DATASET FOR PUBLICATION: USEFULNESS OF SYNTHETIC DATASETS FOR DIATOM AUTOMATIC DETECTION USING A DEEP-LEARNING APPROACH

Version 1.0

Laviale, Martin; Venkataramanan, Aishwarya, 2023, "Dataset for publication: Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach", https://doi.org/10.12763/UADENQ, Université de Lorraine, V1, UNF:6:WAcYIGvggM/oSHwrm9hPqQ== [fileUNF]

Description

This repository contains the dataset and code used to generate synthetic dataset as explained in the paper "Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach". Dataset : The dataset consists of two components: individual diatom images extracted from publicly available diatom atlases [1,2,3] and individual debris images. - Individual diatom images : currently, the repository consists of 166 diatom species, totalling 9230 images. These images were automatically extracted from atlases using PDF scraping, cleaned and verified by diatom taxonomists. The subfolders within each diatom specie indicates the origin of the images: RA[1], IDF[2], BRG[3]. Additional diatom species and images will be regularly updated in the repository. - Individual debris images : the debris images were extracted from real microscopy images. The repository contains 600 debris objects. Code : Contains the code used to generate synthetic microscopy images. For details on how to use the code, kindly refer to the README file available in `synthetic_data_generator/`.

Subject

Computer and Information Science; Earth and Environmental Sciences

Keyword

Synthetic dataset, Images, Diatoms, Automatic detection, Deep learning

Related Publication

Aishwarya Venkataramanan, Pierre Faure-Giovagnoli, Cyril Regan, David Heudre, Cécile Figus, et al.. Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach. Engineering Applications of Artificial Intelligence, Elsevier. doi: 10.1016/j.engappai.2022.105594

???file.dataFilesTab.terms.list.license???

Files | Metadata | Terms | Versions

Export Metadata

???dataset.metadatalanguage.view.guidance???

Citation Metadata

Dataset Persistent ID

doi:10.12763/UADENQ

Publication Date

2023-07-21

Title

Dataset for publication: Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach

Author

Laviale, Martin (LIEC ; Université de Lorraine, CNRS ; France) - ORCID: 0000-0002-9719-7158
Venkataramanan, Aishwarya (LIEC ; Université de Lorraine, CNRS ; France) - ORCID: 0000-0002-6100-0034

Contact

Use email button above to contact.

Laviale, Martin (LIEC ; Université de Lorraine, CNRS ; France)

Description

This repository contains the dataset and code used to generate synthetic dataset as explained in the paper "Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach". Dataset : The dataset consists of two components: individual diatom images extracted from publicly available diatom atlases [1,2,3] and individual debris images. - Individual diatom images : currently, the repository consists of 166 diatom species, totalling 9230 images. These images were automatically extracted from atlases using PDF scraping, cleaned and verified by diatom taxonomists. The subfolders within each diatom specie indicates the origin of the images: RA[1], IDF[2], BRG[3]. Additional diatom species and images will be regularly updated in the repository. - Individual debris images : the debris images were extracted from real microscopy images. The repository contains 600 debris objects. Code : Contains the code used to generate synthetic microscopy images. For details on how to use the code, kindly refer to the README file available in `synthetic_data_generator/`.

Subject

Computer and Information Science; Earth and Environmental Sciences

Keyword

Synthetic dataset
Images
Diatoms
Automatic detection
Deep learning

Related Publication

Aishwarya Venkataramanan, Pierre Faure-Giovagnoli, Cyril Regan, David Heudre, Cécile Figus, et al.. Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach. Engineering Applications of Artificial Intelligence, Elsevier. doi: 10.1016/j.engappai.2022.105594 https://doi.org/10.1016/j.engappai.2022.105594

Depositor

Laviale, Martin

Deposit Date

2022-11-18

Kind of Data

Image

Data Sources

Bey, M.Y., Ector, L., 2013. Atlas des diatomées des cours d'eau de la région rhône-alpes. tome 1. Centriques, Monoraphidées. tome 2. Araphidées, Brachyraphidées. tome 3. Naviculacées: Naviculoidées. tome 4. Naviculacées: Naviculoidées. tome 5. Naviculacées: Cymbelloidées, Gomphonematoidées. tome 6. Bacillariacées, Rhopalodiacées, Surirellacées, Direction Régionale de l'Environnement, de l'Aménagement et du Logement Rhône-Alpes, ISBN:978-2-11-129817-0, p.1182.; Lalanne-Cassou, C., Voisin, J.F., 2013. Atlas des diatomées d'ile de france, Direction Régionale et Interdépartmentale de l'Environnement et de l'Energie d'Île-de-France, p. 734.; Peeters, V., Ector, L., 2018. Atlas des diatomées des cours d'eau du territoire bourguignon. volume 2: Monoraphidées, Brachyraphidées, Direction Régionale de l'Environnement, de l'Aménagement et du Logement, Bourgogne-Franche-Comté. Dijon, p. 271.