Journal: AMT
Title: Analysis of 2D airglow imager data with respect to dynamics using machine learning
Author(s): René Sedlak, Andreas Welscher, Patrick Hannawald, Sabine Wüst, Rainer Lienhart, and Michael Bittner
MS No.: amt-2023-25
MS Type: Research Article

General Comments

1. The authors describe the development and results of applying a neural network based classification method to sequences of OH* nightglow images obtained from a very high resolution InGaAs-based camera. The goal of the work was to find an efficient method of identifying images that captured turbulence or had the potential to show turbulence from which the energy dissipation rate in the turbulence could be calculated. In order to achieve this goal, the classification scheme would segregate images into one of three categories namely “cloudy”, “dynamic” and “calm”. Only the dynamic images carried the information sought. The variation in images that display turbulence in terms of the scale size and persistence makes the task of identifying dynamic images very challenging.

2. As explained by the authors, attempting to develop a classification method by applying an artificial intelligence approach directly to the images would be computationally very demanding. An alternative approach based on eight features (two basic; three textual, and three based on power spectral density calculations) representing each image was employed as the input to a neural network (NN). The features chosen together with the characteristics they represent are clearly described. Sequences of these image features, rather than features of individual images, are the input to the neural network because this helps the classification process – just as it does in the case of manual classification. Two sets of sequences are used – a 13 image sequence and a 61 image sequence (both chosen symmetrically around the time of the image to be classified) – to provide sensitivity to both short-term and longer term events. Thus two neural network (Temporal Convolution Network (TCN)) instances were constructed. The output of the two neural networks was combined to calculate a probability for each of the label classes (cloudy, dynamic or calm).

3. The manuscript is well structured and clearly written. It provides a great deal of detail on the classification algorithm employed and the architecture of the TCNs as well as the training procedure and evaluation of the results. It is to the credit of the authors that the high level of technical detail and the reasons for the decisions are clearly explained.

4. Cameras of the type used in this study are being employed more and more frequently in recent years for atmospheric studies. They produce enormous quantities of data, which mush be archived and then analysed thereby making the work described here very interesting and a valuable contribution to this field of study. The text includes an appropriate set of references. The work is suitable for publication in AMT, provided that the following points are addressed.

Specific Comments

1. Explain the rationale behind the decision to use 70% of the available dataset for training; 20% for validation and 10% for testing.

2. Explain the reason for 100 epochs (line 157) (see also lines 267-268). What determines the number? For example, could it be 50 or 200?

3. One can readily appreciate the difficulty of manual classification as discussed in line 409 and following. However, the number of time steps in the test dataset that were deemed to be misclassified (1110 and 1040) in Table 4 is a cause of concern. If I understand this correctly, these were manual classifications. The manuscript describes the negative impact of these manual misclassifications on the statistical measures of the neural network classifier (lines 425-427; lines 430-435 and lines 503-504). Why do the authors not repeat the calculation of the statistical measures using the "correct" classification to establish the "true" value of the statistical measures?

4. Lines 430-435 state that the NN-classifier is superior to manual classification at distinguishing between "calm" and "dynamic" episodes, which is indeed good news for the method, but leaves the reader wondering if the statistical measures have a great deal of validity.

5. A second even more disturbing issue arises with the large number of incorrect manual classifications. Since the test data is only 10% of the total (70% training; 20% validation in line 152), perhaps a large proportion of the training and validation were manually classified incorrectly to start with, thereby having a negative effect on both the training and the validation.

Minor technical corrections and suggestions

Title; Possible alternative title: "Application of machine learning for the detection of turbulence in 2D airglow imager data"

Line 26; Omit b in "2017b".

Line 29/30; replace "OH* measurements are also possible from satellite (see table 1 of Wüst et al. (2023) for limb instruments). They can be made in limb or nadir. " by "OH* measurements are also possible from satellite where they can be made in limb or nadir viewing geometry (see table 1 of Wüst et al. (2023) for limb instruments)."

Line 38; replace "Former studies " by "Previous studies".

Line 42; suggest "propagate" in place of "exist".

Line 54; consider "reliable" in place of "resilient".

Line 69; replace "disbalance" by "imbalance".

Line 71; suggest "very difficult" in place of "hardly possible".

Line 73; replace "data basis" by "database".

Line 81; suggest "The system has been described already … " in place of "The system is already described…".

Line 83; insert a space between "640" and "pixels".

Line 88; insert a space after "175" and "13.9". Replace "pixels^{-1}" by "pixel^{-1}".

Line 89; insert a space before "km".

Line 90; replace "data basis" by "database". Replace "June, 11th 2019 and February, 25th 2020." by "June 11th, 2019 and February 25th, 2020.

Line 91; replace "measurement have been performed in 258 nights" by "measurements have been performed on 258 nights".

Line 92; replace “cloud coverage and do not allow” by “cloud cover and prevent”.
Line 93; insert “of” after “all”.
Line 94; replace “one night has more than one clear episode as soon as it is interrupted by cloudy episodes.” by “a single night may have several clear episodes interspersed with cloudy episodes.”.
Line 97; recommend “the average image of each episode (a pixel-wise mean of all images in that episode)” in place of “the average image, a pixel-wise mean of all images, of each episode”
Line 107; insert “(compared with structures in “dynamic” class)” after “structures”.
Line 117; replace “distinguishing” by “to distinguish”.
Line 122; Replace “Mean and standard deviation of the label class “dynamic” are to be expected between “calm” and “dynamic”.” by “Both mean and standard deviation of the label class “dynamic” are expected to have values intermediate between those of the “calm” and “cloudy” class.”
Line 125: replace “and a way” by “and is a method”.
Line 133; insert a space before “pixels”.
Line 135; replace “power spectral density (PSD) in dependence of the horizontal wave number k is derived.” by “power spectral density (PSD) as a function of horizontal wave number k is derived.”
Line 140; consider “takes into account the fact that clouds tend to cause stronger fluctuations over time than during clear sky episodes.” in place of “takes the fact into account that, for example, clouds are causing stronger fluctuations over time than during clear sky episodes.”
Line 157; Explain the reason for 100 epochs. See also lines 267-268. What determines the number?
Line 162; consider “inadvertent” instead of “unnoticed”.
Line 168; consider “exceed” instead of “surpass”.
Line 196; omit “also”.
Line 237; omit “an” before “neurons”.
Line 291; omit the space in the word “precision”.
Line 328 and throughout the manuscript; adopt a consistent approach to the use of double quotation marks around the various classes, e.g., “clouds”, not „clouds”.
Line 332; replace “no or more than one” by no label class or more than one”.
Line 335; insert space in “Table 1,Table 2 andTable 3” as follows: “Table 1, Table 2 and Table 3”.
Line 337; this sentence appears to be incorrect based on the layout of Table 1 and Table 2 and is a source of considerable confusion for the reader until she/he realizes that it is incorrect. It should state “the manual classifications are plotted in the vertical direction and the automatic predictions in the horizontal direction”. To make both Tables completely clear, the left hand column should be labelled “Manual classification”, and the top row should be “Classifier prediction”.
Line 365; insert a space before “Table 2”.
Line 368; omit the word “by”.
Line 382; ““calm” predictions” rather than “calm predictions”; i.e., put quotation marks around “calm”.
Line 393; insert “a” before “further”.
Table 5; how do the values in column 2 (total 6666) relate to the values in Table 1?
Line 428; omit “On the one hand,”.
Line 429 replace “On the other hand” by “In addition”.
Line 456; replace “(with 13 resp. 61 time steps)” by “(with 13 and 61 steps respectively)”.
Line 483/484; the superscript “-1” in “kg^{-1}” has become separated onto two lines. Please correct this.

Line 492; replace “data basis” by “database”.
Line 493; consider “reliable” in place of “resilient”.
Line 509/510; omit “from local case studies to investigations of global extent”. Despite the reference to “global information” in line 54, this manuscript does not address this issue in a substantive way.
Line 541; “Horak” should begin on a new line.
Line 552; “Marsh” should begin on a new line.
Line 555; “Murphy” should begin on a new line.
Line 607; replace “basing” by “based”.