



HAL
open science

SATELLITE IMAGE QUALITY ASSESSMENT USING DEEP LEARNING

Bouchra Harnoufi, Ségolène Bourrienne, Mathias Ortner, Renaud Fraise

► **To cite this version:**

Bouchra Harnoufi, Ségolène Bourrienne, Mathias Ortner, Renaud Fraise. SATELLITE IMAGE QUALITY ASSESSMENT USING DEEP LEARNING. Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle (APIA2021), Jul 2021, Bordeaux, France. <hal-03874228>

HAL Id: hal-03874228

<https://hal.science/hal-03874228v1>

Submitted on 27 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

SATELLITE IMAGE QUALITY ASSESSMENT USING DEEP LEARNING

Bouchra Harnoufi, Ségolène Bourrienne, Mathias Ortner, Renaud Fraisse

Airbus Defence and Space, Toulouse (France)

ABSTRACT

The Modulation Transfer Function (MTF) is one of the key indicators regarding Image Quality of Earth Observation systems. It characterizes the level of contrast that can be maintained by the optical system and is monitored during the whole life of the satellite. Due to the strong acquisition constraints as well as the gradually complexity of its estimation with future systems, it becomes necessary to increase the reactivity with a method free of acquisition constraints. In this paper, we present a model able to estimate the absolute MTF (or blur) level as well as its prediction uncertainty without any reference image or user parameter using deep learning techniques.

Index Terms— Satellite Imagery, No-Reference Image Quality, Focus, Defocus, Deep Learning

1. INTRODUCTION

The Modulation Transfer Function (or MTF) of an optical system is a measurement of its ability to reproduce various levels of details (spatial frequencies) from the object to the image. More formally, it indicates its capability to transfer contrast at a particular resolution: it shows how well frequency information is transferred from object to image. As a function of spatial frequency, its unit is the ratio of the image contrast over the object contrast (Fig. 1). For high frequencies it is limited by the optical instrument and its analog / digital chain. Hence, it is an important part of the system requirements during the satellite design phase and is the subject of a close monitoring during the Maintenance in Operational Condition (MCO) phase where its level is checked and corrected with refocusing. It is possible to refocus the instrument in orbit by modifying the temperature of the telescope. The range of variation depends on the instrument. At each temperature change, the MTF level is measured: when the highest value of MTF is reached at Nyquist frequency (0.5 when the frequencies are normalized) $MTF_{Nyquist}$, the defocus is corrected (Fig. 2).

Currently, the MTF estimation for a given temperature is notably achieved via dedicated acquisitions on specific areas containing well-known ground patterns (Fig. 3). There are a few patterns around the earth. Viallefont-Robinet et al. [1] compares the MTF measurements using the edge method. This method is widely used. However, for on-orbit MTF assessment, it involves strong constraints.

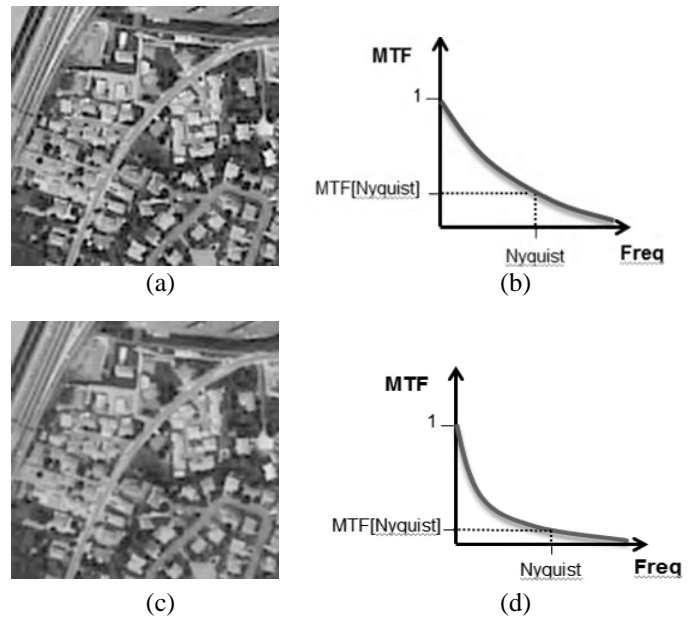


Fig. 1 (a) and (c) represent the same image with the same resolution but with different levels of MTF (or blur). (b) and (d) correspond to the MTF for image (a) and (c) respectively. The lower the MTF is, the less contrast is transferred, and thus the blurrier the image is (in this case (c) is blurrier than (a)).

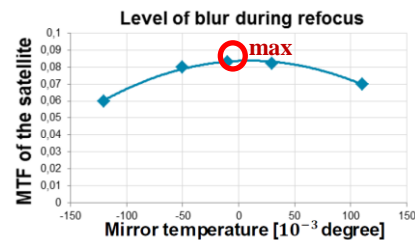


Fig. 2 Evolution of the MTF at Nyquist frequency during thermal refocus. At each temperature variation, the MTF is computed and in particular at Nyquist frequency. The highest MTF value corresponds to the defocus correction.

First of all, a few acquisitions are needed for each focus point and can take a long time depending on the “revisit period” of the satellite. These patterns must be visible which implies cloud-free acquisitions. Consequently, the mission is



Fig. 3 Ground pattern examples (Baoutou). These patterns are currently used for MTF estimation. There are a few around the earth.

interrupted in order to achieve such specific acquisitions. Besides, with new satellite generations, instruments tend to defocus due to their sensitive materials to heat. Monitoring the MTF is thus challenging while the need is increasing.

The main aims are thus to reduce the mission time spent for refocus operations increasing the reactivity and to avoid mission interruption. Instead of manually estimating the MTF using dedicated acquisitions, a new deep learning neural network model is proposed to estimate the MTF using nominal acquisitions in natural scenes. Recent advances in deep learning approaches are tackling the problem of image quality assessment mainly through comparing images notably leading to a relative measure or using reference images (full references or reduced references). The applications of such techniques still require adjustments to satellite imagery. Instead, we used a classification approach in order to get the level an image belongs to without any reference. The next part details more specifically the methodology followed

2. DEEP LEARNING FOR NO-REFERENCE SATELLITE IMAGE QUALITY ASSESSMENT

Given a set of images each associated with its own defocus MTF level, the objective is to assess the blur level of each image. Comparing two images in order to rank them using notably siamese networks [2] was our first approach. However, it requires a threshold selection which may need user parameters. We chose to use a discrete representation of the defocus MTF levels and to implement a classifier in order to predict the blur level instead of using a regression more suited for continuous variables. Classifiers have demonstrated very powerful results especially in images classification using Convolutional Neural Networks. Our approach is similar to Yang et al. method [3] which assesses microscope image focus quality. However, instead of using a ranked probability score loss function, we used the cross entropy loss function. The model uses this categorical cross entropy to learn to give a high probability to the correct blur level and a low probability to the other levels. We wanted to evaluate if the model is able to learn itself an order without

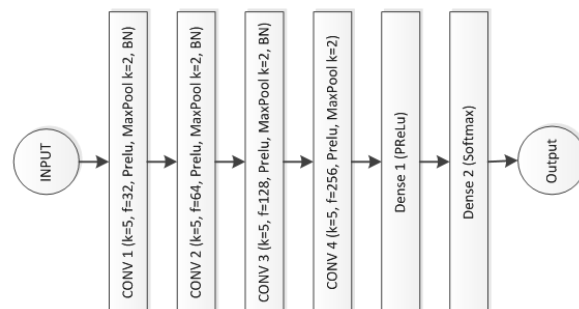


Fig. 4 Neural network model architecture. The convolutional blocks are defined by **k**: kernel size, **f**: feature maps, the activation function, the max pooling layer and **BN**: batch normalization, A probability distribution over the discrete defocus classes is predicted for each input image.

prior knowledge: indeed, the human eye cannot distinguish very close blur levels. Additionally, we used batch normalization regularization technique which significantly improves the results. The network consists of four convolutional blocks followed by two dense blocks. A convolutional block is composed of a convolution layer with a Parametric Rectified Linear Unit (PReLU) activation function, a max-pooling layer and a batch normalization layer except for the last one. The output feature maps are then fed to fully connected layers after having been flattened to finally produce the probability distribution over the classes considered. The softmax activation function ends the neural network since it highlights the target level and normalizes the outputs so that they sum to 1. Consequently, they can be directly treated as probabilities over the output (Fig. 4).

3. EXPERIMENTS

3.1. Data

Data needed are the different MTF levels and the images as described in the previous section. A first simulator produces realistic defocus MTF levels equidistant at Nyquist frequency meaning that from a level to the very next level the amount of “blur” added is constant (from the top higher curve to the lower one). In Fig. 5 each curve represents a MTF level: the lower the point at Nyquist is, the less contrast is transferred and consequently the blurrier is the image. The MTF range considered at Nyquist frequency is [05.25, 12.75]. Two separation powers have been considered: 10% (or 0.5 point MTF) with 16 defocus levels (equidistant at Nyquist frequency) generated and 5% (or 0.25 point MTF) with 31 defocus levels simulated. In both cases, the human eye struggles to distinguish an image with a given defocus MTF level and the same image with the very next defocus MTF level.

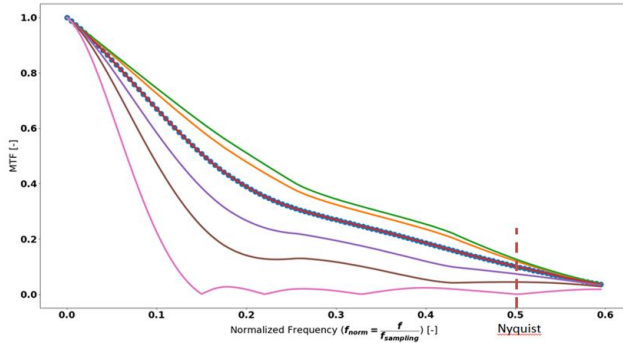


Fig. 5 Different MTF level curves representation generated by the MTF simulator. Each curve represents the level of contrast restored and corresponds to a temperature variation. At Nyquist frequency, the lower the point is, the blurrier the image is (less contrast restored).

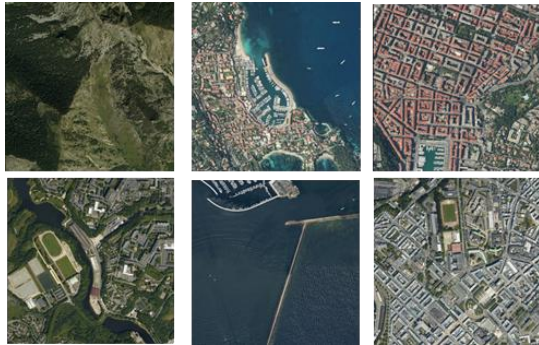


Fig. 6 Examples of images used during the training and the test phase from France. These images contain various frequencies: higher in downtown areas and lower in seas areas for instance.

The available images are aerial images with a resolution of 10 cm. A radiometric simulator has also been developed in order to:

- apply the MTF (blur level) to the images
- resample images to 50 cm: the target resolution
- perform corrections and conversions: remove gamma correction, convert to 12-bit images ...

The model has been trained using simulated images from France. These images contain various land cover types such as: some rural regions, homogeneous structures, seas, city centers, buildings (Fig. 6). The diversity of land covers helps evaluating the MTF on a wide range of spatial distributions. To evaluate the performances we used some other simulated images not used in the training data set. Each image is divided into adjacent tiles of size 128 x 128 pixels from which each defocus MTF level is applied to. The dataset is balanced: there are as many examples for each defocus MTF level. Given training examples of 12-bit 128 x 128 pixels input images patches and the corresponding defocus MTF level, the model predicts the

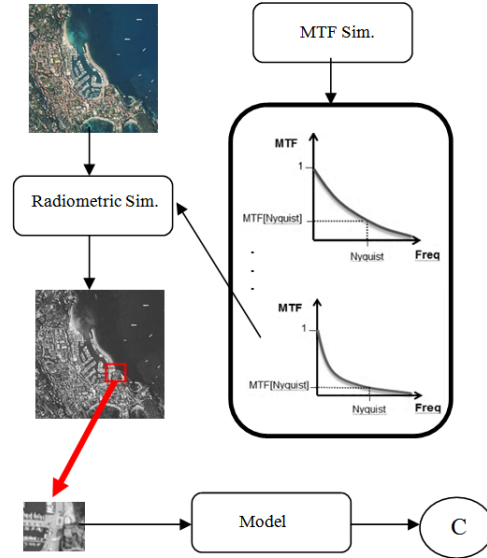


Fig. 7 A MTF simulator generates defocus MTF levels given a MTF range at Nyquist frequency and a separation power. These defocus MTF levels are applied to each image. The images are then divided into tiles to be fed to the neural network. The model output “C” corresponds to the predicted defocus MTF level.

probability distribution over these classes (Fig. 7). The optimizer is Adam to update the network weights with a learning rate value of $5e-6$. The model has been trained during 300 epochs and the framework used is PyTorch, in a cloud environment, specifically in a virtual machine containing one GPU NVIDIA Tesla P100.

3.2. Results

The most probable class and the prediction certainty given by the trained model can be visualized for each image patch as a colored border, with a color indicating the predicted class (defocus MTF level) and the transparency denoting the certainty as Yang et al [3] representation (Fig. 8). Tests have been performed using 16 and 31 classes.

Even if the model has been trained not using a specific ordered loss function, it learnt the order (Fig. 9). Indeed, when it fails predicting the right class, it predicts the neighbor class. Moreover, the model is able to produce interesting results even in areas it has not seen before and seems to be able to generalize (no Lyon city image was in the training data set Fig. 8). Another behavior we wanted to check is regarding homogenous areas like seas or rivers where we expect low certainty. Without pre-processing techniques including filtering some areas, the model predicts a defocus MTF level with low probability in areas where frequencies are very low: in Fig. 8 the river area has a very high transparency traducing low certainty. Finally, the separation power obtained using this classifier is 0.4 point

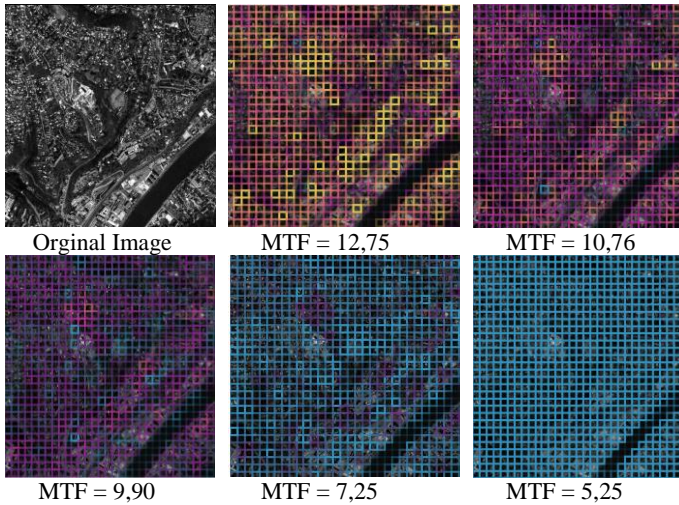


Fig. 8. Examples of result in Lyon city, (a city not contained in the training dataset). Each patch has a color and a transparency representing the MTF level and certainty respectively. The certainty is very low in the river area.

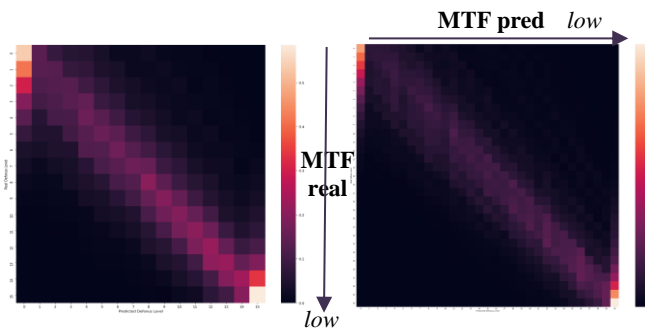


Fig. 9 Covariance matrix: the left one corresponds to 16 ordered blur levels and the right one, to 31. Without using a specific order loss function, the model is able to learn the rank. Indeed, it tends to predict either the true blur level or the previous or next one.

MTF which is equivalent to 8% when $MTF_{Nyquist}=0.05$ and 3% when $MTF_{Nyquist}=0.12$. Given that it is very close and sometimes even better than manual methods, these results are very promising.

4. DISCUSSION

One of the main advantages of this approach is the simulators developed. Indeed, both the defocus MTF level simulator and the radiometric simulator enable us to generate data with their ground truth (i.e. their defocus MTF level) without having to label them manually. It offers also the opportunity to simulate a large number of data which is a requirement when training a deep learning neural network. Moreover, future work includes noise robustness (necessary to real data applications) and can also be simulated. Another

improvement is data augmentation. Yang et al. [3] demonstrates the importance in using this technique to train a model. It ensures the diversity of examples learnt by the model. In addition, as our model performs better in low MTF levels, we can also add more examples of images with high defocus MTF levels in the training dataset to help the model distinguish them.

Regarding the neural network model architecture, several optimizations can improve the performances. First of all, we can also use a ranked probability score loss to help the model to focus on learning the MTF level order. Furthermore, as batch normalization has improved our result (without it, the model does not converge), we could use new classification architectures [4] that do not include batch normalizations and have been designed for models sensitive to data statistics. Finally, the requirement of using a defined size patch (here 128 x 128) may be not necessary if the architecture model is fully convolutional.

An extension of this application is to apply such model while performing other image enhancements so as to know how modifying another image parameter affects the blur level of the image. Some algorithms should not alter the blur level such as noise reduction. A measure using our method can confirm if it respects this property

5. CONCLUSION

We have proposed a deep learning model able to assess the MTF or blur level of any satellite image without needing dedicated acquisitions. To train and validate such method, we have also developed a simulator capable of simulating images at different defocus MTF levels. Experiments conducted in this frame showed promising results in terms of accuracy of MTF levels as well as generalization to unseen images. It opens up a wide range of new applications needing a MTF measure and several improvements have been identified to improve performances before being challenged to real data.

6. REFERENCES

- [1] Viallefont-Robinet, Françoise, et al. "Comparison of MTF measurements using edge method: towards reference data set." *Optics express* 26.26 (2018): 33625-33648.
- [2] Niu, Yuzhen, et al. "Siamese-network-based learning to rank for no-reference 2D and 3D image quality assessment." *IEEE Access* 7 (2019): 101583-101595.
- [3] Yang, Samuel J., et al. "Assessing microscope image focus quality with deep learning." *BMC bioinformatics* 19.1 (2018): 1-9.
- [4] Brock, Andrew, et al. "High-Performance Large-Scale Image Recognition Without Normalization." *arXiv preprint arXiv:2102.06171* (2021).