Removing cloud shadows from ground-based solar imagery

Amal Chaoui
  École Centrale de Lille

Jay Paul Morgan
  Université de Toulon, Aix Marseille Univ, CNRS, LIS

Adeline Paiement (✉ adeline.paiement@univ-tln.fr)
  Université de Toulon, Aix Marseille Univ, CNRS, LIS

Jean Aboudarham
  Observatoire de Paris/PSL, LESIA, CNRS

Additional Declarations: No competing interests reported.

Removing cloud shadows from ground-based solar imagery

Amal Chaoui[1], Jay Paul Morgan[2], Adeline Paiement[2*],
Jean Aboudarham[3]

[1]Université de Lille, Centrale Lille, Lille, France.
[2]Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France.
[3]Observatoire de Paris/PSL, LESIA, CNRS, Paris, France.

*Corresponding author(s). E-mail(s): adeline.paiement@univ-tln.fr;
Contributing authors: amal.chaoui@centrale.centralelille.fr;
jay.morgan@univ-tln.fr; Jean.Aboudarham@obspm.fr;

Abstract

The study and prediction of space weather entails the analysis of solar images showing structures of the Sun's atmosphere. When imaged from the Earth's ground, images may be polluted by terrestrial clouds which hinder the detection of solar structures. We propose a new method to remove cloud shadows, based on a U-Net architecture, and compare classical supervision with conditional GAN. We evaluate our method on two different imaging modalities, using both real images and a new dataset of synthetic clouds. Quantitative assessments are obtained through image quality indices (RMSE, PSNR, SSIM). We demonstrate improved results with regards to the traditional cloud removal technique and a sparse coding baseline, on different cloud types and textures.

Keywords: Image cleaning, solar imaging, deep learning, U-Net, C-GAN
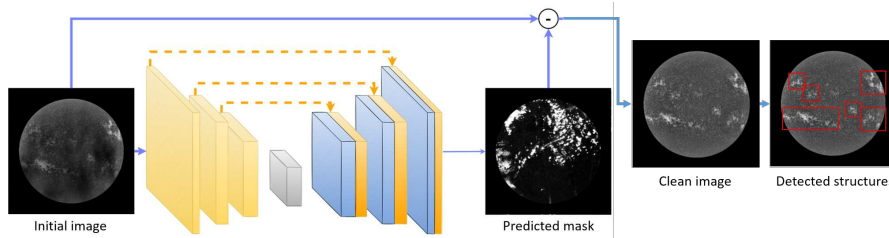
Acknowledgments

Fig. 1 Overview of cloud removal from solar imagery. A cloud mask is predicted and subtracted from the original image, to prepare for a subsequent detection of solar structures (here ARs in red boxes). Our proposed method only deals with the cloud removal stage.

1 Introduction and previous works

This paper addresses the problem of cleaning ground-based solar imagery from cloud shadow contaminants, as illustrated in Fig. 1, in support to forthcoming analyses with solar physics and space weather applications. Monitoring the activity of the Sun is a task of growing importance in the frame of space weather research. The main sources of severe space weather disturbances at Earth are energetic outbursts from the Sun, *e.g.* solar flares and coronal mass ejections. They are strongly associated with solar structures of the chromosphere (*e.g.* active regions (AR) and filaments, as illustrated in Fig. 3), visible from the Earth ground in the light emitted by ionised calcium (Ca-II, at a wavelength of 3933.7Å) and the H-α line of neutral hydrogen (6562.8Å). Despite their affordability, images taken from the ground present a particular difficulty: they are prone to be tainted by cloud shadows when the sky is not clear. These shadows are rarely uniform and may be mistaken for solar structures (*e.g.* filaments). Thus, it is substantial to remove them before automatic solar structure detection and monitoring.

Very few previous works addressed this problem. [1, 2] designed denoising methods for H-α images. Feng *et al.* [1] proposed a metric for assessing the level of cloud coverage in an image by considering the homogeneity of image intensity between different quadrants of the solar disk. Additionally, they proposed a denoising method that first estimates the pixel-wise transmittance of clouds $\hat{I}_t$ from a cloud-free temporal neighbour using a low-pass morphological filter to smooth out solar structures (*i.e.* filaments). Second, it uses it to normalise the original image I_{ini}:

$$\hat{I}_{\text{clean}} = I_{\text{ini}}/\hat{I}_t \tag{1}$$

In [2], Fuller and Aboudarham perform two stages of median filtering to identify large scale intensity variations that correspond to clouds, while thresholding out the solar structures. In both methods, the use of low-pass filtering to distinguish clouds from solar structures is a possible limitation when their spatial scales are similar in at least one direction.

In other application domains, some works removed large contaminants, such as clouds and shadows, from remote sensing or natural images. Due to clouds transparency, the task of removing clouds in Sun observations resembles to a degree to these tasks. While there is a large literature on denoising methods *e.g.* [3], we focus

on the removal of large and semi-transparent contaminants, as this is the closest to our task. Methods that perform inpainting over opaque contaminants, such as [4] for thick clouds, are not considered so as to avoid introducing artefacts in the study of solar structures (images with thick clouds are therefore not usable). We also discard methods that exploit the physics of image formation for natural images, *e.g.* [5–8], as solar imaging follows a different physics. Thin cloud and shadow removal methods are good candidates for our task of removing semi-transparent clouds while preserving the background solar structures.

Some shadow suppression methods *e.g.* [9] rely upon the sparse approximation theory. A local dictionary of atoms is learnt from image patches as a basis for sparse coding. This allows decomposing the image into reflectance (invariant to shading), geometry (fine structural details), and shadow components, the latter being excluded during image reconstruction. The use of local dictionaries, rather than generic ones (*e.g.* wavelets, curvelets), helps in learning atoms that adapt to each single image. However, this approach cannot gather further insights from multiple examples.

This limitation is addressed by (deep) learning-based methods. [10, 11] use encoder-decoder deep neural networks (DNN) to remove shadows by learning a mapping between the initial images I_{ini} and their binary mask $\hat{M}$ of shadow ratio, which is reminiscent of the transmittance image of [1]:

$$\hat{I}_{\text{clean}} = I_{\text{ini}}/\hat{M} \qquad (2)$$

[12] and pix2pix [13] in [4] rely on skip connections (U-Net-style) to preserve spatial resolution, and a conditional generative adversarial network (C-GAN) for more efficient training. They obtain state-of-the-art (SoTA) results in shadow and thin cloud removal. In [12], an additional C-GAN outputs the denoised image from predicted shadow mask, while pix2pix of [4] outputs the cleaned image directly.

We explore the applicability of these denoising methods for the new application of removing cloud shadows from solar imagery. We evaluate sparse representation-based and deep learning-based approaches, and compare against [1, 2] (Section 4). Inspired by [4, 12], we use a combination of U-Net and C-GAN (Section 3) to obtain the new SoTA for this task.

The methods are comparatively assessed on two imaging modalities. To support these experiments, we introduce two new datasets of Sun images with artificial and real clouds.

2 Data

Our images are from Paris-Meudon (PM) observatory. They are 1024×1024 px full-disk spectroheliograms of two modalities: Ca-II (for studying ARs), and H-α (for filaments), as highlighted in Fig. 3. Observations span different periods of the solar cycle in order to reflect variations in the Sun's activity and number / size of observed structures. 96 Ca-II and 111 H-α images are sampled between May 2002 and June 2017. Sampling uses the metric of cloud coverage level proposed by [1] to select the images that are tainted with cloud shadows.
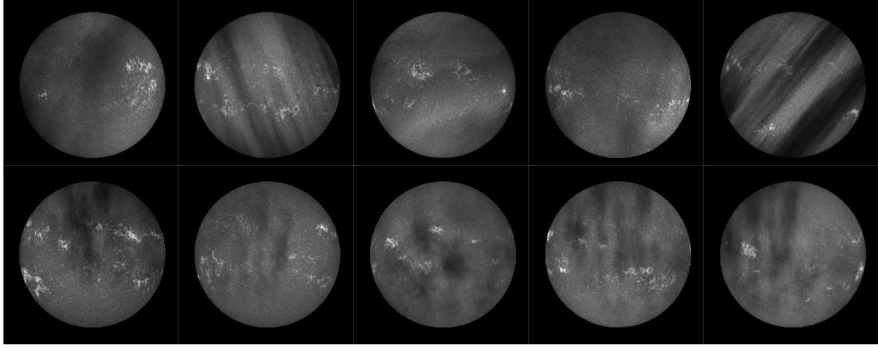
Fig. 2 Examples of real (top) and synthetic (bottom) cloud contamination in Ca-II images.

PM images are taken in principle during clear sky days. However, even during those days, small thin clouds emerging in the horizon cannot be prevented and result in light cloud stains. Cloud shadows can come with different textures. In PM observatory, the Sun disk is scanned by a thin moving slit over 20 to 100 sec. [14]. In case of fast enough cloud motion, due to the slit movement alongside the motion of clouds, cloud spots may be shifted along the scanning process, creating straight lines on the image (denoted as "streaked" texture). Slower cloud motions result in the "fluffy" texture that is customary of many terrestrial clouds. The "streaked" texture may provide more challenges to denoising methods, due to having different spatial scales in different directions. Fig. 3 illustrates samples of streaked and fluffy textures with different cloud cover thicknesses and streak orientations.

The cloudy images don't have cloud-free counterparts for training DNNs. Thus, we introduce a dataset of real (initially cloud-free) images with added synthetic cloud shadows. Using the cloud coverage level metric of [1], we selected reasonably cloud-free images in the same time period as above. A basic cloud texture was obtained from [15]. For each cloud-free image, a new cloud shadow was composed by summing a random number (in range [2, 3]) of duplicates of the basic texture with horizontal and vertical random resizing between 50-100% of the original size, random flips (around one or two axes), and random transparencies (in range [0.1, 0.4]). It was then overlapped to the cloud-free image. This procedure provides realistic images of fluffy and semi-transparent cloud shadow contamination. The Ca-II and H-α synthetic datasets comprise respectively 319 and 367 pairs of shadow/shadow-free images, split into 179/44/96 and 205/51/111 training/validation/testing pairs. The combination of our real and artificial cloud datasets provides a variety of streaked and fluffy cloud textures. They are publicly available[1], along with the pretrained cloud-removal DNNs[2].

Images have their intensity normalised in range $[0, 1]$, centered solar disks with normalised radii, and background pixels set to zero outside of the solar disk to remove any distracting features from background stars and imaging artefacts.

---

[1]https://doi.org/10.5281/zenodo.7684200
[2]https://github.com/jaypmorgan/cloud-removal

3 Methodology

We adopt the deep learning approach due to its SoTA results in large contaminants removal. As in [12, 13], we use skip connections to preserve spatial resolution and details[3]. Inspired by the results of previous methods, we compare the classical fully supervised and the C-GAN training strategies.

We experiment with our DNNs predicting either the cleaned image directly, or a cloud shadow mask, to be used as in [1, 10, 11]. To avoid division by zero and improve stability, we compute the clean image as:

$$\hat{I}_{\text{clean}} = \frac{I_{\text{ini}}}{\hat{M} + \epsilon} \tag{3}$$

We empirically verified that the exact value of ϵ has little impact on the results, and we set it to $1e^{-5}$ in our experiments. Regressing a shadow mask instead of the cleaned image directly is motivated by the physics of solar imaging as in [1]. In addition, estimating a transmittance or shadow ratio, rather than the pixel's absolute intensity, reduces the risk to create artefacts that would hinder the forthcoming study of solar structures. Finally, estimating a shadow mask should be a less complex task for the DNN, which may be learnt more efficiently. Indeed, reproducing the exact patterns of the more complex and diverse solar structures in the output image would be an unnecessary additional task for the DNN.

With Eq. 3, the loss gradient varies as $I_{\text{ini}}/\hat{M}^2$, meaning that the darkest pixels of the images, where cloud thickness is high, see higher loss gradients and require a finer adjustment of their mask value. In an attempt to improve stability during training through more consistent loss gradient magnitudes, we also experiment with an alternative implementation *i.e.* adding in the intensity that was removed by clouds. This is similar to predicting a residual mask:

$$\hat{I}_{\text{clean}} = I_{\text{ini}} + \hat{M} \tag{4}$$

We compare two training strategies: a fully supervised DNN (denoted as FS) and a C-GAN following the pix2pix setup. For ease of comparison, the FS DNN and C-GAN generator have the same U-Net-style architecture[3]: down-sampling and up-sampling blocks along with skip connections, each block as defined in [16]. We experiment with a varying number of blocks to determine the optimal complexity and receptive field needed to account for both large scale contaminants and smaller scale solar structures. Due to memory constraints, we go up to 6 down/up-sampling blocks, and find it to be the best performing architecture. In the future, and on a different hardware, more blocks could be tried.

We use a linear activation in the last block when predicting $\hat{I}_{\text{clean}}$ directly or $\hat{M}$ as in Eq. 4. For Eq. 3, we find that imposing the output to be in range $[0, 1]$ is necessary to avoid instability issues, so we use $0.5 \tanh +0.5$ as activation.

The discriminator in the C-GAN is the same as in pix2pix [13]: a PatchGAN network that considers image patches (small receptive fields) to capture high-frequencies.

---

[3]We use the U-Net implementation from https://github.com/mateuszbuda/brain-segmentation-pytorch.

5

Table 1 Comparison of the three DNN outputs on the Ca-II synthetised dataset. Results are in format mean(std). SSIM and RMSE are as $e-2$.

DNN	Output	PSNR ↑	SSIM ↑	RMSE ↓
Fully sup.	$\hat{I}_{\text{clean}}$	30.2(0.4)	98.2(0.5)	3.9(0.1)
	$\hat{M}$ Eq. 3	28.9(0.2)	98.8(0.1)	4.5(0.1)
	$\hat{M}$ Eq. 4	**30.6(0.3)**	**98.9(0.1)**	**3.7(0.1)**
C-GAN	$\hat{I}_{\text{clean}}$	29.4(0.6)	97.7(0.4)	4.1(0.2)
	$\hat{M}$ Eq. 3	27.3(0.6)	97.6(0.8)	5.1(0.2)
	$\hat{M}$ Eq. 4	**30.0(0.4)**	**98.8(0.2)**	**3.9(0.1)**

We experimentally verified that the recommended receptive field of 70×70 is also optimal in our application. The discriminator is applied to the cleaned image rather than predicted shadow mask to assess the preservation of solar structures.

For the C-GAN, we use the loss function suggested in [13]. The loss function of the FS network and of the C-GAN generator uses the pixel-wise L_1-norm since it promotes sparsity and encourages less blurring as noted in [13]. The pixel-level loss is computed between the target I_t and cleaned image $\hat{I}_{\text{clean}}$:

$$\mathcal{L}_{L_1}(I_t, \hat{I}_{\text{clean}}) = \frac{1}{N} \sum_{i=1}^{N} |I_t^i - \hat{I}_{\text{clean}}^i| \tag{5}$$

For both DNNs, we use the default U-Net[3] training scheme and batch size of 3. We train for 100 epochs and test on the epoch with best validation loss (FS DNN) or generator's training loss (C-GAN, following the usual practice for GANs). During testing of the C-GAN, the generator is run with dropout, and batch normalisation considers the statistics of the batch, as in [13].

4 Results

4.1 Choice of DNN output

We compare the three possible outputs of our DNNs, *i.e.* direct cleaned image and shadow masks to be used as in Eqs. 3 (transmittance) and 4 (residual). Table 1 reports the metrics on the Ca-II synthetic dataset. For both DNNs, Eq. 3 is less stable than Eq. 4 during training, with more noisy loss curves. This more unstable training results in a poor convergence. Poorer results are observed in the sensitive cases of strong cloud coverage (*i.e.* low values in I_{ini}) where more subtle adjustments are needed in the mask's values to correct for the clouds. Predicting a residual mask $\hat{M}$ for Eq. 4 performs slightly better than the direct prediction of $\hat{I}_{\text{clean}}$, which, in rare cases, created artefacts that may be confused with solar structures. The better performance of Eq. 4 is in line with other works on residual prediction. Following these observations, we choose Eq. 4 to be used in the rest of this paper.

Table 2 Performance on synthetised datasets. Results are in format mean(std). SSIM and RMSE are as $e-2$.

Data	Method	PSNR ↑	SSIM ↑	RMSE ↓
Ca-II	[1] (15 failures)	21.90	94.8	8.6
	[2]	26.0	98.1	5.4
	[9]	23.2	92.7	7.3
	Fully sup.	**30.6(0.3)**	**98.9(0.1)**	**3.7(0.1)**
	C-GAN	30.0(0.4)	98.8(0.2)	3.9(0.1)
H-α	[1] (5 failures)	14.8	91.4	19.6
	[2]	23.3	98.4	7.0
	[9]	21.0	96.0	9.2
	Fully sup.	**28.6(0.3)**	**98.9(0.1)**	**4.5(0.2)**
	C-GAN	28.3(0.5)	98.9(0.2)	4.6(0.2)

4.2 Comparative evaluation of the methods

We compare the proposed DNNs against the SoTA [1, 2] and the sparse approximation method of [9] as non-learning baselines. [1] requires a clean temporal neighbour image (within 24 hours), which may not be always available. We consider these cases a failure and report their numbers. The first reflectance step of [9] is skipped, as reflection plays no role in solar imaging. We use the PSNR, SSIM and RMSE metrics to assess the quality of the predicted shadow-free images. For [1], the metrics are computed over non-failure cases. For the DNNs, they are averaged over 10 runs.

Table 2 presents results on the synthetic cloud datasets and Fig. 3 illustrates how the algorithms perform on real/synthetic clouds. [1] obtains the worse results on both imaging modalities. It struggles to preserve solar structures when their spatial scale is similar to the scale of clouds in at least one direction, as often happens in case of streaked textures. It is also less robust to strong cloud thickness. [2] also struggles with the spatial scale of streaked clouds. [9] fails to recover the underlying solar structures and results in poor contrast.

The proposed DNNs outperform non-learning methods on all datasets. On synthesised images, the DNN's respective metrics are not significantly different. Upon visual inspection of artificial and real images, FS produces a slightly more uniform contrast across the solar disk, while the C-GAN's discriminator, focusing on small patches, allows more diversity of contrasts. However, FS tends to over-compensate strong clouds resulting in a lower contrast there, while C-GAN benefits from enforcing a patch-wise visually pleasant output.

It is worth noting that in our real-cloud dataset, most clouds are streaked, while the DNNs are trained on the fluffy synthetic textures. Both DNNs generalised well to this new cloud texture. While, in some cases, faint streaks remain from the cleaned images, this happens for the strongest clouds and generally doesn't prevent the restoration of underlying solar structures. C-GAN performs slightly better than FS in this particular case of strong streaked clouds.

Both DNNs provide a detailed restoration of ARs, with the aforementioned contrast differences. Filaments suffer from a lower contrast, likely due to their dark appearance making them harder to distinguish from clouds. This is confirmed by the thicker and
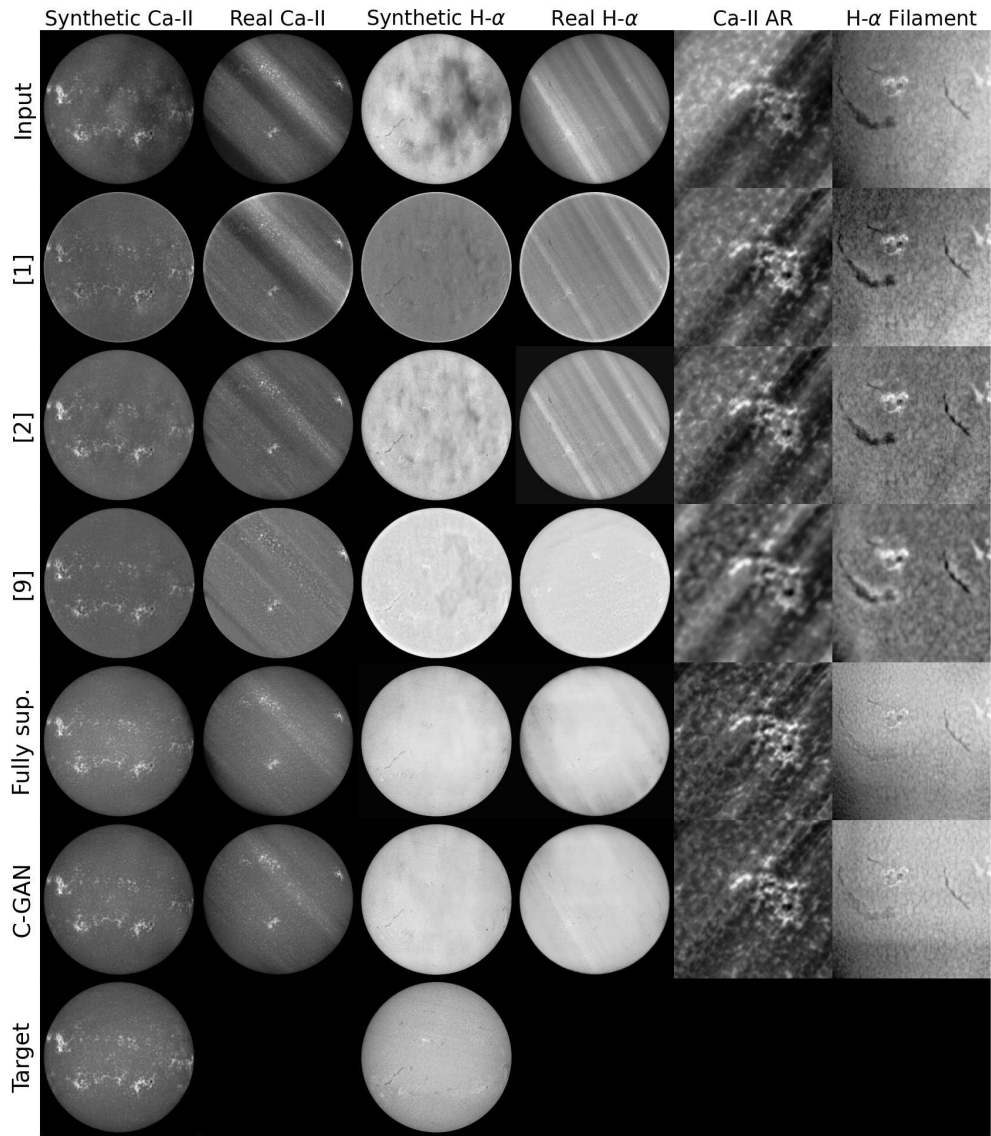
Fig. 3 Examples of removal of real and synthetic clouds from Ca-II and H-α images by compared algorithms.

therefore blurrier ones being the most affected (erased) by this contrast reduction. C-GAN preserves slightly better these thick filaments.

5 Conclusion

We investigated deep learning methods for removing large cloud shadow contaminants from ground-based solar imagery. We adopted the U-Net-style architecture for

more detailed outputs. We compared three possible inference methods (*i.e.* direct and through a transmittance or residual mask), and two training setups: fully-supervised and C-GAN. The DNNs obtained better results than the domain's SoTA and a non-leaning method based on sparse approximation, with better restoration of the underlying solar structures, and are the new SoTA. The training setups have different strengths and weaknesses: full supervision produces more homogeneous solar disks, and C-GAN better handles strong clouds. Future work may further focus on filament restoration, and evaluate the effect of cloud removal on a subsequent image analysis *e.g.* the automatic detection of solar structures.

Supplementary information.   Not applicable

Declarations

The data used in this study is available at https://doi.org/10.5281/zenodo.7684200. The code is available at https://github.com/jaypmorgan/cloud-removal.

Adeline Paiement and Jean Aboudarham contributed to the study conception and design. Data preparation was performed by Jay Paul Morgan, experiments were performed by Amal Chaoui and Jay Paul Morgan, analysis was performed by Amal Chaoui, Jay Paul Morgan and Adeline Paiement. All authors contributed to the manuscript.

References

[1] Feng, S., Lin, J., Yang, Y., Zhu, H., Wang, F., Ji, K.: Automated Detecting and Removing Cloud Shadows in Full-Disk Solar Images. New Astronomy **32**, 24–30 (2014)

[2] Fuller, N., Aboudarham, J.: Automatic detection of solar filaments versus manual digitization. In: Knowledge-Based and Intelligent Information and Engineering Systems (KES), pp. 467–475 (2004)

[3] Chen, Z., Zeng, Z., Shen, H., Zheng, X., Dai, P., Ouyang, P.: DN-GAN: Denoising Generative Adversarial Networks For Speckle Noise Reduction in Optical Coherence Tomography Images. Biomedical Signal Processing and Control **55**, 101632 (2020)

[4] Zheng, J., Liu, X.-Y., Wang, X.: Single Image Cloud Removal Using U-Net and Generative Adversarial Networks. IEEE Transactions on Geoscience and Remote Sensing **59**(8), 6371–6385 (2021)

[5] Finlayson, G.D., Hordley, S.D., Lu, C., Drew, M.S.: On the removal of shadows from images. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 59–68 (2006)

[6] Yang, Q., Tan, K.H., Ahuja, N.: Shadow removal using bilateral filtering. IEEE Trans. Image Process. **21**, 4361–4368 (2012)

[7] Guo, R., Dai, Q., Hoiem, D.: Paired regions for shadow detection and removal. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 2956–2967 (2013)

[8] Yu, X., Li, G., Ying, Z., Guo, X.: A new shadow removal method using color-lines. In: Computer Analysis of Images and Patterns, pp. 307–309 (2017)

[9] Sasi, R.K., Govindan, V.K.: Shadow Removal Using Sparse Representation Over Local Dictionaries. Engineering Science and Technology, an International Journal **19**(2), 1067–1075 (2016)

[10] Fan, H., Han, M., Li, J.: Image shadow removal using end-to-end deep convolutional neural networks. Applied Sciences (2019)

[11] Qu, L., Tian, J., He, S., Tang, Y., Lau, R.W.H.: Deshadownet: A multi-context embedding deep network for shadow removal. In: Computer Vision and Pattern Recognition, pp. 2308–2316 (2017)

[12] Wang, J., Li, X., Yang, J.: Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In: Computer Vision and Pattern Recognition, pp. 1788–1797 (2018)

[13] Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Computer Vision and Pattern Recognition, pp. 5967–5976 (2017)

[14] Malherbe, J.-M., Dalmasse, K.: The new 2018 version of the meudon spectroheliograph. Solar Physics **294** (2019)

[15] pngkey.com: Transparent Cloud Texture Download. https://www.pngkey.com/maxpic/u2a9o0u2i1i1y3r5/

[16] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention (MICCAI) (2015)