






RESEARCH ARTICLE

10.1029/2021SW002976

Attention-Based Machine Vision Models and Techniques for Solar Wind Speed Forecasting Using Solar EUV Images

Special Section:

Heliophysics and Space Weather Studies from the Sun-Earth Lagrange Points

Edward J. E. Brown^{1,2,3} , Filip Svoboda¹, Nigel P. Meredith² , Nicholas Lane^{1,4}, and Richard B. Horne² 

¹Department of Computer Science and Technology, University of Cambridge, Cambridge, UK, ²Space Weather and Atmosphere Team, British Antarctic Survey, NERC, Cambridge, UK, ³BAS AI Lab, British Antarctic Survey, NERC, Cambridge, UK, ⁴Samsung AI Center, Cambridge, UK

Edward J. E. Brown and Filip Svoboda contributed equally to this work.

Key Points:

- Attention-based machine vision models and methodological enhancements are developed to improve solar wind speed forecasts from solar images
- Attention-based architectures outperform convolutional models, motivating their use in future studies and production systems
- The models perform best when solar activity is driven by coronal holes, such as in the declining phase of the solar cycle

Correspondence to:

E. J. E. Brown,
ejeb4@cam.ac.uk

Citation:

Brown, E. J. E., Svoboda, F., Meredith, N. P., Lane, N., & Horne, R. B. (2022). Attention-based machine vision models and techniques for solar wind speed forecasting using solar EUV images. *Space Weather*, 20, e2021SW002976. <https://doi.org/10.1029/2021SW002976>

Received 1 NOV 2021

Accepted 27 JAN 2022

Author Contributions:

Conceptualization: Edward J. E. Brown, Filip Svoboda

Data curation: Edward J. E. Brown, Filip Svoboda

Formal analysis: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith

Investigation: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith

Methodology: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract Extreme ultraviolet images taken by the Atmospheric Imaging Assembly on board the Solar Dynamics Observatory make it possible to use deep vision techniques to forecast solar wind speed—a difficult, high-impact, and unsolved problem. At a 4 day time horizon, this study uses attention-based models and a set of methodological improvements to deliver an 11.1% lower RMSE and a 17.4% higher prediction correlation compared to the previous work testing on the period from 2010 to 2018. Our analysis shows that attention-based models combined with our pipeline consistently outperform convolutional alternatives. Our study shows a large performance improvement by using a 30 min as opposed to a daily sampling frequency. Our model has learned relationships between coronal holes' characteristics and the speed of their associated high-speed streams, agreeing with empirical results. Our study finds a strong dependence of our best model on the phase of the solar cycle, with the best performance occurring in the declining phase.

Plain Language Summary Solar images contain rich information that can be used to forecast conditions at Earth. This study develops a robust methodology for processing solar images and trains machine learning models that can use them to predict the solar wind speed. Combined, these deliver a very significant 17.4% improvement in the correlation between the prediction and the ground truth over previous works. The models perform better during the quieter, declining phase of the solar cycle when the solar activity is driven by coronal holes. Finally, the trained models learn properties of coronal holes that agree with prior empirical studies.

1. Introduction

The solar wind is a stream of charged particles that is emitted from the upper atmosphere of the Sun. The speed, density, temperature and the magnitude and direction of the associated magnetic field of the solar wind are constantly varying affecting the way in which it ultimately interacts with the Earth's magnetosphere. High-speed solar wind streams (HSS) emanating from coronal holes are particularly effective at coupling with the Earth's magnetosphere. The weak storms they produce tend to have long-lasting recovery phases which often result in prolonged and enhanced substorm activity (Meredith et al., 2011; Tsurutani et al., 1995). This results in repeated injections of suprathermal electrons into the inner magnetosphere and significant increases in the fluxes of relativistic electrons in the outer radiation belt, increasing the risk to satellites via surface charging and internal charging respectively (e.g., Borovsky & Denton, 2006). Indeed, it has been suggested that satellites at geostationary orbit are more likely to be at risk from an extreme HSS-driven storm than a Carrington type event (Horne et al., 2018). Furthermore, prolonged and enhanced substorm activity associated with HSS-driven storms results in increased thermospheric densities and satellite drag (Chen et al., 2012). Consequently, accurately forecasting the solar wind speed associated with coronal holes is very important for our modern society.

Coronal holes are large dark areas on the Sun as seen in extreme ultraviolet (EUV) and soft X-ray images (Cranmer, 2009). They are regions of open magnetic field and cooler plasma, leading to the production of high-speed solar wind streams. Coronal holes are long-lasting features that can persist from one solar rotation to the next, giving rise to a 27 day periodicity in the arrival of HSS at Earth. The occurrence rate of coronal holes peaks during the declining phase of the solar cycle (Burlaga & Lepping, 1977) and high-speed streams observed at Earth during these intervals tend to be coronal-hole driven. The distribution of speeds in high-speed streams associated with coronal holes ranges from 400 to 800 kms^{-1} (Kilpua et al., 2017). While these streams do not

Resources: Edward J. E. Brown, Filip Svoboda
Software: Edward J. E. Brown
Supervision: Nigel P. Meredith, Nicholas Lane, Richard B. Horne
Validation: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith, Nicholas Lane
Visualization: Edward J. E. Brown, Filip Svoboda
Writing – original draft: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith
Writing – review & editing: Edward J. E. Brown, Filip Svoboda, Nigel P. Meredith, Nicholas Lane, Richard B. Horne

result in major geomagnetic storms (Richardson et al., 2006), they have extensive recovery phases, typically lasting from 5 to 10 days, and, as a result, may deposit more energy in the magnetosphere than larger storms (Kozyra et al., 2006; Turner et al., 2006).

Coronal holes are not the only source of high-speed solar wind at Earth. Coronal mass ejections (CMEs) also cause high-speed solar wind, although not all CMEs are associated with high solar wind speeds (Kilpua et al., 2017). CMEs are large explosions on the Sun that hurl vast amounts of plasma into space. The occurrence rate of CMEs peaks at solar maximum (St. Cyr et al., 2000) so that most periods of high solar wind speed observed during these periods tend to be CME-driven. The distribution of speeds in interplanetary CMEs and sheath regions associated with CMEs on the Sun ranges from 250 to 950 km s⁻¹ (Kilpua et al., 2017). Unlike coronal holes, CMEs are not associated with long lasting features on the Sun. In contrast, they are best observed in coronagraph images where they appear as expanding shells of material.

In this study, we build a machine learning model to use solar images to forecast the solar wind speed at Earth. This technique is expected to perform best when there are associated visible features on the Sun. The method is thus expected to work well for coronal holes, which are large features on the solar disk. In contrast, CMEs are barely noticeable within EUV images and so the ML model would not be expected to work well for these events.

The field of machine learning has built a lot of momentum over the last 10 years. This has largely been the result of improvements in algorithmic capability, availability of data, funding and hardware. Not to be overlooked though is the creation of field benchmarks like ImageNet (Deng et al., 2009) and open-source software such as PyTorch (Paszke et al., 2019) which dramatically shortened the development cycle in the field and greatly increased its standardization.

Deep (Machine) Learning excels where rich data exists in large quantities, because models with deep structures and therefore many parameters need to consume richly varied data sources to build complex internal representations of the data generating system. This is the essence of deep learning. Recently, curated solar image datasets have been created such as the SDOML data set (Galvez et al., 2019) which contains images of the Sun taken at various EUV wavelengths. These data allow the rapid application of machine learning algorithms to consume solar images.

In this paper, we use the EUV images taken by the Solar Dynamics Observatory (SDO) using the Atmospheric Image Assembly (AIA; Lemen et al., 2011) to forecast the solar wind speed at the Lagrangian L1 point. We present results for forecasting at a 4 day lag from a single 211 Å image—but this forecast could be used for any lag up to 4 days. We also explore the model's learned behavior by examining relationships between the peak solar wind speed and the coronal hole area and intensity. Previous works and the datasets are presented in Sections 2 and 3, respectively. In Section 4, we discuss our general methodology and model architectures. Our results are presented and discussed in 5. Finally, our conclusions are summarized in Section 6.

2. Previous Works

The works of Wintoft and Lundstedt (1997) and Wintoft and Lundstedt (1999) were the first to use neural networks to forecast the solar wind speed. These are small, so-called fully connected, models that could learn nonlinear relationships between a limited set of pre-computed feature inputs, such as the flux tube expansion factor, and the solar wind speed. More recently, similar studies were performed by D. D. Liu et al. (2011), Yang et al. (2018), Chandorkar et al. (2019), and Bailey et al. (2021) using similar non-image-based inputs to the models, albeit with more advanced models than the earlier works.

Upendran et al. (2020) was the first study aiming to forecast solar wind speed from solar EUV images using deep learning techniques. The work uses images from both 193 and 211 Å wavelengths to forecast the solar wind speed at a 1 day resolution. Upendran uses GoogleNet (Szegedy et al., 2014), trained on the ImageNet data set (Deng et al., 2009), as a feature extractor for each image. The extracted per-image features are then passed into an LSTM Recurrent Neural Network (Hochreiter & Schmidhuber, 1997) to produce the predicted solar wind speed. The study achieves a best performing model at a lag of 3 days and a history of 4 days, with a correlation of 0.55 and an RMSE 80.28 km/s. This study will build on this insightful initial work.

Next, Raju and Das (2021) proposed a smaller three-layer convolutional feature extractor, which they train on the 193 Å wavelength solar EUV images. Their method targets a subtly different task than that of Upendran et al. (2020). While Upendran et al. (2020) use present solar images to forecast future solar wind speeds at a fixed lag in the future, Raju and Das (2021) backcast current solar wind speed based on flexible-lag past images. Specifically, Raju and Das (2021) use the current solar wind speed to infer which past image was likely to have caused the recorded solar wind speed, and then pass this image into their model with the expectation that the model will be able to correctly reconstruct the observed solar wind speed. The difference becomes clearer when the models are to be deployed as live solar wind speed predictors. Under the forecasting setup, today's images can be used to produce the predicted solar wind speed 4 days from now. In contrast, under the backcasting setup, the inference process by which images are paired with time stamps does not guarantee a unique prediction for each time stamp, and so some future time stamps can be expected to receive multiple solar wind speed predictions, while others would get none. Thus, this model is not comparable to Upendran et al. (2020). Nevertheless, they provide results for a model specially trained at a fixed 4 day forecast horizon (their Table 4), with the year 2018 held out as a test set. They report 78.3 km/s RMSE and a prediction correlation of 0.55. This would be comparable to Upendran et al. (2020), except they provide no results for 2018 alone. Their test results are from across multiple years. Therefore, our study will compare to Upendran et al. (2020) for dates across an 8.5 years range and then run a separate training run to compare to Raju and Das's (2021) fixed 4-day model, just evaluating on the year 2018.

3. Data

3.1. Solar Images

The image data set consists of EUV images from NASA's SDO taken by the AIA (Lemen et al., 2011) that have been processed by performing various instrumental corrections, downsampled to useable spatial and temporal resolutions and synchronized both spatially and temporally to form the SDOML data set (Galvez et al., 2019). The resulting data set contains 8 and a half years of images every 6 min from May 2010 to December 2018. These images are monochromatic and the pixel values represent the intensity of light. This study uses the EUV images at 211 Å.

3.2. Solar Wind Speed

The solar wind speed data are taken from the OMNIWeb service. Specifically, we use the solar wind speed, measured in km/s, at a 1 min time resolution for the period of the SDOML data set. The data comes from WIND and the Advanced Composition Explorer spacecraft, both positioned at the L1 point, about 1.5 million km from Earth.

The solar wind speed is highly auto-correlated with itself over hourly time periods and is still at 0.7 after 1 day. By 4 days, the correlation has dropped to negligible amounts. Notably, at 27 days, there is a spike in the auto-correlation. This is because the Sun has a synodic rotation period of approximately 27 days and some longer lasting features, such as coronal holes, come around again causing similar solar wind speed conditions at L1. This auto-correlation is important since it has implications for which images are included in training and test sets due to their dependence on each-other. This is further discussed in Section 4.1.7.

4. Methodology

4.1. Methodological Improvements

Here, we discuss changes in our methodology to the only previous work (Upendran et al., 2020), covering all the date ranges available from the SDOML data set.

4.1.1. Image Pre-Processing

The EUV images at their provided resolution are too large to practically process on standard computing hardware. Previous works elected to down-sample the full 512 by 512 pixel image to 224 by 224 by max pooling. Instead, we take a 300 by 300 pixel square whose corners are approximately at the edges of the solar disk, and then down sample this cropped image to the desired 224 by 224 image size. This results in lower loss of information content in the relevant section of the Sun because (a) the cropped solar poles are unlikely to contain features

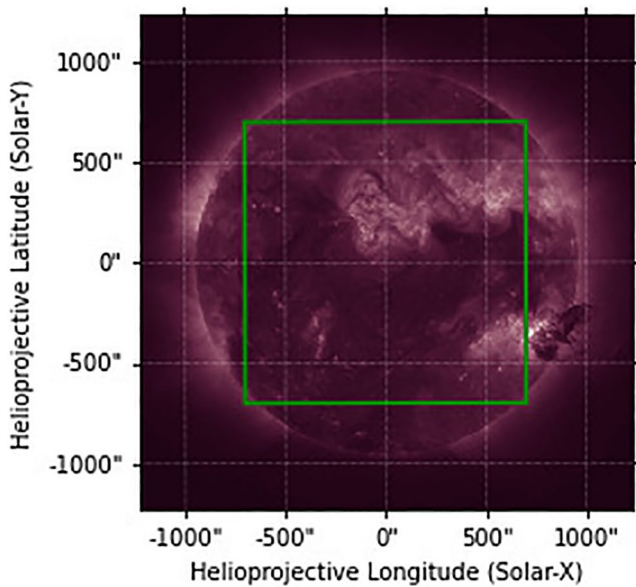


Figure 1. Solar Dynamics Observatory Atmospheric Image Assembly 211 Å image taken at 7 June 2011 06:33:02 UT (Lemen et al., 2011). The green line represents the crop taken.

that affect the solar wind speed at L1, (b) the cropped features at the eastern limb have not yet had time to rotate more centrally and become relevant and the response from the western limb has come and gone, (c) this allowed us to down-sample the central, relevant, portion of the image less aggressively. Figure 1 shows an example of our cropping technique.

Regarding scaling the cropped image images, the same method as used in Upendran et al. (2020) is employed by clipping the pixels to have values between a minimum of 25 and a maximum of 2,500 and taking the natural logarithm. However, after this we rely on a batchnorm layer to learn an optimal scaling, as opposed to fixing it (further detailed in Section 4.2).

4.1.2. Sampling Frequency

We replace the previously used daily sampling resolution with a 30 min schedule, because solar wind speeds can change significantly even on a 30 min time scale.

4.1.3. Carrington Rotation

The Sun rotates on average every 27.28 days as viewed from Earth, this is one Carrington rotation (Ridpath, 2012). As such, the solar features that affected the solar wind speed at a given point come back approximately 27 days later and produce similar effects. Thus, the solar wind speed is also auto-correlated at the Carrington rotation periodicity with a value of 0.42 at 27 days. As this value is available to all forecasters operating at lower than 27 days forecast horizon, it should be used as an input to our models.

4.1.4. North-South Augmentation

We augment the data set by randomly flipping the training images north to south, as features, such as coronal holes, produce a similar increase in solar wind speed regardless of which side of the solar equator they are on. Although it is not claimed these are valid physical suns.

4.1.5. Single Image Versus Sequence

The previous work relies on a convolutional feature extractor pre-trained on ImageNet in combination with an LSTM cell and a fully connected layer (Upendran et al., 2020). Up to four images were sequentially passed through the convolutions. Separate for each image, the model's activations at multiple layers were extracted, concatenated, and passed into the LSTM as individual time steps. The convolutions remained parametrized by the weights obtained on ImageNet and only the other layers' parameters were trained. The high auto-correlation of solar images is likely to, again, exaggerate the model's multi-collinearity in hidden features while providing little additional context. Thus, we replaced the LSTM feeding into a fully connected output layer with two consecutive fully connected layers.

4.1.6. Feature Extractor Re-Training

This study will use pre-trained vision models at the core of the model architecture (see Section 4.2 for more details). Rather than to use the fixed pre-trained ImageNet weights, the model will be initialized with these weights but they will not be fixed. This we believe to be strictly necessary due to the wide gap between the EUV and the ImageNet datasets.

4.1.7. Training, Validation, and Test Sets

For this study, fivefold cross-validation is employed to evaluate the models. Solar wind speed is auto-correlated up to a period of about 4 days. For the period of June 2010 to December 2018, the auto-correlation is as high as 0.70 at 1 day. This means that if timestamps are too close to each-other between training, validation and test sets, it is not a fair reflection of the performance of a model, since the Sun has not changed much in for example, 30 min. Furthermore, this will mean that the model overfits on the validation sets, meaning they will not generalize as well. To create more independent training and test sets, a method similar to that used in Upendran et al. (2020) is employed whereby the timestamps from 2010 to 2018 are split into chunks of 20 days. However,

a buffer period of 4 days between each chunk is discarded to ensure the independence of the training, validation and test sets. It is noted that this throws out approximately one fifth of all the data. However, this is justified to ensure the independence of datasets while also covering as many parts of the solar cycle as possible. Appropriating the chunks into train, validation and test buckets is not a random shuffle of the data, but it follows a cyclic pattern. The first three chunks are put in the train set, the fourth in the validation set, and the fifth in the test set. This pattern is then repeated until no chunks are left to create the first fold. This pattern is then cyclically permuted to produce each fold. This means each chunk serves its turn in the test set in one of the five folds. For each fold, a model is trained on the training set and evaluated on the validation set for 100 epochs (1 epoch is a full pass over the training data). The model is saved every epoch. The version of the model that performs best on the validation set is the final model. This final model is then applied and evaluated on the unseen test set. Figure 2a shows the training sets in orange, the validation sets in blue and the test sets in yellow. White buffer sets of 4 days are included between the 20 days chunks. Chunking the data as in Figure 2a, results in 124 20-day chunks of data. This scheme results in fivefold of approximately 64,000, 21,000, and 21,000 data points for the train, validation and test sets respectively. These respectively approximate to 1300, 440, and 440 days worth of data. The reason this is lower than 8.5 years (May 2010 to December 2018) is due to both the removed buffer data as well as missing data in the underlying data set. The reported RMSE and Correlation is averaged over the five folds and reported.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}$$

$$\text{Correlation} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

where y_i is the real solar wind speed, x_i is the predicted solar wind speed, \bar{y}_i is the mean real speed, \bar{x}_i is the mean predicted speed, and n is the total number of data points.

4.2. Model Architectures

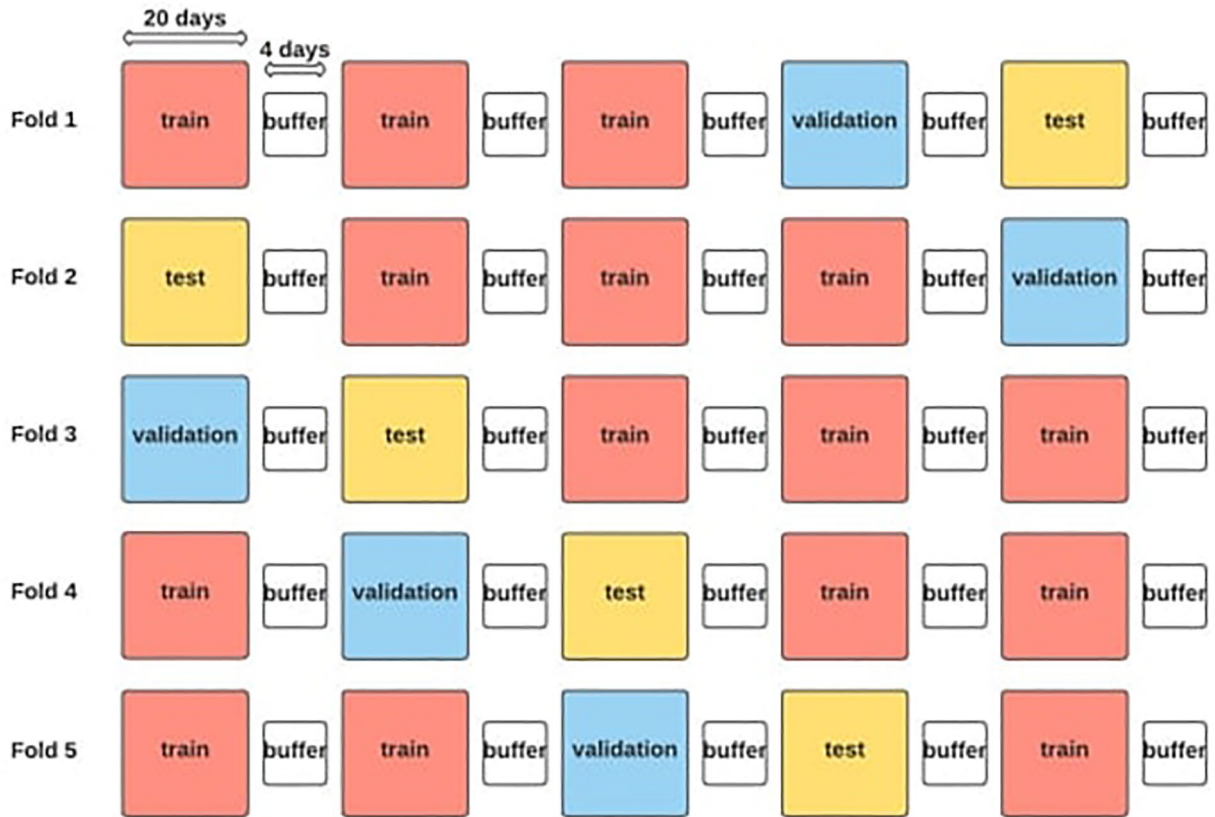
For this study, the architectures for the different models will follow the format in Figure 3. The image will pass through a batch norm layer that will rescale it. Then it is passed into the candidate architecture, be it a CNN or a vision transformer. The outputs from this model as well as the solar wind speed from one Carrington rotation ago are then passed into two final consecutive nonlinear projections that produce the model's solar wind speed prediction.

In all cases, the models are trained in their entirety on the EUV data. That is, after their parameters are initialized using either random, or when available, pre-set weights the algorithm iteratively updates them with the goal of incrementally decreasing the mean squared error of its prediction.

4.2.1. Benchmark CNN-Based Models

In general, every deep model can be seen as a layered composition of nonlinear projections, each forming a separate layer. Model inputs, solar images in our case, can be seen as the zero-th layer, while, model outputs, the predicted solar wind speed, can be treated as the last layer. Each layer in between is a nonlinear projection that receives inputs from the preceding layer, and that outputs its value to the next layer. Commonly, several layers are grouped into modules and used as a type of meta-layer. Modern architectures are defined by the features that build on and expand this basic structure.

Previous work used convolutional models in the forecasting of solar wind (Raju & Das, 2021; Upendran et al., 2020). These models are designed to process images, each of which has three dimensions—the height, the width, and the number of channels. A standard color image has three channels: red, green, and blue. Convolutions are operations that split the image into a grid of patches and then use a three dimensional kernel to compute weighted averages per each patch. The same kernel is used on each patch and the averages it produces become the pixel values the layer outputs. Multiple kernels may be employed, in which case their outputs are treated as separate channels of the outputted image.



(a) 5-fold cross validation with buffer data thrown out. Pattern is repeated across the May 2010 to December 2018 range.



(b) Dataset split with 2018 as hold-out test set for comparison with Raju and Das (2021)

Figure 2. Training, validation, and test sets. (a) Fivefold cross validation with buffer data thrown out. Pattern is repeated across the May 2010 to December 2018 range. (b) Data set split with 2018 as hold-out test set for comparison with Raju and Das (2021).

GoogleNet, also known as InceptionNet v1, is the convolutional architecture at the heart of Upendran et al.'s (2020) work. It is a convolutional architecture that replaces layers with modules. Each module computes several, rather than just one convolution. These are computed in parallel, and are meant to complement each other. The desired effect is to make the model's computation more parallelizable, thus faster, while improving the model's ability to fit complex patterns in the data (Szegedy et al., 2014).

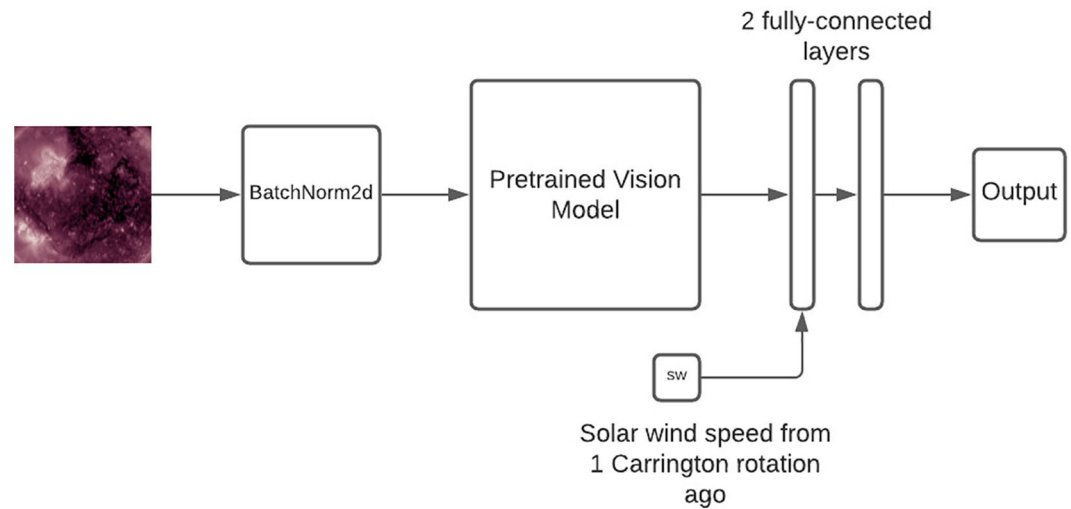


Figure 3. The solar model architecture.

InceptionNet v2 is a second generation and a refinement of the GoogleNet. The architecture builds on GoogleNet's inception modules by decomposing their convolutions serially. Specifically, more computationally expensive, that is larger-kernel convolutions, are replaced by a series of much cheaper smaller-kernel convolutions carried out one after the other. The desired effect is to make the working set of this algorithm smaller, while further improving the model's capacity, that is, its ability to fit complex data patterns (Szegedy et al., 2016).

ResNet is a predecessor of GoogleNet. ResNet's modules consist of two consecutive convolutions, and a so-called residual connection. The residual connection is a bypass that circumvents the two convolutions. In effect, this results in a block that outputs both its convolution's output as well as the original inputs to the block. This trick helps to propagate the training gradients through the network, mitigating the vanishing gradient problem. The architecture was the first one to breach the 20 layer depth ceiling (He et al., 2016).

DenseNet is a generalization of ResNet that adds multiple residual connections to each module. The beginning of a block of convolutions, is connected not only to the output of that same module, but also to the outputs of all modules down-stream from it (Huang et al., 2017).

4.2.2. Attention-Based Models

This paper proposes using attention, rather than convolution, as the core model feature. Attention is a deep learning mechanic that, rather than learn a weight per each input pixel or a patch of pixels, learns a method for generating these weights from the input data. Consequently, the models can weight each patch based on what its position is and what the rest of the image depicts (Vaswani et al., 2017). In contrast, convolutions are designed to analyze each patch of each input image using the same kernel of weights, regardless of what the image depicts outside of the patch and what its position is. Formally, convolutions enforce translation invariance, while attention models do not. Translation invariance in computer vision is achieved when the model maintains the same output even if the objects in the image are moved around.

Attention's ability to judge each image patch in the context of its position in the image and the contents of the rest of the image is critical for making sound solar wind speed predictions from the EUV data. First, the attention mechanism allows the model to assign higher importance to features on the Sun's surface if they appear in the equatorial region. Moreover, the model is able to learn to distinguish between situations when an active region interferes with a coronal hole, and when it does not. The weights it places on the patches of the image with the coronal hole in it will depend not only on its position in the image, but also on whether the model identified an interference from an active region. In contrast, convolution-based models were designed to identify an object anywhere in the input image field. Therefore, they place equal weight on each image patch as they process it using the same fixed-weight convolution kernel. It was assumed that multiple layers of convolutions would learn increasingly complex representations by deriving higher-layer features from simple lower-layer ones. Recently, however, it was shown that convolutional models do not recognize complex features, instead they aggregate

low-level texture features from across the input image and then make their prediction based on which texture prevails in the input image (Geirhos et al., 2018). Consequently, attention-based models will make better and more theory-sensible predictions as it, for example, will account for and internalize the higher importance of features in the equatorial region and the interference of active regions with coronal holes while convolution will fail to do so.

The *Vision Transformer* was the first transformer architecture successfully used in image recognition (Dosovitskiy et al., 2020). The architecture combines large image patches with the attention mechanism. Each patch is first individually passed through a linear projection, then, the attention mechanism applies context-derived weights on each. The result is then passed into two consecutive nonlinear projections, sometimes called fully connected layers, before being outputted. An important point of comparison is the size of the model's patches. While all benchmark models only consider patches of no more than 5×5 pixels, our Vision Transformer works with patches of 16×16 . This is meant to allow it a larger receptive field and to steer clear of focusing on textures.

The *Transformer in Transformer* follows the same general architecture as the original Vision Transformer, the crucial difference is that the linear projection at the beginning of the outer transformer is replaced by an inner transformer that is modeled as a smaller version of the same original Vision Transformer (Han et al., 2021). Therefore, the input image is first split into 16 by 16 patches. Each of these patches is then passed into the inner Vision Transformer, as if they were images in their own right. This splits them into smaller (4×4) patches still, derives the attention weight for each sub-patch based on the rest of each patch, and outputs the processed image back to the outer transformer. The outer transformer then uses these processed patches to derive its attention weights per each patch based on what the rest of the full image's processed patches are like. Then the outer transformer uses two consecutive nonlinear projections to produce the final output.

The *Swin Transformer* is similar to the Vision Transformer except it builds hierarchical feature maps by merging image patches, as opposed to treating image patches separately as in the Vision Transformer (Z. Liu et al., 2021). The idea is that the model is able to treat features on different scales, whereas the vanilla vision transformer is limiting itself to a predetermined patch size. Furthermore, a feature of the algorithmic construction is a linear scale in computational complexity based on image size.

These pre-trained attention-based models, as well as the benchmark CNN models, all accept three-channel RGB images normally. In order to use these powerful models, the solar images have to be repeated three times to form the three channels. Normally, one would use the advised normalization schedule from the papers that produced these models. In this case, however, since the models are not RGB in the first case, it was decided that an initial batch norm layer is applied before the model, so that the best normalization schedule can be learned and not fixed.

4.3. Other Experimental Details

4.3.1. Missing Data

Missing images are substituted with valid observations no more than 30 min removed from the missing datum. Missing solar wind speed data are interpolated from available data but if there is no data within 30 min of a timestamp, that timestamp is thrown out. The remaining points of time, which both have a speed after interpolation and an image after we have looked for a suitable replacement image if missing, are used as the datapoints for the model.

4.3.2. Hyper-Parameter Selection

Hyper-parameters are chosen using a Bayesian parameter sweep using the software *Weights and Biases* (Biewald, 2020) based on the performance of the validation set. For cost reasons, the sweep is conducted at 120 min resolution for only 30 epochs.

4.3.3. Training Process

The loss function of the network is the default implementation of PyTorch's mean squared error (squared L2 norm; Paszke et al., 2019). The optimizer method to update the weights of the network is the default implementation of the Adam optimizer in PyTorch as well (Kingma & Ba, 2014). Batch size is fixed at 64.

Table 1

Performance of Our Solar Models Compared to Upendran et al. (2020) Forecasting Solar Wind Speed Using the Extreme Ultraviolet Data at a 4 day Forecast Horizon in the Period May 2010 to December 2018

	RMSE	% Improvement	Correlation	% Improvement
Model				
Persistence (4 day)	127.59	-57.1%	0.080	-85.2%
Persistence (27 day)	100.86	-24.2%	0.426	-21.1%
Upendran et al. (2020)	81.21	-	0.54	-
Our models				
Solar InceptionNet v4	74.09	8.8%	0.609	12.7%
Solar DenseNet	73.92	9.0%	0.611	13.1%
Solar GoogleNet	73.71	9.2%	0.619	14.6%
Solar ResNet	73.52	9.5%	0.618	14.4%
Solar TNT	72.70	10.5%	0.629	16.5%
Solar Vision Transformer	72.66	10.5%	0.630	16.7%
Solar Swin Transformer	72.21	11.1%	0.634	17.4%

4.3.4. Computation

All experiments were run on V100 Nvidia GPU, resulting in a total compute of about 900 GPU hr.

4.4. Year 2018 Evaluation

Solar activity can vary significantly based on position in the solar cycle, so only testing on 2018 only gives the performance of the model in that part of the solar cycle. It therefore cannot be representative of the generalization of the model to other periods of the solar cycle. However, Raju and Das (2021) provide results for a model trained on solar imaging data with the entire year of 2018 held out for evaluation. As an extra experiment and to compare to their study, a model will be trained with the training and test set schedule shown in Figure 2b. Notably, Figure 2b features a 27 day test buffer before the start of the 2018 test set. This buffer is present because of Raju and Das' concern of 27 day resurgence causing the training and test sets to not be independent. Our view is that since this model is forecasting at a 4 day forecast, any image before that 4 days could be used to train a model in a production system to make that 4 day forecast (especially using the method of online learning). Despite the dependence, this 27-day old image would be one of the most important images you would want to train on. Where the dependence matters for forecasting purposes is crucially when the images are less than the forecast horizon apart. This explains our choice of 4 day buffer otherwise. However, for the point of comparison, this 27 day buffer is kept. Otherwise, all experimental procedures as detailed will remain the same as with the fivefold split.

5. Results and Discussion

5.1. Comparison to Previous Works

Table 1 shows the comparison of our methodological and modeling pipeline, used with a range of feature extractors, against the most recent state of the art forecasting model in the field and two naive persistence model benchmarks. Notably, all of the models trained under our pipeline improve on the work by Upendran et al. (2020) by at least 8.8% in RMSE and 12.7% in correlation. Indeed, our pipeline with the GoogleNet feature extractor, which is the same feature extractor as was used in the Upendran et al. (2020) model, demonstrated the total improvement our pipeline has delivered. It lowered the RMSE by 9.2% and increased the correlation by 14.6%. Furthermore, our best performing model, based off the Swin Vision Transformer, improves on the state of the art by 11.1% in RMSE and 17.4% in correlation. The model also outperforms at the 1, 2, and 3 day time horizon because the

Table 2

Performance of Our Solar Models Relative to (Raju & Das, 2021) Predicting Solar Wind Speed Using Extreme Ultraviolet Data at a 4 Day Forecast Horizon in for the Year 2018

	RMSE	% Improvement	Correlation	% Improvement
Model				
Persistence (4 day)	118.76	−52.3%	−0.027	−104.9%
Persistence (27 day)	85.16	−9.2%	0.464	−15.6%
Raju and Das (2021)	78	–	0.55	–
Our model				
Solar Swin Transformer	71.65	8.3%	0.644	17.1%

4 day forecast could also be used for those. Finally, transformer feature extractors outperformed convolutional ones by about 1%–2% in either metric when used in our model pipeline.

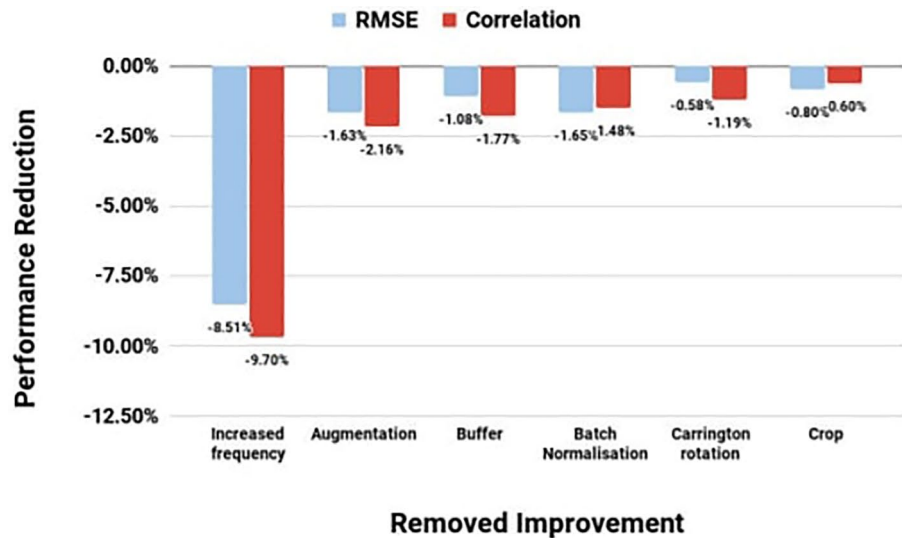
Table 2 compares the performance of our best performing model, that is the one based on the Swin Transformer feature extractor, and the two persistence benchmarks against the predictions Raju and Das (2021) produced for the year 2018. This setup differs from that of Table 1, in that table tests the models on data examples sampled from the whole data set, and thus across the solar cycle. The present comparison is made solely with respect to the solar cycle conditions present in the year 2018, as chosen by Raju and Das (2021). Our model shows a significant improvement of 8.3% in RMSE and 17.1% in correlation over the performance achieved by Raju and Das (2021).

5.1.1. High-Speed Enhancements

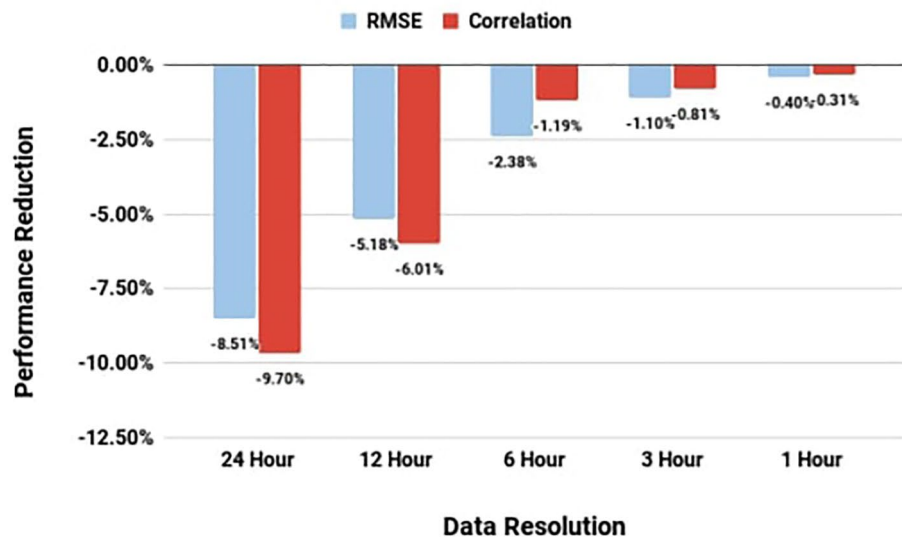
Regarding the forecasting of specific events, namely high-speed enhancements (HSEs), the same evaluation technique is employed for the identification of HSEs as described in Jian et al. (2015) (See their Section 8: Validation for Slow-to-Fast Stream Interactions, for a full description). Furthermore, because our data partitioning is discarding the buffer zones, all HSEs that occurred over those buffers are discarded. For comparison directly with Upendran et al. (2020), the true skill score is reported. Our best model achieves a true skill score of 0.387. This compares similarly with Upendran et al.'s (2020) 0.357. For the HSE that the model successfully captured, the RMSE in the peak is 99.1 km/s. However, noting the model's tendency to under-predict strong solar wind, the RMSE drops to 82.1 km/s after multiplying the prediction peaks by a corrective factor of 1.09.

5.2. Ablation Study

To demonstrate the stand-alone effect of our suggested techniques on the results, we conducted a study whereby each improvement is removed one at a time and the performance reduction reported. In the case of dropping the buffers, the no-buffer condition was implemented by making those buffers between the validation and training sets become part of the validation set, thus removing the separation between the two sets whilst adhering to a test-validation-train split that is comparable to that of the original condition. Figure 4a shows that the dominant improvement has been the adjustment of the sampling frequency, excluding it causes 8.51% performance reduction in RMSE and 9.70% in correlation. The solar wind speed at Earth changes on timescales that are much faster than 1 day (Meredith et al., 2011), suggesting that a higher sampling rate would capture extra information. In order to demonstrate the relationship between the sampling rate and performance, further training runs were completed at different resolutions. Figure 4b shows how the model performance improves with a higher sampling rate. By 1 hr cadence, the performance reduction is only 0.4% in RMSE and 0.31% in correlation. These results show that the more fine-grained the resolution the better, but clearly with diminishing returns. At 1 hr resolution, there is half the amount of data compared with 30 min cadence, so computational constraints will also dictate how high a resolution will be used.



(a) Performance reduction resulting from removing one improvement at a time.



(b) Performance reduction compared to 30 min resolution

Figure 4. Ablation study results. (a) Performance reduction resulting from removing one improvement at a time. (b) Performance reduction compared to 30 min resolution.

The other four methodological improvements deliver performance reductions between 0.58% and 1.63% in RMSE and between 0.6% and 2.16% in correlation. While these figures are modest in magnitude, it ought to be pointed out that the benefits appear uncorrelated between the methods, and when they are all combined, they deliver a significant improvement over the previous works. The removal of the Carrington rotation results in a performance reduction of 0.5% in RMSE and 1.19% in correlation. Again, although slight, this result justifies our inclusion of it. It also opens up the possibility of adding other useful values into the network before the final processing layers. An example might include the angle of the tilt of the Sun onto the plane of the sky (as observed from Earth), which can vary by a few degrees depending on the time of year. Augmenting the data set by flipping north to south also improves the model RMSE and correlation. It is not necessary for the augmented image to be

expected to produce the exactly same speed, the speed would just have to be highly correlated with the original image. Lastly, the inclusion of the batch normalization layer also results in a minor performance improvement. This was to be expected, as it can be viewed as a learned input normalization, which was established in the field to aid numerical stability of gradient descent methods and thus improve their convergence.

5.3. Prediction Analysis

Next, we analyze the predictions made by the best performing Swin Transformer model to get a better understanding of what aspects of the solar wind speed prediction task it gets right, and where it is limited.

5.3.1. Distribution

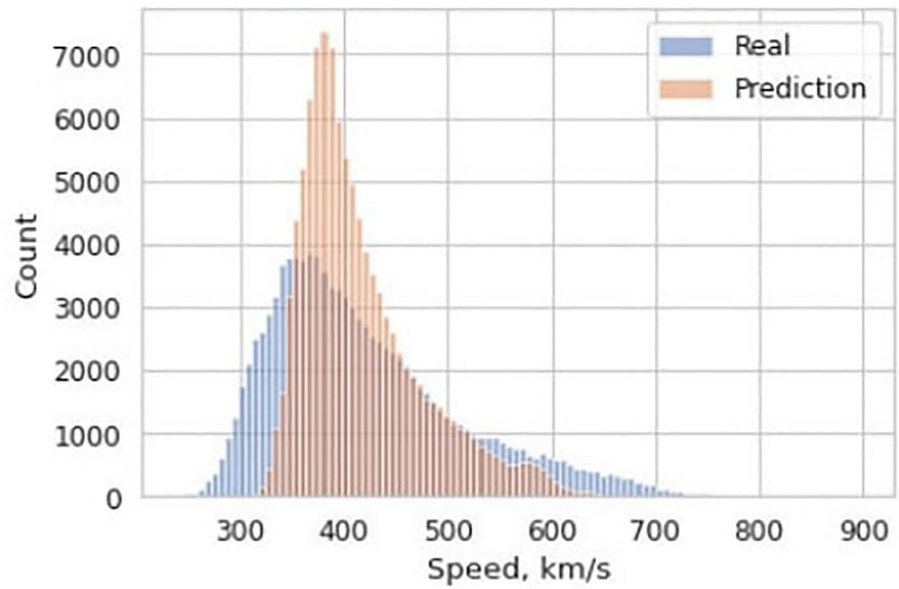
Figure 5a shows the distributions of the solar wind speeds predicted by the top model and the underlying ground truth. Both distributions are roughly centered around the same mean with a positive skewness, that is, they have long right-hand tails. The distributions differ significantly in their kurtosis. The real data has lower kurtosis, that is, it has more observations in both its right and left tails. The model's predictions have notably higher kurtosis, as it has a much more pronounced peak at around its mean and much fewer observations in its tails. This is to be expected as the L2 loss function chosen, which all models in this domain use, is known to prioritize the average fit of the model over fitting the extremities. The distributions by themselves, however, do not tell the full story. For that we need to look at Figure 5b, which shows the confusion matrix of binned solar wind speeds. Both predicted and actual solar wind speeds are split into four distinct class bins incremented by 100 km/s and 2 catch all classes one at each extreme of the distributions. Each block of the confusion matrix corresponds to one combination of a predicted class and a ground truth, that is, real, class. The value in the block represents the fraction of that real class that were classified as the predicted class. Under a perfect prediction, the blocks would read 1.0 along the diagonal and 0 everywhere else. This would mean that all solar wind speeds were correctly predicted in their class. As it is however, our model shows a tendency to over-predict the lower real solar wind speeds while under-predicting the higher solar wind speeds. Indeed, no solar wind speeds that were in the 700–900 km/s range were correctly predicted as such. Similarly, no solar wind speeds in the 100–300 km/s range were correctly predicted. This confirms our suspicion that it is the tail observations that are being regressed toward the mean that is driving both the error in the confusion matrix and the difference in the prediction and ground truth distributions.

5.3.2. Solar Cycle Variability

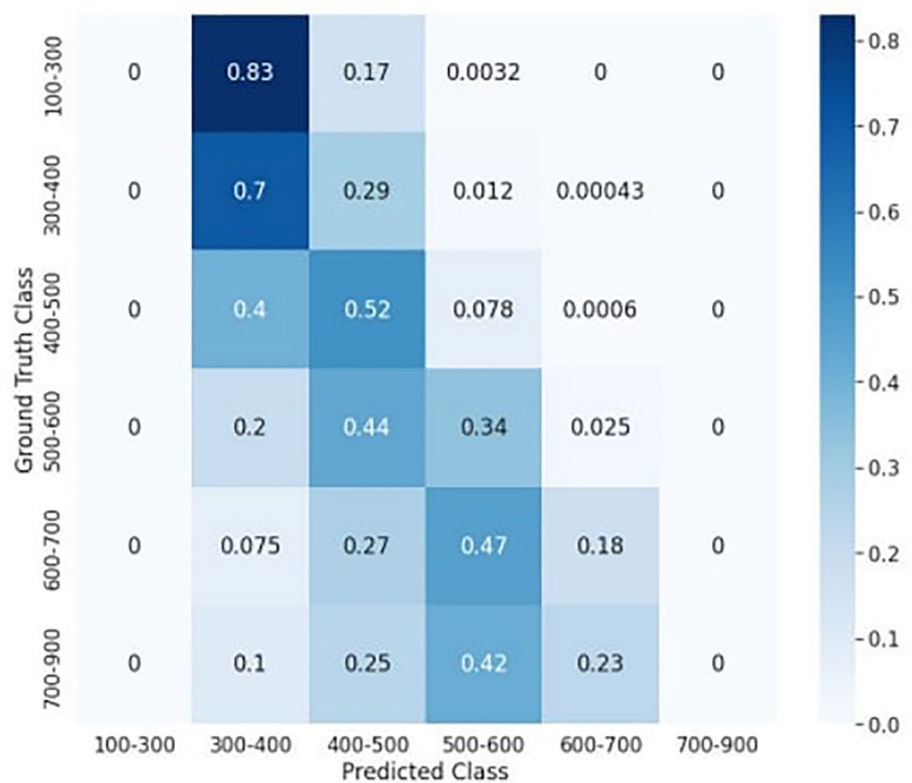
The measured and predicted solar wind speeds are shown in Figure 6 for the period 2010–2018. The performance of the model is highly dependent on the phase of the solar cycle with the model performing better during the declining phase of the solar cycle in 2016–2018. We examine this in more detail in Figure 7 where we plot the correlation of the model prediction with the ground truth at 6 months intervals (blue trace) against the sunspot number (red) in the same interval. The model's prediction correlation to the ground truth is strongest during the declining phase and worse around solar maximum. This relationship is confirmed when we view the data as correlation-sunspot number couples and visualize them in a scatter plot. This is shown in Figure 7b. We observe a strong, 0.78, negative correlation of the number of sunspots and the model prediction correlation to the ground truth. Since sunspot number is used to measure the solar cycle, this suggests that the model performance is highly dependent on the solar cycle and more specifically on the prevalent type of solar activity in a given period.

Indeed, a key component of the model's performance across the solar cycle is the type of encountered solar features. The top two panels of Figure 8 show the model's performance in early 2012, with 80.81 RMSE and 0.45 correlation, and in late 2016, with 73.32 RMSE and 0.81 correlation. The solar wind behavior in the later half of 2016, was driven by coronal holes and the high-speed solar wind streams associated with them. Whereas, 2012 had a much higher sunspot number and had far more Earth-directed CMEs.

We observe a marked difference in performance between predictions driven by different solar events—CMEs and coronal holes. Figures 8b and 8c show how the model captures the longer lasting, speed profile of a coronal hole quite well, while missing the speed profile of the sudden CME. This offers an explanation to the pronounced variability in the model's prediction quality. The solar activity in the declining phase is driven by coronal holes. These are more easily picked up by the models. Since the Sun in the later half of 2016 was in the declining phase, the models' performance was much better. In 2012, a year with far more CMEs, the model performance was reduced, as the models struggled to catch the CMEs.



(a) Distribution of predicted and real speeds



(b) Confusion matrix of binned speeds (km/s)

Figure 5. Distribution and confusion matrix of predicted speeds. (a) Distribution of predicted and real speeds. (b) Confusion matrix of binned speeds (km/s).

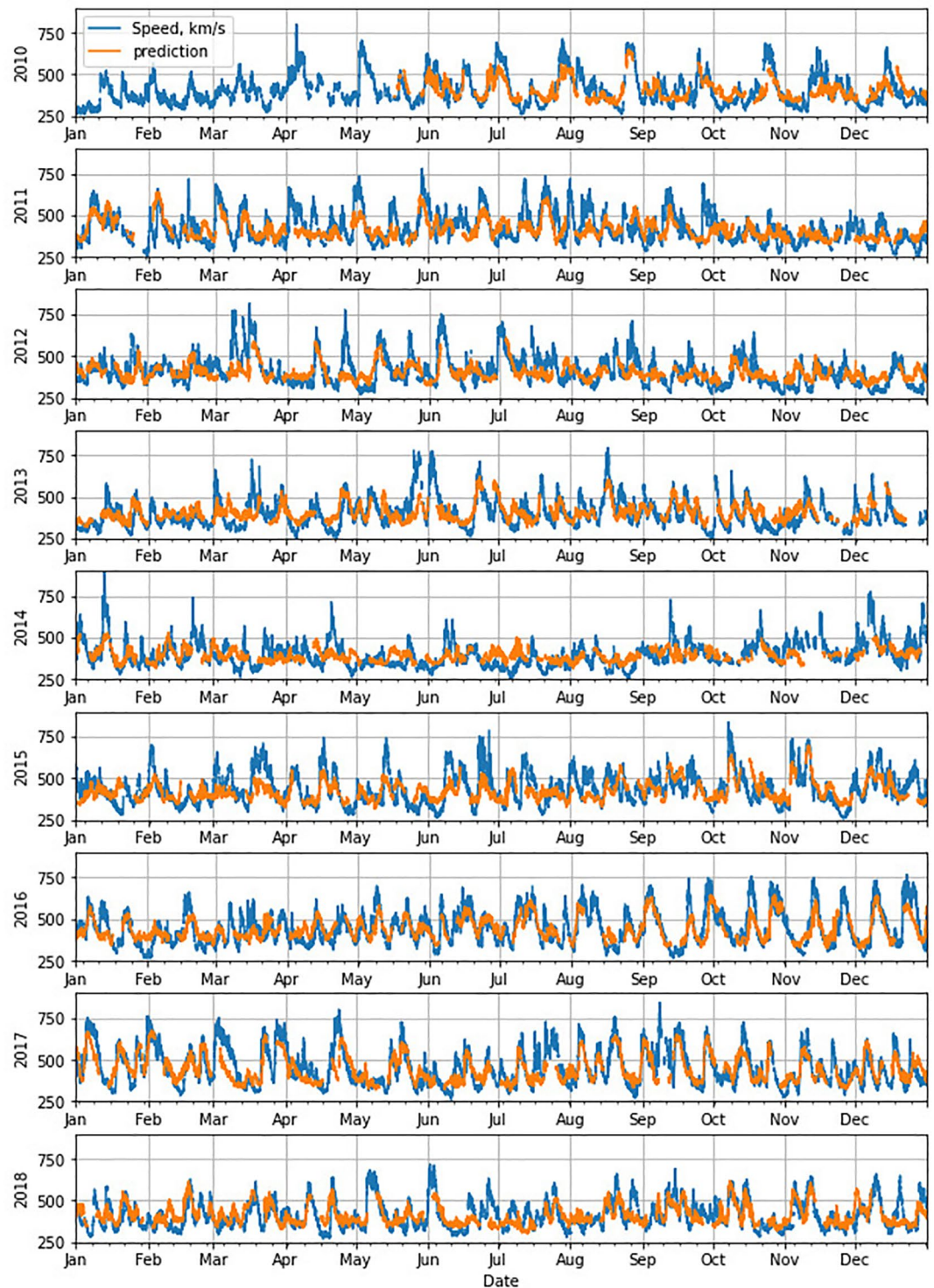
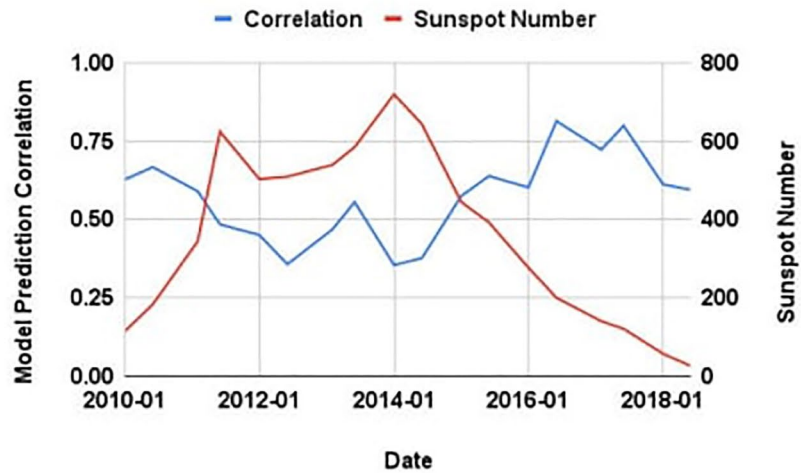
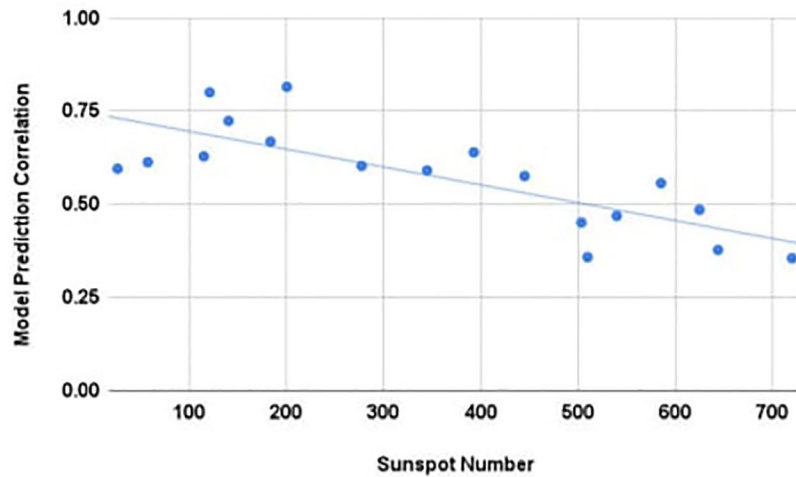


Figure 6. Plots of the measured (blue) and predicted (orange) solar wind speeds for the period 2010–2018. (a) Model prediction correlation (blue trace) and sunspot number (red space) as a function of UT date. (b) Plot of the model prediction correlation as a function of sunspot number. The plotted blue trace is the fitted linear relationship.

Since extreme events are, by their very nature, the events that are most important to society, the failure to fit on the more sudden CMEs is a chief limitation of the models developed in this space. It can be ascribed to the lack of significant and persistent CME-related features in the EUV images, preventing them from being captured by the



(a) Model prediction correlation (blue trace) and sunspot number (red trace) as a function of UT date



(b) Plot of the model prediction correlation as a function of sunspot number. The plotted blue trace is the fitted linear relationship.

Figure 7. Model performance compared to sunspot number.

models. We note that ML models using solar EUV images alone to forecast other space weather related parameters such as geomagnetic activity as measured by the AE or Kp indices or suprathermal electrons at geostationary orbit would most likely suffer from the same limitation resulting in a similar pattern of behavior with the best correlations during the declining phase of the solar cycle and the worst correlations around solar maximum.

5.3.3. Coronal Hole Area

It has been empirically established that there is a linear relationship between coronal hole area at low latitudes and peak solar wind speed (Hofmeister et al., 2018; Nolte et al., 1976). In order to test whether our model has learned this relationship we need to devise a way of obtaining images with specified coronal hole sizes at the desired latitude. We chose to generate our images using a background of enlarged uneventful solar region and a

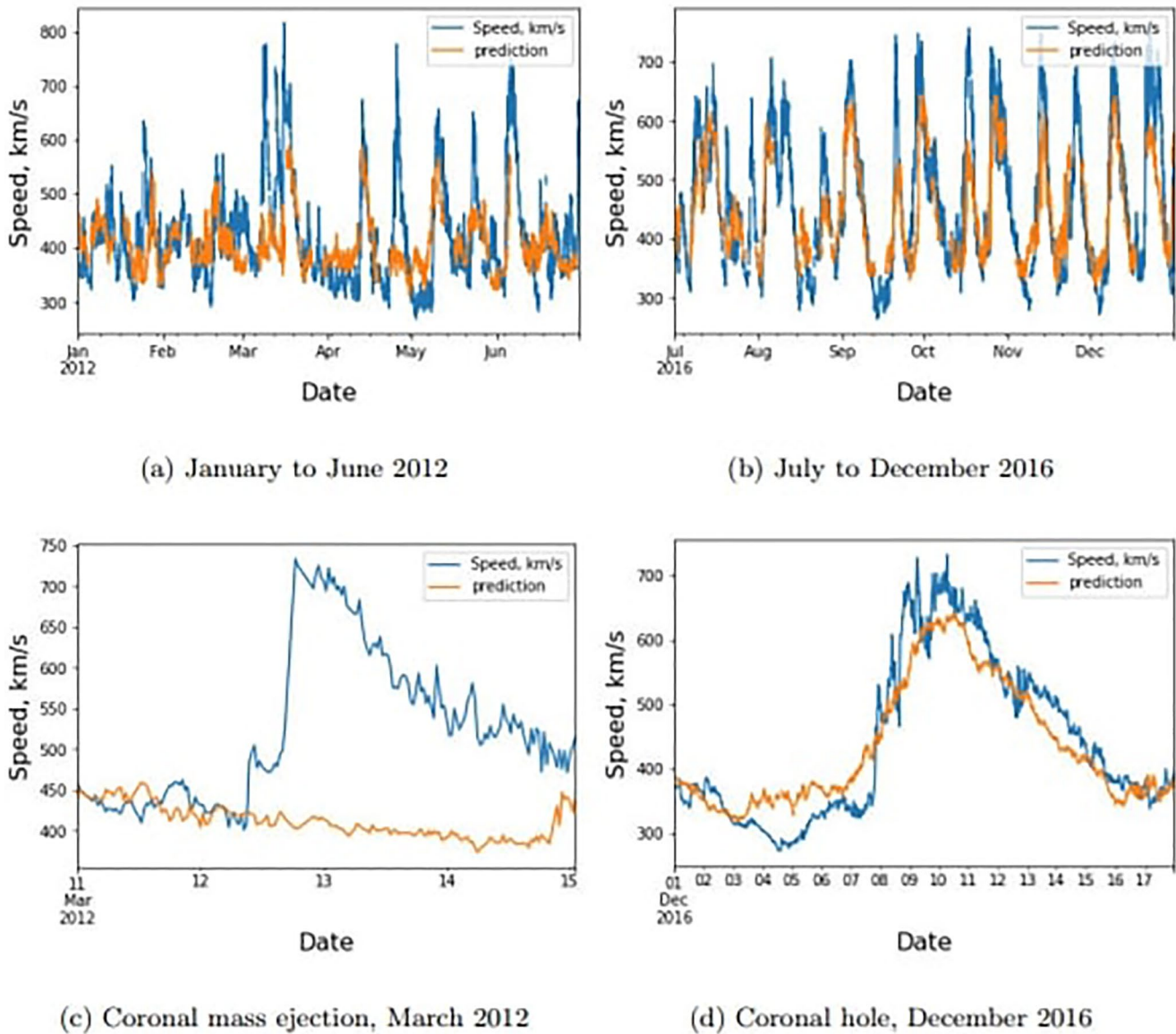


Figure 8. Solar Swin Transformer performance in different parts of the solar cycle and on different solar phenomena. (a) January to June 2012. (b) July to December 2016. (c) Coronal mass ejection, March 2012. (d) Coronal hole, December 2016.

patch extracted from a coronal hole that can be sized as desired. Each patch size is moved horizontally across the center of image, and the model's peak prediction for that size is recorded. Figure 9 plots the predicted peak solar wind speeds against the patch sizes in blue. The red line is a fitted linear function of best fit, with a coefficient of determination (R^2) of 0.953. It shows that our model succeeded to learn a close linear relationship as described by Nolte et al. (1976) and Hofmeister et al. (2018).

5.3.4. Coronal Hole Position

We investigate the role of the position of a coronal hole on the forecasted solar wind speed. A hole of fixed area in the plane of the image, 40 pixels by 40 pixels—which is about 280 arcsec by 280 arcsec in helioprojective coordinates (Thompson, 2006) and corresponds to 1,600 pixel area as shown in Figure 9, is moved around an image of quiet solar background to see the effect of its position on the forecast. The results are presented in Figure 10. To clarify, the color of the square at $(-675, -675)$ in the figure represents the solar wind speed forecasted 4 days later with a coronal hole centered at those coordinates. The model forecasts higher solar wind speeds for simulated coronal holes that are closer to the equator. This agrees with empirical relationships established in works such

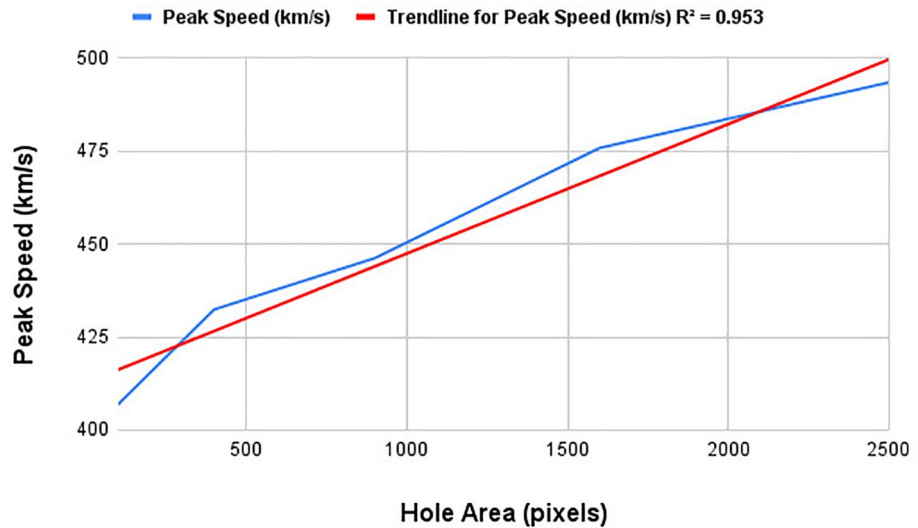


Figure 9. Peak speed of coronal holes (blue trace) at solar equator versus coronal hole area. Red trace shows the fitted linear relationship with an R^2 of 0.953.

as Hofmeister et al. (2018) where the observed solar wind speed from a given coronal hole is lower the further from the equator it is. Notably, the model gives higher solar wind speeds for holes on the right of the image. If the solar wind from a coronal hole took exactly 4 days to reach L1, we would expect the heatmap to show the highest speeds in the center. However, the solar wind, when elevated, takes less than 4 days to reach the Earth. This is why the image is brighter on the right hand side of the image, because the forecasted speed is for 4 days later than the image, but the solar wind takes less than that due to the presence of the coronal hole. A limitation of the model is however noticeable from this figure, as small movements in the position result in swings in the outputted speed.

5.3.5. Coronal Hole Intensity

Finally, Obridko et al. (2009) found that the darker the coronal hole, the larger is the peak of the associated high-speed stream. We test whether our model learned this empirical relationship by incrementally increasing the minimum brightness of a coronal hole. At each step, any pixel value below the minimum threshold is increased to the minimum value. Figure 11 shows the predicted solar wind speed for a large coronal hole visible on the day of 6 December 2016 at 00:00:00 UT at various minimum intensities. As we increase the brightness of the coronal hole, the model starts to forecast lower solar wind speeds. This suggests that the model has learned the Obridko et al. (2009) empirical relationship that the darker the hole, the stronger the solar wind.

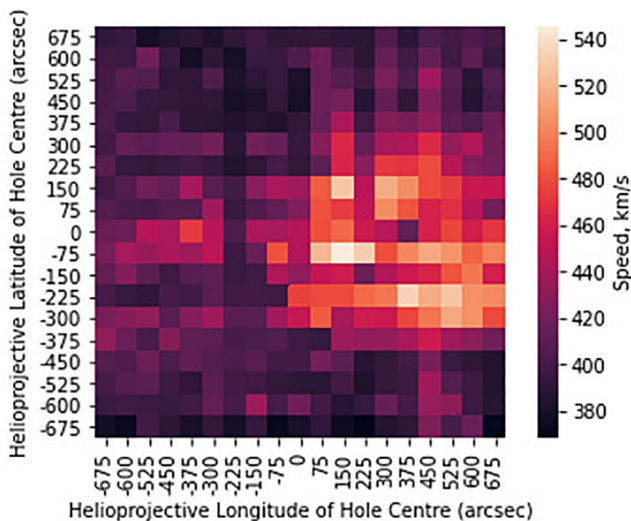


Figure 10. Central positions of coronal hole and its forecasted speed.

6. Conclusions

This study uses attention-based machine vision models and a set of methodological and modeling improvements to forecast the solar wind speed at L1 using solar images at 211 Å wavelength. These improvements result in 11.1% lower RMSE and 17.4% higher prediction correlation with the ground truth when compared to previous works. The most significant improvement comes from moving from a daily to a 30 min sampling rate. Additionally, this study observed that attention-based architectures in general have about 2%–3% performance edge in both RMSE and correlation over the previously used convolutional alternatives. The model's performance is highly dependent on the position in the solar cycle. The model performance is strongly negatively correlated with the sunspot number, as the model performance is better in the declining phase of the solar cycle when the solar wind behavior is dominated by coronal hole activity. Finally, the model has independently learned three empirical relationships between coronal features and their associated solar

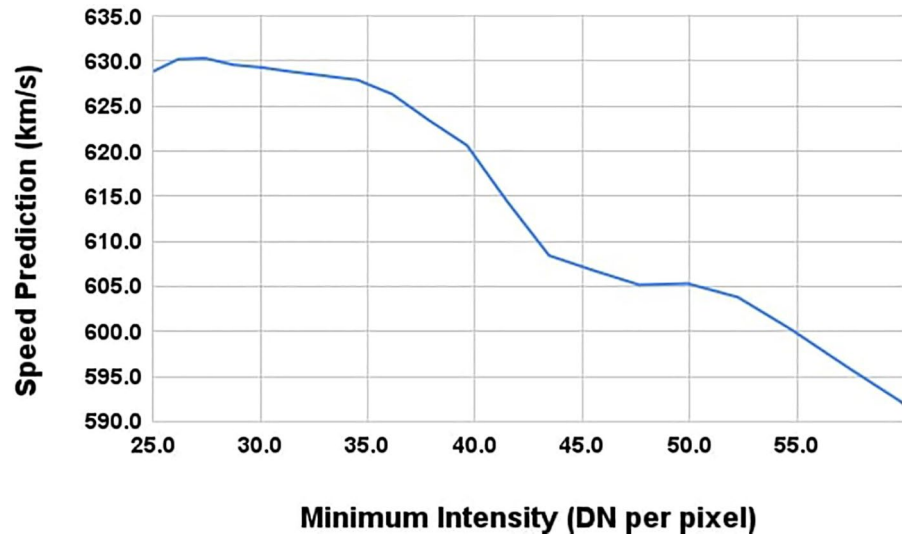


Figure 11. Plot of the predicted solar wind speed as a function of minimum pixel intensity for an image with a large coronal hole observed on 6 December 2016.

wind speeds established by previous publications. First, it complied with the observed linear relationship between coronal hole area and the peak solar wind speed associated with it. Second, it learned that equatorial coronal holes are associated with higher solar wind speeds when compared to those at higher latitudes. Lastly, the model learned that the darker the coronal hole, the stronger the solar wind speed associated with it.

Data Availability Statement

The SDO ML 211 Å image data is available here: <https://purl.stanford.edu/vk217bh4910>. The OmniWeb solar wind data is available here: https://omniweb.gsfc.nasa.gov/form/omni_min.html. Code for analyzing model output as well as the python code for the model, is supplied here: <https://github.com/eddbrown/solar-swin-transformer-output-data>.

Acknowledgments

For this study, the authors acknowledge extensive use of the SDO ML (Galvez et al., 2019) data set. These images are processed versions of images taken by the AIA instrument aboard the SDO. Furthermore, the authors acknowledge the use of the OMNIWeb service and their provided measurements of the solar wind speed at L1. Regarding software, the authors acknowledge extensive use of python and the python packages NumPy (Harris et al., 2020) and PyTorch (Paszke et al., 2019). For experiment tracking and general project organization, the authors acknowledge extensive use of the software provided by Weights and Biases (Biewald, 2020). Additionally, the authors would like to thank the BAS AI Lab and BAS IT team for their computing facilities and general advice/support. N. P. Meredith and R. B. Horne would like to acknowledge support from the NERC grants NE/V00249X/1 (Sat-Risk) and NE/R016038/1.

References

- Bailey, R. L., Reiss, M. A., Arge, C. N., Möstl, C., Henney, C. J., Owens, M. J., et al. (2021). Using gradient boosting regression to improve ambient solar wind model predictions. *Space Weather*, 19(5), e2020SW002673. <https://doi.org/10.1029/2020SW002673>
- Biewald, L. (2020). *Experiment tracking with weights and biases*. Retrieved from <https://wandb.ai/site>
- Borovsky, J. E., & Denton, M. H. (2006). Differences between CME-driven storms and CIR-driven storms. *Journal of Geophysical Research*, 111(A7). <https://doi.org/10.1029/2005JA011447>
- Burlaga, L., & Lepping, R. (1977). The causes of recurrent geomagnetic storms. *Planetary and Space Science*, 25(12), 1151–1160. [https://doi.org/10.1016/0032-0633\(77\)90090-3](https://doi.org/10.1016/0032-0633(77)90090-3)
- Chandorkar, M., Furtlehner, C., Poduval, B., Camporeale, E., & Sebag, M. (2019). Dynamic time lag regression: Predicting what & when. In *International Conference on Learning Representations*. <https://openreview.net/pdf?id=SkxybANtDB>
- Chen, G.-M., Xu, J., Wang, W., Lei, J., & Burns, A. G. (2012). A comparison of the effects of CIR- and CME-induced geomagnetic activity on thermospheric densities and spacecraft orbits: Case studies. *Journal of Geophysical Research*, 117(A8). <https://doi.org/10.1029/2012JA017782>
- Cranmer, S. R. (2009). Coronal holes. *Living Reviews in Solar Physics*, 6, 366. <https://doi.org/10.12942/lrsp-2009-3>
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–255). <https://doi.org/10.1109/CVPR.2009.5206848>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, abs/2010.11929. <https://arxiv.org/abs/2010.11929>
- Galvez, R., Fouhey, D. F., Jin, M., Szenicer, A., Muñoz-Jaramillo, A., Cheung, M. C. M., et al. (2019). A machine-learning data set prepared from the NASA solar dynamics observatory mission. *The Astrophysical Journal – Supplement Series*, 242(1), 7. <https://doi.org/10.3847/1538-4365/ab1005>
- Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2018). Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*. <https://arxiv.org/abs/1811.12231>
- Han, K., Xiao, A., Wu, E., Guo, J., Xu, C., & Wang, Y. (2021). *Transformer in transformer*. <https://arxiv.org/abs/2103.00112>
- Huang, G., Liu, Z., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261–2269).
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). <https://doi.org/10.1109/CVPR.2016.90>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hofmeister, S. J., Veronig, A. M., Temmer, M., Vennerstrom, S., Heber, B., & Vrsnak, B. (2018). The dependence of the peak velocity of high-speed solar wind streams as measured in the ecliptic by ace and the stereo satellites on the area and co-latitude of their solar source coronal holes. *Journal of Geophysical Research: Space Physics*, 123, 1738–1753. <https://doi.org/10.1002/2017JA024586>
- Horne, R. B., Phillips, M. W., Glauert, S. A., Meredith, N. P., Hands, A. D. P., Ryden, K. A., & Li, W. (2018). Realistic worst case for a severe space weather event driven by a fast solar wind stream. *Space Weather*, 16(9), 1202–1215. <https://doi.org/10.1029/2018SW001948>
- Jian, L., MacNeice, P., Taktakishvili, A., Odstrcil, D., Jackson, B., Yu, H.-S., et al. (2015). Validation for solar wind prediction at earth: Comparison of coronal and heliospheric models installed at the CCMC. *Space Weather*, 13(5), 316–338. <https://doi.org/10.1002/2015SW001174>
- Kilpua, E., Balogh, A., Von Steiger, R., & Liu, Y. (2017). Geoeffective properties of solar transients and stream interaction regions. *Space Science Reviews*, 212(3), 1271–1314. https://doi.org/10.1007/978-94-024-1588-9_9
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. <https://arxiv.org/abs/1412.6980>
- Kozyra, J. U., Crowley, G., Emery, B. A., Fang, X., Maris, G., Mlynczak, M. G., et al. (2006). Response of the upper/middle atmosphere to coronal holes and powerful high-speed solar wind streams in 2003. In *Recurrent magnetic storms: Corotating solar wind streams* (pp. 319–340). American Geophysical Union (AGU). <https://doi.org/10.1029/167GM24>
- Lemen, J. R., Akin, D. J., Boerner, P. F., Chou, C., Drake, J. F., Duncan, D. W., et al. (2011). The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO). In *The solar dynamics observatory* (pp. 17–40). Springer.
- Liu, D. D., Huang, C., Lu, J. Y., & Wang, J. S. (2011). The hourly average solar wind velocity prediction based on support vector regression method. *Monthly Notices of the Royal Astronomical Society*, 413(4), 2877–2882. <https://doi.org/10.1111/j.1365-2966.2011.18359.x>
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *CoRR*, abs/2103.14030. <https://arxiv.org/abs/2103.14030>
- Meredith, N. P., Horne, R. B., Lam, M. M., Denton, M. H., Borovsky, J. E., & Green, J. C. (2011). Energetic electron precipitation during high-speed solar wind stream driven storms. *Journal of Geophysical Research*, 116(A5). <https://doi.org/10.1029/2010JA016293>
- Nolte, J., Krieger, A., Timothy, A., Gold, R., Roelof, E., Vaiana, G., et al. (1976). Coronal holes as sources of solar wind. *Solar Physics*, 46(2), 303–322. <https://doi.org/10.1007/BF00149859>
- Obridko, V., Shelting, B., Livshits, I., & Askerov, A. (2009). Relationship between the contrast of coronal holes and parameters of the solar wind streams. *Astronomy Reports*, 53(11), 1050–1058. <https://doi.org/10.1134/S1063772909110109>
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 32, pp. 8024–8035). Curran Associates, Inc.
- Raju, H., & Das, S. (2021). CNN-based deep learning model for solar wind forecasting. *Solar Physics*. <https://doi.org/10.1007/s11207-021-01874-6>
- Richardson, I. G., Webb, D. F., Zhang, J., Berdichevsky, D. B., Biesecker, D. A., Kasper, J. C., et al. (2006). Major geomagnetic storms (Dst less than or equal to -100 nt) generated by corotating interaction regions. *Journal of Geophysical Research*, 111(A7). <https://doi.org/10.1029/2005JA011476>
- Ridpath, I. (2012). *A dictionary of astronomy*. Oxford University Press.
- St. Cyr, O. C., Howard, R. A., Sheeley, N. R., Jr., Plunkett, S. P., Michels, D. J., Paswaters, S. E., et al. (2000). Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998. *Journal of Geophysical Research*, 105(A8), 18169–18185. <https://doi.org/10.1029/1999JA000381>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. E., Anguelov, D., et al. (2014). Going deeper with convolutions. *CoRR*, abs/1409.4842.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818–2826). <https://ieeexplore.ieee.org/document/7780677>
- Thompson, W. (2006). Coordinate systems for solar image data. *Astronomy & Astrophysics*, 449(2), 791–803.
- Tsurutani, B. T., Gonzalez, W. D., Gonzalez, A. L. C., Tang, F., Arballo, J. K., & Okada, M. (1995). Interplanetary origin of geomagnetic activity in the declining phase of the solar cycle. *Journal of Geophysical Research*, 100(A11), 21717–21733. <https://doi.org/10.1029/95JA01476>
- Turner, N. E., Mitchell, E. J., Knipp, D. J., & Emery, B. A. (2006). Energetics of magnetic storms driven by corotating interaction regions: A study of geoeffectiveness. In *Recurrent magnetic storms: Corotating solar wind streams* (pp. 113–124). American Geophysical Union (AGU). <https://doi.org/10.1029/167GM11>
- Upendran, V., Cheung, M. C. M., Hanasoge, S., & Krishnamurthi, G. (2020). Solar wind prediction using deep learning. *Space Weather*, 18(9), e2020SW002478. <https://doi.org/10.1029/2020SW002478>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). *Attention is all you need*. <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
- Wintoft, P., & Lundstedt, H. (1997). Prediction of daily average solar wind velocity from solar magnetic field observations using hybrid intelligent systems. *Physics and Chemistry of the Earth*, 22(7), 617–622. (Low-altitude investigation of dayside magnetospheric boundary processes). [https://doi.org/10.1016/S0079-1946\(97\)00186-9](https://doi.org/10.1016/S0079-1946(97)00186-9)
- Wintoft, P., & Lundstedt, H. (1999). A neural network study of the mapping from solar magnetic fields to the daily average solar wind velocity. *Journal of Geophysical Research*, 104(A4), 6729–6736. <https://doi.org/10.1029/1998JA000183>
- Yang, Y., Shen, F., Yang, Z., & Feng, X. (2018). Prediction of solar wind speed at 1 au using an artificial neural network. *Space Weather*, 16(9), 1227–1244. <https://doi.org/10.1029/2018SW001955>