Software Impacts 17 (2023) 100568

Contents lists available at ScienceDirect

Software Impacts

journal homepage: www.journals.elsevier.com/software-impacts

Original software publication

SOM-QE ANALYSIS: A biologically inspired technique to detect and track meaningful changes within image regions

John Mwangi Wandeto [a,*], Birgitta Dresp-Langley [b]

[a] Department of Information Technology, Dedan Kimathi University of Technology, P.O. Box 12305-10100 Nyeri, Kenya
[b] ICube UMR 7357 CNRS - University of Strasbourg, France

ARTICLE INFO

Keywords:
Self-Organizing Map
Quantization Error
Medical images
Satellite images
Viral Proliferation Models

ABSTRACT

In practice, finding evidence for subtle changes in critical image regions through visual expert inspection of serial imaging data can be challenging. For scans taken at relatively short intervals, relevant changes may be too small to be noticed, yet diagnostically meaningful. The earliest such detection generates critical insights into potential risks, the fast it permits setting up early control mechanisms or strategies for clinical treatment. SOM-QE algorithm automatically detects subtle but significant changes in image time series providing information likely to be meaningful for experts. It is implemented in Python to analyze medical, geographic, or behavioral data.

Code metadata

Current code version	v2
Permanent link to code/repository used for this code version	https://github.com/SoftwareImpacts/SIMPAC-2023-308
Permanent link to reproducible capsule	https://codeocean.com/capsule/3541740/tree/v2
Legal code license	*MIT License*
Code versioning system used	git
Software code languages, tools and services used	python
Compilation requirements, operating environments and dependencies	Pyhton 3 with:numpy==1.24.2 matplotlib==3.7.1 MiniSom==2.3.1
If available, link to developer documentation/manual	
Support email for questions	john.wandeto@dkut.ac.ke, ndetos@gmail.com

1. Introduction

SOM-QE ANALYSIS is a software tool that permits analyzing temporal variations in image data with to-the-single-pixel precision. The output metric (the SOM-QE) creates a label that can be used to uniquely identify a specific state reflected by an image of a given time series. If the pixel content changes, SOM-QE analysis will detect this change and assign a corresponding image label. It can be used to differentiate between images in situations such as when two sets of medical images taken at different times from a patient are compared, [1]. If the images were taken from a patient during two different clinical exams, then the radiologist is able to get insight into otherwise non-detected changes in a patient's condition between two clinical sessions. The SOM-QE value represents the Quantization Error (QE) between the final weights of a Self-Organizing Map (SOM) and the values in the image it has learned. SOM is an unsupervised neural network that employs competitive learning to produce lower-dimension representation of some input data, [15], a process that closely imitates the working of the human brain.

The code (and data) in this article has been certified as Reproducible by Code Ocean: (https://codeocean.com/). More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Corresponding author.
E-mail addresses: john.wandeto@dkut.ac.ke (J.M. Wandeto), birgitta.dresp@unistra.fr (B. Dresp-Langley).

https://doi.org/10.1016/j.simpa.2023.100568
Received 28 June 2023; Received in revised form 25 July 2023; Accepted 5 August 2023

The weights of the map can be randomly initialized, and on a learning iteration the weight whose value is closest to the input vector becomes the winner. In addition, this weight's value is adjusted to be closer to that of the winning input. Weights neighboring the winner will have their weights adjusted too, but their adjustment is of lower margin than that done on the winning weight. After several iterations, the map settles, with each weight having won inputs to become a stable zone. Though weights win over input vectors, their values may not equals those of the inputs, making room for the existence of QE. The average of QE values in an image gives its SOM-QE value, used to identify the image state at that time.

2. Self-Organizing Map (SOM) and Quantization Error (QE)

The Self-Organizing Map may be described formally as a nonlinear, ordered, smooth mapping of high-dimensional input data onto the elements of a regular, low-dimensional array [15]. It is assumed that the set of input variables is definable as a real vector x, of n-dimension. A parametric real vector m_i of n-dimension is associated with each element in the SOM. Vector m_i is a model and the SOM is therefore an array of models. Assuming a general distance measure between x and m_i denoted by $d(x, m_i)$, the map of an input vector x on the SOM array is defined as the array element m_c that matches best (smallest $d(x, m_i)$) with x. During the learning process, the input vector x is compared with all the m_i in order to identify m_c, Fig. 1.

Models topographically close in the map up to a certain geometric distance, indicated by h_{ci} and shown by the bigger cycle, will activate each other to learn something from their common input x. This results in a local relaxation or smoothing effect on the models in this neighborhood, which in continuous learning leads to global ordering. SOM learning is represented by the equation

$$m(t+1) = m_i(t) + \alpha(t)\, h_{ci}(t) \left[x(t) - m_i(t) \right] \tag{1}$$

where $t = 1, 2, 3...$ is an integer, the discrete-time coordinate, $h_{ci}(t)$ is the neighborhood function, a smoothing kernel defined over the map points which converges towards zero with time, $\alpha(t)$ is the learning rate, which also converges towards zero with time and affects the amount of learning in each model. At the end of the winner-take-all learning process in the SOM, each image input vector x becomes associated to its best matching model on the map m_c. The difference between x and m_c, $||x - m_c||$, is a measure of how close the final SOM value is to the original input value and is reflected by the quantization error, QE. The average QE of all x (X)in an image is given by:

$$QE = 1/N \sum_{i=1}^{N} \left\| X_i - m_{c_i} \right\| \tag{2}$$

where N is the number of input vectors x in the image. Eq. (2) gives the SOM-QE value of an image and when calculated for a series of images it is different for each image. This is because it depends on the content of each image, thus when a series of images is taken over the same location and then subjected to the same SOM, the difference in their SOM-QE value reflects the changes that have occurred in that image location over time.

3. Impact overview

Thus, SOM-QE expresses difference between an input value and its model representation at any given moment in time, and has been used successfully to detect, in minimal computation time, meaningful changes across medical and satellite time series images. By quantifying changes translating the effect of evolution with time in the image contents, it helps track a patient's condition by showing changing states, to be interpreted as either improving, deteriorating, or stable.

Successful computation of the SOM-QE opens a window to further investigate its significance in a given context. This allows understanding what it means when an image's SOM-QE value is zero, and the potential extent of differences between images with different SOM-QE values. Visualizing the consistency of increases or decreases in differences between SOM-QE then provides further insight into the potential extent and meaning of temporal change between images.

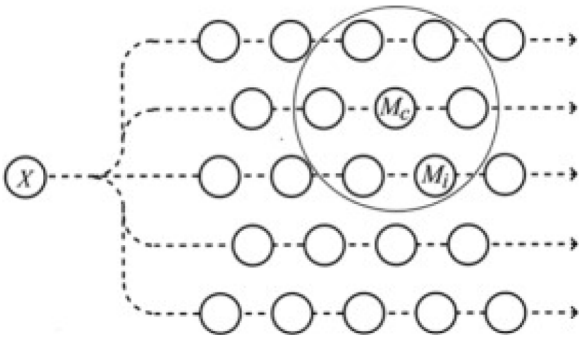

Fig. 1. Illustration of a Self-Organizing Map. An input data item X is broadcast to a set of models M_i, of which M_c matches best with X. All models that lie in the neighborhood (larger circle) of M_c in the array will be updated together in a training step and finally match better with X than with the rest, [16].

For example, determining the efficiency of novel drugs for treatment of a specific condition has long been of interest to physicians and surgeons. The earliest such can be determined, from the date of first administration of the treatment, the better. SOM-QE detects small progress/remiss of a disease condition that human observers could not detect [1,2], and thereby enables the programming of clinical treatment early rather than when it may be too late, i.e. when large differences become detectable. SOM-QE accomplishes this task fast, without requiring extensive computing resources [3–5].

Wandeto et al. [1], reported a finding whereby an expert wrongly classified all cases with 1% artificial lesion growth, and only achieved an accuracy of 20% for cases with 5% growth, while they were all reliably captured by the SOM-QE. The same expert correctly classified all cases with a minimum of 22% lesional growth. These experiments with human observers confirm that small growths in lesions are often impossible to detect for humans, while they are reliably captured by the SOM-QE. SOM-QE ANALYSIS thus provides a novel and parsimonious technique for the pre-analysis of large bodies of imaging data reflecting temporal evolution/change in critical physical states.

SOM-QE has been shown to work in various areas, such as in the classification of cell imaging data, where it may provide helpful information prior to any further comparison of representative experimental data to mathematical models of viral propagation in host cells [6,7], for example. In a computer-generated database of images, viral particle increase was simulated by a one-by-one pixel increase across images in black or gray single pixel content representing dead or partially infected cells. Hypothetical remission was simulated by a one-by-one increase in white pixels coding for living/healthy cells in the original image model. This characterizes single stages of viral proliferation in cells, or cell recovery, after focal infection of cells in vitro. Here, the SOM-QE is successfully exploited to automatically classify the model images as a function of the cellular changes represented therein.

The algorithm has also been applied to detect certain aspects of somatosensory cognition, cf. the basis of human ability to manipulate and transform objects of the physical world by grasping them with the right amount of force. This made a contribution [8,9] in the area of new technologies for monitoring grip forces during hand and finger movements in nonstandard task contexts such as robot-assisted surgery. Functionally motivated spatiotemporal analysis of individual average grip forces, computed for time windows of constant size in the output of

a restricted amount of task-relevant sensors in the dominant (preferred) hand, reveal finger-specific synergies reflecting robotic task skill. As a neural network metric, SOM-QE reliably captures the differences between novice and expert performance in terms of grip force variability. The analyses lead the way towards grip-force monitoring in real time. This will permit tracking task skill evolution in trainees, or identify individual proficiency levels in human robot-interaction, which represents unprecedented challenges for perceptual and motor adaptation in environmental contexts of high sensory uncertainty. Cross-disciplinary insights from systems neuroscience and cognitive behavioral science, and the predictive modeling of operator skills using parsimonious Artificial Intelligence (AI), will contribute towards improving the outcome of new types of surgery, in particular the single-port approaches such as NOTES (Natural Orifice Transluminal Endoscopic Surgery) and SILS (Single-Incision Laparoscopic Surgery).

SOM-QE metric is also capable of detecting and scaling symmetry uncertainty in response to visual patterns [10]. Such capacity is tightly linked to the metric's proven selectivity to local contrast and color variations in large and highly complex image data.

In addition, SOM-QE has been shown [3,11], to be capable of reliably discriminating between fine differences in local contrast intensity and sign. While this capability of the QE is akin to functional characteristics of a specific class of retinal ganglion cells (the Y-cells) in the visual systems of the primate and the cat, the sensitivity of the QE surpasses the capacity limits of human visual detection. In this case, SOM-QE is found to reliably signal changes in contrast or color when contrast information is removed from or added to the image, but not when the amount and relative weight of contrast information is constant and only the local spatial position of contrast elements in the pattern changes. While the RGB Mean reflects coarser changes in color or contrast well enough, the SOM-QE is shown to outperform the RGB Mean in the detection of single-pixel changes in images with up to five million pixels. This could have important implications in the context of unsupervised image learning and computational building block approaches to large sets of image data (**big data**), including deep learning blocks, and automatic detection of contrast change at the nanoscale in Transmission or Scanning Electron Micrographs (TEM, SEM), or at the sub-pixel level in multi-spectral and hyper-spectral imaging data.

SOM-QE can be applied to detect structural change in geographic regions too, [12,13]. For instance, it was used to process satellite images of Lake Mead - a water reservoir formed by the Hoover Dam on the Colorado River in the Southwestern United States - and determine the yearly raising/falling of water levels in the lake, which is significant in determining effects of drought, climate change, etc. In time series of satellite images taken in the years 1984 to 2011, SOM-QE reliably and consistently captures the changing water levels of Lake Mead and proves significantly correlated with related measures from the Hoover Dam Control Room during that same time period, and other public data made available by organizations like the National Bureau of Reclamation, or NASA's Earth Observatory.

SOM-QE code is available at [14], where one can edit, change parameters, and re-run it.

4. Future improvements

The performance of SOM-QE can be improved by creating a module for standardizing input. Diagnostic medical imaging or satellite equipment often misses slow evolution because critical change may be obscured by variations in pixels position or intensity between scans. The images may then not have accurate information on a specific state change. This brings in the need for establishing suitable criteria for pre-processing input images.

CRediT authorship contribution statement

John Mwangi Wandeto: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Writing – original draft, Writing – review & editing. **Birgitta Dresp-Langley:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Scholarly publications enabled by SOM-QE

[1] J.M. Wandeto, H. Nyongesa, Y. Rémond, B. Dresp-Langley, Detection of small changes in medical and random-dot images comparing self-organizing map performance to human detection, Inform. Med. Unlocked 7 (2017) 39–45, http://dx.doi.org/10.1016/j.imu.2017.03.001.

[2] J. Wandeto, H. Nyongesa, B. Dresp, Detection of smallest changes in complex images comparing self-organizing map to expert performance, 2017, Perception ECVP 2017 Abstract Supplement.

[3] J.M. Wandeto, B. Dresp-Langley, The quantization error in a self-organizing map as a contrast and colour specific indicator of single-pixel change in large random patterns, Neural Netw. 119 (2019a) 273–285, http://dx.doi.org/10.1016/j.neunet.2019.08.014.

[4] B. Dresp-Langley, J.M. Wandeto, Fast artificial intelligence for the detection of viral particle proliferation in cellular imaging data, in: European Society for Medicine General Assembly 2022, August 4-6, online.

[5] J.M. Wandeto, B. Dresp, Ultrafast automatic classification of SEM image sets showing CD4 cells with varying extent of HIV virion infection, 2019, http://dx.doi.org/10.13140/RG.2.2.25942.29760.

[6] B. Dresp-Langley, J.M. Wandeto, Pixel precise unsupervised detection of viral particle proliferation in cellular imaging data, Inform. Med. Unlocked 20 (2020) 100433, http://dx.doi.org/10.1016/j.imu.2020.100433.

[7] B. Dresp-Langley, J.M. Wandeto, Unsupervised classification of cell-imaging data using the quantization error in a self-organizing map, in: H.R. Arabnia, K. Ferens, D. de la Fuente, E.B. Kozerenko, J.A. Olivas Varela, F.G. Tinetti (Eds.), Advances in Artificial Intelligence and Applied Cognitive Computing, Transactions on Computational Science and Computational Intelligence, Springer International Publishing, Cham, 2021, pp. 201–209, http://dx.doi.org/10.1007/978-3-030-70296-0_16.

[8] R. Liu, J. Wandeto, F. Nageotte, P. Zanne, M. de Mathelin, B. Dresp-Langley, Spatiotemporal modeling of grip forces captures proficiency in manual robot control, Bioengineering 10 (59) (2023) http://dx.doi.org/10.3390/bioengineering10010059.

[9] B. Dresp-Langley, R. Liu, J.M. Wandeto, Surgical task expertise detected by a self-organizing neural network map, 2021, http://dx.doi.org/10.48550/arXiv.2106.08995.

[10] B. Dresp-Langley, J.M. Wandeto, Human symmetry uncertainty detected by a self-organizing neural network map, Symmetry 13 (2021) 299, http://dx.doi.org/10.3390/sym13020299.

[11] J.M. Wandeto, B. Dresp-Langley, The quantization error in a self-organizing map as a contrast and colour specific indicator of single-pixel change in large random patterns, Neural Netw. 120 (2019) 116–128, http://dx.doi.org/10.1016/j.neunet.2019.09.017, Special Issue in Honor of the 80th Birthday of Stephen Grossberg.

[12] B. Dresp, H. Nyongesa, J. Wandeto, Vision-inspired automatic detection of water level changes in satellite images: the example of Lake Mead, Perception 48 (1) (2019).

[13] J.M. Wandeto, H.O. Nyongesa, B. Dresp-Langley, Detection of structural change in geographic regions of interest by self organized mapping: Las vegas city and Lake Mead across the years, 2018, http://dx.doi.org/10.48550/arXiv.1803.11125.

References

[14] John Mwangi Wandeto, Birgitta Dresp-Langley, SOM-QE: A biologically inspired technique to detect and track meaningful changes within image regions [Source Code], 2023, http://dx.doi.org/10.24433/CO.3246066.v2.

[15] T. Kohonen, Automatic formation of topological maps of patterns in a self-organizing system, in: Proceedings of the 2nd Scandinavian Conference on Image Analysis, 1981.

[16] T. Kohonen, MATLAB Implementation and Applications of the Self-Organizing Map, Unigrafia Oy, Helsinki, Finland, 2014, http://docs.unigrafia.fi/publications/kohonen_teuvo/index.html.