

# BLUR-SPECIFIC NO-REFERENCE IMAGE QUALITY ASSESSMENT FOR MICROSCOPIC HYPERSPECTRAL IMAGE FOCUS QUANTIFICATION

Laura Quintana, Samuel Ortega, Himar Fabelo, Gustavo M. Callico

Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria, Spain

## ABSTRACT

Hyperspectral (HS) imaging is a novel technique that allows for better understanding of materials, being an improvement in multiple applications. However, one of its main drawbacks is the focus assessment. This issue has already been covered for RGB images. Thus, in this study, it is going to be revised several no reference RGB image quality assessment algorithms (NR-IQA). To this aim, a HS image database was created by capturing different images of the same specimen at different working distances. NR-IQA algorithms were tested over monochromatic images extracted from the HS images. Additionally, a study through each independent wavelength was carried out. Results showed that some algorithms perform better for calibration samples and another ones for biological samples. Furthermore, focus differences were found in the initial and final wavelengths. In conclusion, HS image results are similar to the one obtained for RGB images but, there is still room for improvement.

**Index Terms**— *Hyperspectral Imaging, hyperspectral microscopy, NR-IQA, focus quantification*

## 1. INTRODUCTION

Nowadays, hyperspectral imaging (HSI) is an emerging technology, being its main strength, the large number of adjacent spectral bands which are possible to acquire and handle. This allows the detection of materials in a scene with higher precision than conventional RGB imaging. Thus, this type of images and the type of information they provide have multiple applications in fields such as medicine [1] or agriculture [2].

However, images may suffer from a variety of distortions, especially focus became really important when working with hyperspectral (HS) microscopic images [3]. Most common problems are related with the depth of field (DOF). DOF is the distance between the nearest and the farthest objects that are in acceptably sharp focus in an image (Figure 1.a).

Equation (1) shows the relationship between the different parameters involved in the DOF for a given circle of confusion ( $c$ ), focal length ( $f$ ), f-number ( $N$ ), and working distance ( $WD$ ). Unfocused images may take place when changes are made on the  $WD$  or the  $f$ .

$$DOF = \frac{2WD^2 Nc}{f^2} \quad (1)$$

On one hand, changes in the  $WD$  may occur because of non-flat microscopic samples. This results in different areas focused at different  $WD$ s. Figure 1.b shows a ray transfer matrix analysis of these kind of distortion and Figure 1.c displays a real example when trying to image a non-flat grape leaf sample using our HS microscope at 20 $\times$  of magnification.

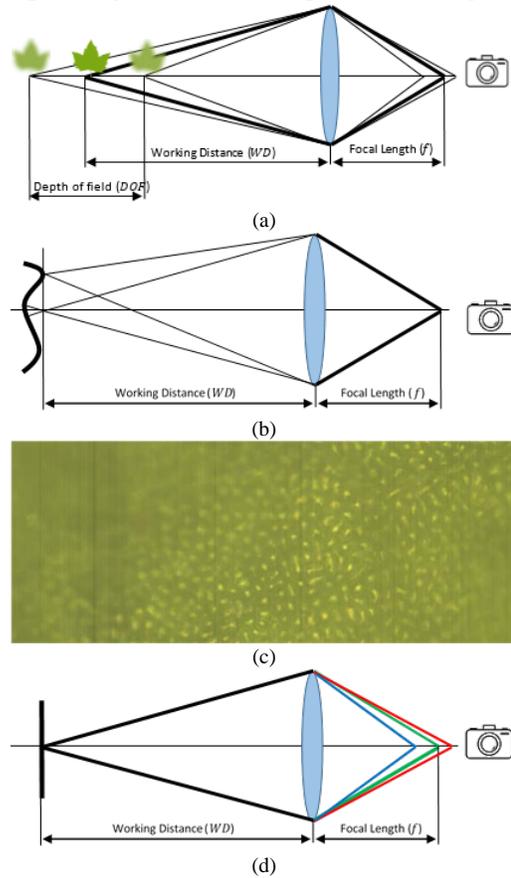


Figure 1. (a) Depth of field illustration. (b) Working distance shift problem. (c) Example of an unfocused grape leaf due to different  $WD$ s in the scene captured at 20 $\times$ . (d) Chromatic focal shift problem.

On the other hand, changes in  $f$  may occur when different wavelengths are involved. A common optical problem occurs when a lens is unable to bring all wavelengths to the same focal plane. This phenomenon is called chromatic aberration (CA) (Figure 1.d).

Sharp images are crucial in most HSI applications, such as cancer diagnosis in histological samples [4]. Therefore, image quality assessment (IQA) becomes a key technique.

Objective IQA can be categorized into full-reference IQA (FR-IQA, e.g., Peak signal to noise ratio (PSNR)[5]), reduced-reference IQA (RR-IQA, e.g., Free-energy-based distortion metric (FEDM)[6]), and no-reference IQA (NR-IQA, e.g., Local phase coherence (LPC)[7]). Due to the unavailability of ground truth images in blur image applications, NR-IQA is preferable but also more challenging.

There are numerous NR-IQA algorithms for RGB images in the literature, such as dubbed blind/referenceless image spatial quality evaluator (BRISQUE) [8] or blind image blur evaluator (BIBLE) [9]. However, there is a limited state-of-the-art for these algorithms applied to HS images. The objective of this study is to analyze NR-IQA algorithms developed for RGB images in HS microscopic images.

## 2. METHODOLOGY

In this section, we report the methodology followed in the dataset acquisition and further describe the different NR-IQA algorithms employed to evaluate our samples.

### 2.1. Database Acquisition

The instrumentation employed in this study consists of an HS camera coupled to a conventional light microscope (Figure 2). The microscope is an Olympus BX-53 (Olympus, Tokyo, Japan). The HS camera is a Hyperspec® VNIR A-Series from HeadWall Photonics (Fitchburg, MA, USA), which is based on an imaging spectrometer coupled to a CCD (Charge-Coupled Device) sensor, the Adimec-1000m (Adimec, Eindhoven, Netherlands). This HS system works in the visual and near-infrared (VNIR) spectral range from 400 to 1000 nm with a spectral resolution of 2.8 nm, sampling 826 spectral channels and 1004 spatial pixels. The push-broom camera performs spatial scanning to acquire an HS cube with a mechanical stage (SCAN, Märzhäuser, Germany) attached to the microscope, which provides accurate movement ( $\pm 3 \mu\text{m}$  accuracy) of the specimens in the 3 axes directions:  $x, y, z$ . The objective lenses are from the LMPLFLN family (Olympus, Tokyo, Japan), which are optimized for infra-red (IR) observations. The light source is a 12 V, 100 W halogen lamp. This system was previously employed to histological HS analysis of brain cancer samples [4].

The database is divided into calibration and biological microscopic samples. In order to acquire a dataset for focus analysis, different images were acquired varying the  $WD$ , i.e. modifying the distance between the specimen and the objective lens. For each specimen 10 images were recorded, being the first one captured in the subjective point of focus. From there, 9 additional images were captured distancing 0.1 mm each time from the previous record. For each HS image captured, white and dark reference images were also taken. Later, calibration of the HS cubes was performed following conventional HS calibration, presented in [3].

A subjective score ( $SC$ ) from 0 to 9 was given to each image according with the order in which it was captured. Images were taken with an increment distance ( $\Delta z$ ) of 0.1 mm respect to the previous one. Therefore,  $WD$  of each image can be calculated following the equation (2), knowing that  $WD_0$  is the point where the HS image is focused by visual inspection. In Figure 4 it can be seen an example of three specimens with the 10 different  $WD$ s.

$$WD = WD_0 + (SC \times \Delta z) \quad (2)$$

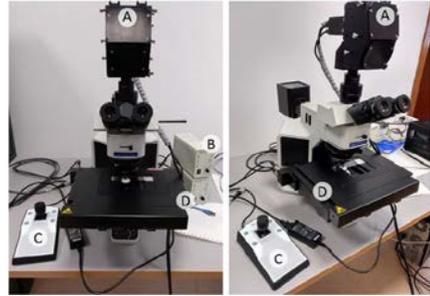


Figure 2. Microscopic HS acquisition system. a) HS camera, b) Halogen light source, c) Positioning joystick and d) XY linear stage [4].

### 2.2. NR-IQA Algorithms

Following the classification made by [10] over NR-IQA methods applied to RGB images, they can be divided into learning free and learning based methods. Here, we are focused studying learning free methods, which can be further divided as shown in Figure 3. From each of the four learning free NR-IQA algorithm categories, the best algorithm with available code for monochromatic images was chosen.

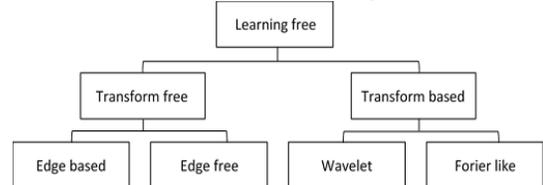


Figure 3. Classification of NR-IQA learning free methods.

#### 2.2.1. Edge Based

Perceptual Sharpness Index (PSI) [7] is based on the statistical analysis of local edge gradients. The first step of the algorithm is to select the most significant edges in the image through an adaptive edge selection procedure. Second, the edge widths of the selected edges are computed by an edge width measurement based on diagonal edge gradients. Third, edge widths above the just noticeable blur (JNB) width are subtracted by the edge slopes. Finally, the local sharpness map is deduced by applying the above three steps in a block-wise way. Since the sharpest regions in an image are most related to human sharpness perception, the global image quality score is determined by the highest  $q$ th percentile average of the local sharpness values.

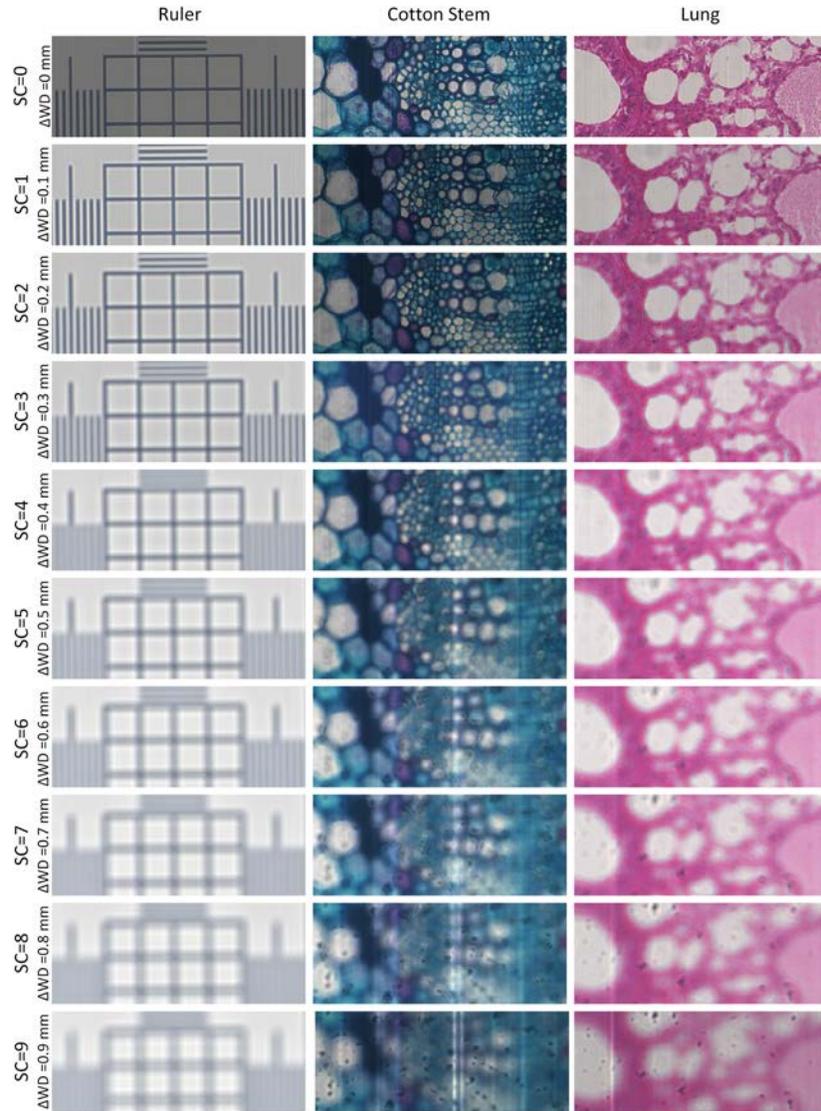


Figure 4. Examples of the RGB representation of the HS images in our database. Column 1 shows the calibration sample, a ruler. Columns 2 and 3 show biological samples, a cotton stem, and a lung histology sample, respectively. Rows show the continuous blurring of the sample as we get further away from the focus point, incrementing  $\Delta WD$  and so, the SC ( $\Delta WD = SC \times \Delta z$ ).

### 2.2.2. Edge Free

Bahrami and Kot [9] defined the maximum local variation (MLV) of each pixel as the maximum intensity variation of the pixel with respect to its 8-neighbors. Since the human vision system is more sensitive to higher variations regions, high variations in the pixel intensities are a better indicator of sharpness. Taking that into account, the pixels' MLVs are subjected to a weighting scheme where heavier weights are assigned to greater MLVs to make the tail end of MLV distribution thicker. Thus, the MLV distribution becomes more discriminative for different blur degrees. Finally, the metric to measure sharpness is taken from the standard deviation of the weighted MLV distribution.

### 2.2.3. Wavelet

Local Phase Coherence (LPC) structures are shown in step edges, but not in blur images. Following this idea, Hassen et al. [11] proposed a blur-specific NR-IQA method based on the strength of this metric near edges and lines, the LPC based sharpness index (LPC-SI). To its calculation, an image must pass through 3-scale 8-orientation log-Gabor filters. Then, LPC strength is computed at one orientation and one spatial location and later, at each spatial location. However, unlike other algorithms, LPC-SI does not employ block-based computation. They proposed an efficient algorithm that largely simplifies the LPC computation, making it easily applicable in practical applications.

### 2.2.4. Fourier like

Vu and Chandler [12] proposed a spectral and spatial sharpness measure (S3). According to the reduction of high-frequency components in blur images, the spectral measure  $S1(x)$  of a block  $x$  is initially calculated, then rectified by a sigmoid function to account for human visual system (HVS). To further consider the contrast effect, the spatial measure  $S2(x)$  of a block  $x$  is calculated based on the local total variation. Then, the overall sharpness map  $S3$  of the image  $I$  is obtained by a geometric mean of spectral and spatial measures in a block-wise way. Finally, to consider the human sharpness perception, the overall sharpness score is calculated as the average of the largest 1% values of the overall sharpness map.

### 2.3. Evaluation Criteria

From each of the aforementioned algorithms, an objective score was obtained when applying them to our HS microscopic image database. However, for the evaluation of the obtained results, Video Quality Experts Group (VQEG) [13] suggest to map each objective score ( $o$ ) to a subjective score ( $s$ ) using equation (4), where  $\tau_1 = \max(s)$ ,  $\tau_2 = \min(s)$ ,  $\tau_3 = \text{mean}(o)$ , and  $\tau_4 = \text{std}(o)/4$ .

$$F(o) = \frac{\tau_1 - \tau_2}{1 + e^{-\frac{o - \tau_3}{\tau_4}}} + \tau_2 \quad (3)$$

Afterwards, three evaluation metrics are calculated to evaluate the method's performance: spearman's rank-order correlation coefficient (SRCC) which employs a monotonic function; pearson's linear correlation coefficient (PLCC) which is a measure of the linear correlation after a nonlinear mapping and; root-mean-square error (RMSE) which is used to measure the differences after the nonlinear mapping. To obtain a good NR-IQA algorithm, the values of SRCC and PLCC are close to 1, while the value of RMSE is close to 0.

These algorithms were evaluated firstly in a set of monochromatic images (computed from the mean of all the wavelengths of the calibrated HS image) at different  $WDs$ , since there is dependency between  $WD$  and focus. Then, they were applied in each band of each image independently to show the dependency between the wavelength and the focus (to prove the existence of any CA).

The different algorithms were implemented in MATLAB® R2019b in a Windows environment (Microsoft Windows 10) with an Intel i5-4210U 1.70 GHz CPU and 16 GB RAM.

## 3. RESULTS

In this section, the results obtained using the previously described methods in our HS database are shown. The performance comparison between different  $WD$  images of the calibration and the biological samples, are shown in Table 2 and Table 3 respectively. Best values are in boldface. Although best methods shown a good correlation (SRCC >

0.8 and PLCC > 0.8), there are still some which did not (SRCC < 0.5 and PLCC < 0.5). As PSI is based on the local edge gradients, it shows good results for the biological samples, where edges are heterogeneous (there is edges everywhere). However, LPC performs better in calibration samples, where there is a homogeneous distribution of edges, due to its computation at several orientations and spatial locations. Furthermore, since later application of these methods is framed in real-life, computational time is also a good estimator of the algorithm's quality. Although overall methods have a small execution time, S3 does not, and it was not considered later on.

Table 1. Calibration sample results

Method	SRCC↑	PLCC↑	RMSE↓	Time (s)
PSI	0.491	0.564	5.931	0.050
MLV	0.806	0.843	2.324	0.113
LPC	<b>0.903</b>	<b>0.871</b>	<b>2.172</b>	<b>1.900</b>
S3	0.576	0.797	2.492	26.790

Table 2. Biological sample results

Method	SRCC↑	PLCC↑	RMSE↓	Time (s)
PSI	<b>0.812</b>	<b>0.810</b>	<b>2.266</b>	<b>0.052</b>
MLV	0.423	0.510	3.619	0.115
LPC	0.441	0.433	3.491	1.903
S3	0.629	0.611	5.792	29.528

NR-IQA methods, but S3 because of its large execution time, were also applied to each wavelength separately. Results are shown in Figure 5. As we can see, correlation indexes are constant for the central wavelengths (~450 to 900 nm) but results seem to be worsened for the initial and final ranges (400-450 nm and 900-1000 nm) in the three methods. This may be produced due to CA and/or because of the low performance of the HS sensor in these spectral ranges as studied in [14] where the same HS camera was employed.

## 4. CONCLUSION

In this work, it has been developed a methodology to capture a focus assessment dataset and followed to acquire 110 HS images with its calibration data. Afterwards, taking advantage of this dataset, it was demonstrated that blur-specific NR-IQA methods developed for RGB images can also quantify focus of HS images. NR-IQA algorithms have been used over HS monochromatic images showing results similar to the ones obtained with regular RGB images. However, some differences were found between calibration and biological samples. This can be due to the limited size of the database we are working with. Moreover, these algorithms were also employed in each wavelength separately and it was shown that focus difference is shown in the initial and final wavelengths. A reason to that could be CA and/or the low performance of the HS sensor at these wavelengths, further investigation is necessary. As [10] also

concluded, there is still a large space for designing robust and effective blur-specific NR-IQA methods for realistic blur images. Future works will include the development of an optimized algorithm for HS images blur detection, targeting z-stacking applications to guarantee the acquisition of high-quality microscopic HS images from non-flat specimens.

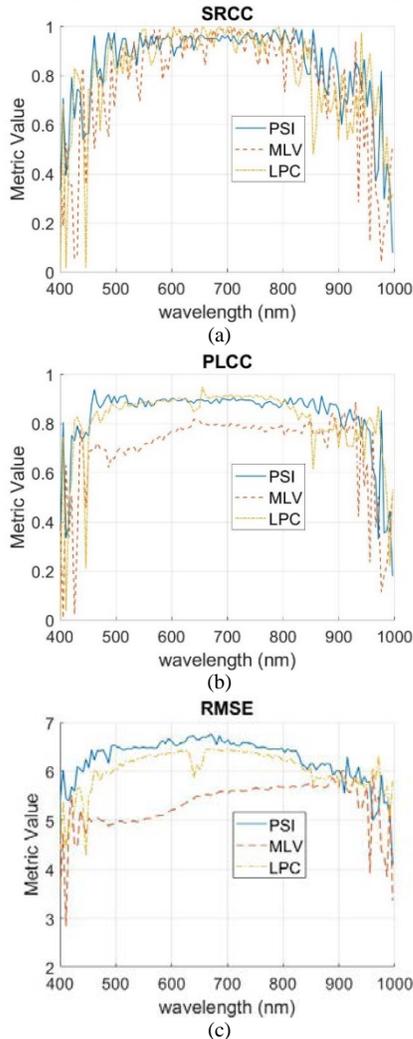


Figure 5. Graph of correlation value vs wavelength for the calibration sample.

## 5. ACKNOWLEDGMENT

This work has been supported in part by the Spanish Government and European Union (FEDER funds) as part of support program in the context of Distributed HW / SW Platform for the Intelligent Processing of Heterogeneous Sensory Information in Supervision Applications of Large Natural Spaces (PLATINO) project, under contract TEC2017-86722-C4-1-R, and by the Canary Islands Government through the ACIISI (Canarian Agency for Research, Innovation and the Information Society), ITHACA project “Hyperspectral Identification of Brain Tumours” under Grant Agreement ProID2017010164.

## 6. REFERENCES

- [1] M. Halicek, H. Fabelo, S. Ortega, G. M. Callico, and B. Fei, “In-Vivo and Ex-Vivo Tissue Analysis through Hyperspectral Imaging Techniques: Revealing the Invisible Features of Cancer,” *Cancers (Basel)*, vol. 11, no. 6, p. 756, May 2019, doi: 10.3390/cancers11060756.
- [2] M. Bacco *et al.*, “Smart farming: Opportunities, challenges and technology enablers,” in *2018 IoT Vertical and Topical Summit on Agriculture - Tuscany (IOT Tuscany)*, 2018, pp. 1–6, doi: 10.1109/IOT-TUSCANY.2018.8373043.
- [3] S. Ortega *et al.*, “Hyperspectral Push-Broom Microscope Development and Characterization,” *IEEE Access*, vol. 7, pp. 122473–122491, 2019, doi: 10.1109/ACCESS.2019.2937729.
- [4] S. Ortega *et al.*, “Hyperspectral Imaging for the Detection of Glioblastoma Tumor Cells in H&E Slides Using Convolutional Neural Networks,” *Sensors*, vol. 20, no. 7, p. 1911, Mar. 2020, doi: 10.3390/s20071911.
- [5] A. Hore and D. Ziou, “Image Quality Metrics: PSNR vs. SSIM,” in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 2366–2369, doi: 10.1109/ICPR.2010.579.
- [6] Guangtao Zhai, Xiaolin Wu, Xiaokang Yang, Weisi Lin, and Wenjun Zhang, “A Psychovisual Quality Metric in Free-Energy Principle,” *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012, doi: 10.1109/TIP.2011.2161092.
- [7] C. Feichtenhofer, H. Fassold, and P. Schallauer, “A Perceptual Image Sharpness Metric Based on Local Edge Gradient Analysis,” *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 379–382, Apr. 2013, doi: 10.1109/LSP.2013.2248711.
- [8] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-Reference Image Quality Assessment in the Spatial Domain,” *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
- [9] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, “No-Reference Image Blur Assessment Based on Discrete Orthogonal Moments,” *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 39–50, Jan. 2016, doi: 10.1109/TCYB.2015.2392129.
- [10] D. Li and T. Jiang, “Blur-Specific No-Reference Image Quality Assessment: A Classification and Review of Representative Methods,” 2019, pp. 45–68.
- [11] R. Hassen, Zhou Wang, and M. M. A. Salama, “Image Sharpness Assessment Based on Local Phase Coherence,” *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013, doi: 10.1109/TIP.2013.2251643.
- [12] C. T. Vu, T. D. Phan, and D. M. Chandler, “S3: A Spectral and Spatial Measure of Local Perceived Sharpness in Natural Images,” *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934–945, Mar. 2012, doi: 10.1109/TIP.2011.2169974.
- [13] VQEG, “Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment,” Video Quality Experts Group, 2000.
- [14] H. Fabelo *et al.*, “In-Vivo Hyperspectral Human Brain Image Database for Brain Cancer Detection,” *IEEE Access*, vol. 7, pp. 39098–39116, 2019, doi: 10.1109/ACCESS.2019.2904788.