional class probabilities. Only one set of class probabilities is produced per grid cell regardless of the number of bounding boxes.
- When the test occurs the conditional class probabilities and the individual box confidence predictions are multiplied, which produces a class -specific confidence for each of the classes appearing in the box and how well the box is situated around the object.

Figure 3 illustrates sunspot detection precision from the YOLO model on an image where a sunspot is just going around the solar limb

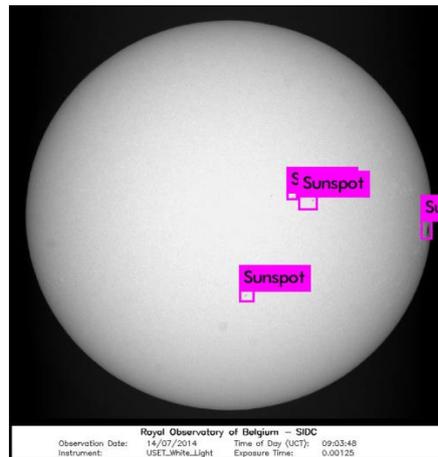


Figure 3. YOLO Sunspot detection

2.5 SSD

The Single Shot Detector was trained using TensorFlow 1.5 because TensorFlow 2 does not have a SSD available yet, and also Detectron 2 where mobile friendly versions were also trained to establish if lighter smaller models meant for less powerful hardware could detect sunspots. Due to the Single Shot Detector being on TensorFlow 1.5 older versions of CUDA and cuDNN had to be used alongside older version of NVIDIA drivers. The different architectures and versions of the SSD were trained.

Overview of SSD training process:

- The input image is used to predict category scores and box offsets for a fixed set of default bounding boxes that are of different scales and aspect ratios
- A set of default anchors then aid in helping to discretize the output space with the various scales of and aspect ratios of the anchor boxes
- These various predictions of the different feature maps are then fused
- This is followed by a non-max suppression step to produce the final output decisions, whereas the early layers are based on high quality classifier networks for good feature extraction

Figure 4 illustrates sunspot detections on a Single Shot Detector using the same image as in 2.2.

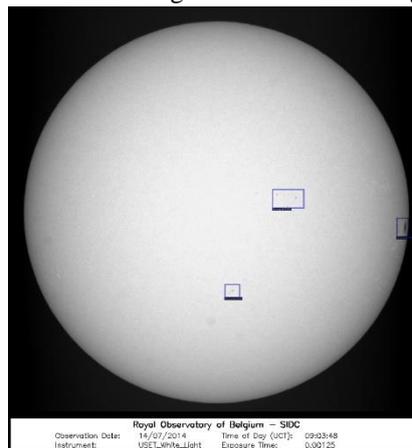


Figure 4 SSD sunspot detection

2.6 EfficientDet

EfficientDet was trained using PyTorch and TensorFlow 2, the model has only a single version and does not have different backbones, but the different versions of the model are the same architecture just with different scales applied across the backbone, image resolution and feature fusion neck. The models go from d0 to d7 getting bigger as you go up from d0. Only d0 to d3 were trained as the model is very memory intensive and requires multi-graphics processing unit set ups to train all the way to d7.

Overview of EfficientDet training process:

- Extract irregular features from the backbone network.
- Fusion occurs in a bidirectional neural network where features are fused for each layer in the backbone using bidirectional cross scale connections repeatedly and fast normalized fusion
- Perform class and box prediction on fused features with a box regression and classification layer.

Figure 5 illustrates detection of sunspots on a group of sunspots and where some sunspots are just moving over the solar limb.

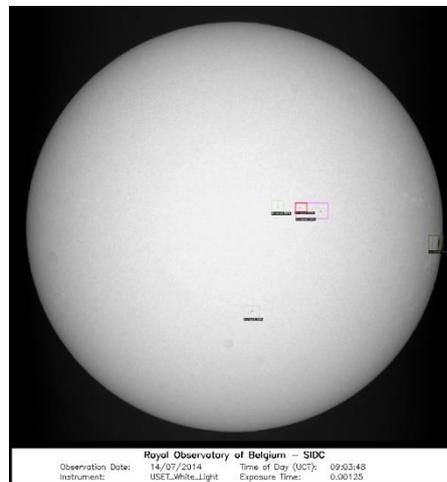


Figure 5. EfficientDet sunspot detection

3. RESULTS

These models were trained and tested using different architectures, different backbones, hyper parameters and frameworks. Models were also trained with and without the augmented data set to establish if the augmentations and enhancements were improving average precision. The results will now be discussed.

3.1 Faster RCNN

The average precision achieved by the Faster RCNN model varied between all the different backbones and hyper parameters. The best performing model achieved an 82.5 % mean average precision, whereas the rest of the Faster RCNN models varied between 79 – 81. The best performing model did not detect any false positives and was very good at picking up sunspots around the solar limbs that were either going out of sight or coming into sight. Sunspots appear in groups also and all the Faster RCNN models were able to contextually learn this, and none of the sunspot pair were seen as two individual sunspots. The Faster RCNN model is very good at sunspot detection and proved very accurate also.

3.2 YOLO

The YOLO models proved that single stage detection techniques can keep up in terms of accuracy with two stage detectors and the best model achieved an 81% accuracy, where no false positives were picked up and the model was very good at detecting sunspots around the solar limbs. However, the yolo models struggled to see a sunspot pair and sunspot group as a single sunspot. This problem was present on all YOLO models except the best performing one and a version which uses a transformer backbone. The YOLO models, however, are very efficient at detection and, are very fast when processing an image for detection.

3.3 SSD

The SSD models performed worse when the backbone network is small, and although this works well in scenarios where objects are large in images, sunspots vary from very large to very small. This is due to the way the SSD approach uses fixed anchors at every layer of the backbone and when the network is small the features are limited. However, when the SSD is used with a large and deep backbone network, the precision increases as the anchor boxes become more. The best mean average precision achieved was 79.6% and the model had no false positives and only one sunspot that was missed. This can be alleviated with more data as the sunspot missed was a very large sunspot and these were very few in the data set.

3.4 EfficientDet

The EfficientDet proved very difficult to train on a single graphics processing unit as the model is very memory intensive. Only d0 – d3 models were trained and between them no meaningful difference in mean average precision was apparent. The best accuracy achieved was 62%, and although the model did not have any false positives, there were overlapping bounding boxes where the model sees a sunspot group and each individual sunspot shown in figure 5. The model did also miss very small sunspots around the solar limbs, but the model has potential and is very efficient at processing an image for detection.

3.5 Results

Model	Backbone	AP %
EfficientDet d2	Bidirectional feature pyramid (Efficient-b2)	62
Faster-RCNN	ResNet-50 conv4 with conv5 as head	79
Faster-RCNN	ResNet-50 conv4 with conv5 as head	80
Faster-RCNN	ResNet-50 conv5 with conv5 dilations	81.6
Faster-RCNN	ResNet -50 and Feature pyramid	80
Faster-RCNN	ResNet-101 conv5 with conv5 dilations	81
Faster-RCNN	ResNext-101 and Feature Pyramid	82.5
Faster-RCNN	Resnet-152 v1	79.4
Faster-RCNN	Inception Resnet v2	80.4
SSD Retinanet	ResNet -50 and Feature pyramid	78.4
SSD Retinanet	ResNet -101 and Feature pyramid	79.6
Yolo v5	CSPDarknet-53	76.8
YoloS	Vision Transformer block	54.2
Yolo v4	CSPDarknet-53	79
Yolo v6	EfficientRep	75
Yolo v7	CSPDarknet-53 With new E-ELAN computation block	81
YoloX	DarkNet-53	80.6

4. CONCLUSION

The Faster RCNN model proved the most accurate at detecting sunspots on any area of the sun and did not have any false positives. Although the other models performed well and did also show very good accuracy the EfficientDet requires more processing power to be trained effectively. The two-stage detection method provides more finer grain detail to do detection as opposed to the single stage detectors which were designed for computational efficiency. The YOLO model showed very good performance on the best model and on the transformer backbone model, indicating that the newer changes to the architectures might be aiding in catching up to the fine grain details of the Faster RCNN backbones. The other model which shows much promise is the SSD as, the best performing model achieved great results, and when considering the SSD was designed for lower power hardware, it might be a good choice to run on satellites or ground based observatories. These models can also be enhanced at pore detection and on other image formats for other solar phenomena. Future work that can be derived from this study includes:

1. Sunspot classification
2. Sunspot real time counting
3. Continuous retraining as detection occurs
4. Space weather forecasting
5. Solar event prediction

REFERENCES

- Channabasava, C and Biabl, B, 2022. Detection and classification of sunspots via deep convolutional neural network. *Global Transitions Proceedings*, Vol. 3, No. 1, pp 177-182.
- Fang, Y *et al*, 2019. Deep Learning for automatic recognition of magnetic type in sunspot groups. *Advances in Astronomy*, Vol. 2019, pp 1-10.
- Hapgood, M, 2017. *Space Weather*. IOP Publishing, Bristol, UK.
- Hanslmeier, A, 2004. *The Sun and Space Weather*. Kluwer Academic Publishers, New York, USA.
- Solanki, S, 2003. Sunspots: An overview. *In The Astronomy and Astrophysics Review*, Vol. 11, No. 2, pp 153-286.
- Wilkinson, J, 2012. *New Eyes on the Sun*. Springer, New York, USA.
- Zhao, Z *et al*, 2019. Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, Vol. p, No. 2, pp 1-21.
- Zhou, Z *et al*, 2020. Computer-aided detection of brain metastases in t1-weighted mri for stereotactic radiosurgery using deep learning. *Radiology*. Vol. 295, No. 2, pp 407-415.