EOPEN

opEn interOperable Platform for unified access and analysis of Earth observatioN data

H2020-776019

D4.1

Change detection techniques in Earth Observation

Dissemination level:	Public
Contractual date of delivery:	Month 16, 28/02/2019
Actual date of delivery:	Month 32, 30/06/2020
Work package:	WP4 Knowledge discovery and content extraction
Task:	T4.1 Change Detection in EO data T4.2 Concept and Event Detection in non EO data T4.3 Similarity fusion from multiple sources for information retrieval T4.4 Data clustering and exchange among federated databases

	T4.5 Community detection in Social Media
Type:	Report
Approval Status:	Approved
Version:	1.0
Number of pages:	72
Filename:	D4.1-Change detection techniques in Earth Observation_2020-06-30_v1.0.docx

Abstract

This deliverable reports on methods used for extracting content from EO and non-EO data. Specifically, it reports in depth the methodology and the research outputs in change detection techniques that are employed in EOPEN. The techniques are adapted to and examined in each use case scenario of EOPEN, aiming to present a complete analysis of the most suitable change detection techniques in the monitoring of floods, climate changes and food security, from Earth Observation imagery. Moreover, the deliverable contains a brief introduction to the EOPEN concept and event detection mechanisms, as well data clustering, similarity search and community detection analytics.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 776019

History

Version	Date	Reason	Revised by	Approved By
0.1	20/12/2018	Initial Draft - ToC	Anastasia Moumtzidou	Ioannis Kompatsiaris
0.2	20/01/2019	Contributions	Stelios Andreadis	Ilias Gialampoukidis
0.3	05/02/2019	Contributions	Dennis Hoppe	
0.3	05/02/2019	Contributions	Anastasia Moumtzidou	
0.4	15/02/2019	Contributions	Ilias Gialampoukidis	Stefanos Vrochidis
0.4	22/02/2019	Contributions	Vasileios Sitokonstantinou	
0.5	20/02/2019	Final Contributions	Anastasia Moumtzidou, Dennis Hoppe, Vasileios Sitokonstantinou, Stelios Andreadis	
0.6	22/02/2019	Final Contributions	Ilias Gialampoukidis	
0.7	25/02/2019	Internal Review	Hakim Boulahya, Bernard Valentin	
0.8	01/03/2019	Updated document after review	Anastasia Moumtzidou, Stelios Andreadis	Stefanos Vrochidis
1.0	30/06/2020	Updated document after addressing reviewer comments	Anastasia Moumtzidou, Stelios Andreadis	Iliias Gialampoukidis

Author list

Organization	Name	Contact Information
CERTH	Anastasia Moumtzidou	moumtzid@iti.gr
CERTH	Stelios Andreadis	andreadisst@iti.gr
CERTH	Ilias Gialampoukidis	heliasgj@iti.gr
CERTH	Stefanos Vrochidis	stefanos@iti.gr
NOA	Vasileios Sitokonstantinou	vsito@noa.gr
USTUTT	Dennis Hoppe	hoppe@hlrs.de
USTUTT	Michael Gienger	gienger@hlrs.de
USTUTT	Li Zhong	zhong@hlrs.de

Executive Summary

This deliverable presents the first version of the change detection techniques in EOPEN and it also provides a short overview of the State-of-the-Art techniques in clustering, concept extraction, event detection and community detection.

As far as the change detection is concerned, it identifies the changes realized on the landscape by considering EO data as input. Change detection addresses all three EOPEN use cases; flood risk assessment, food security and climate change. However, the climate change use case contains several aspects, such as weather conditions, but we put our focus on the road maintenance since the other aspects are covered in other EOPEN deliverables. Specifically, for the flood risk assessment and climate change use cases, we discuss the baseline algorithms developed for identifying floods in general and floods related to road passability. In case of floods both optical and radar satellite data are handled in two different approaches. For radar satellite data an automatic thresholding approach is tested, while for the optical data a Modified Normalized Difference Water Index (MNDWI) and a discriminant analysis with post-processing is applied. All approaches generate water-bodies masks in order to tackle the water-mapping task. In case of the road passability scenario, which is one of the aspects of the climate change use case, a fine-tuned deep neural network was trained that recognized roads with water. Finally, with regard to the Food Security use case, an EO-based change detection framework was designed and implemented to update land cover maps of past years, so as to be used as training datasets for supervised classifiers.

The next topic that was covered was concept detection that is responsible for adding visual tags on social images by using a fine-tuned deep neural network. Furthermore, we discuss event detection task which aims at discovering incidents that can be interesting or alerting regarding to the explored use cases and producing respective notifications in the EOPEN system. A preliminary investigation is performed on social media data about floods in Italy for the period of one year, taking into consideration the number of collected tweets per day. The detected events are manually associated with real incidents in order to show that the fluctuation of posts correlates with events.

In the sequel, we discuss similarity fusion in non-EO data. Specifically, the State-of-the-Art techniques are presented and then we present the modalities describing tweets and how single modality similarity is realized.

Another clustering procedure is performed on textual information, originating from the text of Twitter posts. A dedicated service estimates the number of clusters, assigns the available posts to the clusters and finds the most frequent terms for each group. Aside from the well-known text clustering, we also perform unsupervised image clustering. The techniques are applied to EO data in order to group images that have similar features (e.g. colour). Specifically, artificial neural networks such as an autoencoder are utilized to efficiently perform clustering on a HPDA infrastructure using Keras with TensorFlow. Results are then mapped onto a two-dimensional plane, where the distance between images represents their similarity. Each identified group is then described through a representative thumbnail.

Community detection task aims at targeting the discovery of user communities based on their interaction. An implemented service consumes the relationships between user

accounts of social media and produces a network of users, along with the detected communities and the most influential users.

Finally, in order to respond to the reviewer's comments, information is added to several parts of the deliverable, while two annexes are included as well. Specifically, the comment on the cooperation with CANDELA project is addressed in the end of Section 2.3 and a related table is added in an appendix that presents the outcome of the discussion with the respective partners, regarding the differences and similarities in the approaches used for Change Detection. As far as it concerns the comment of using OpenStreetMap for road detection, a paragraph is added inside Section 2.6.2. Regarding the use of VHR satellite imagery, a paragraph is added in the introductory paragraph of Section 2.2 . Moreover, the comment related to the use of other sources for monitoring flood events is answered by adding a paragraph in Section 2.4.1, while the use of the ECMWF Global Flood Awareness System is now discussed in a paragraph in Section 2.4.1, that explains a scenario where this system would be of interest in EOPEN. The comment on the relation between the H2020 projects beAWARE and EOPEN is addressed in a paragraph added in Section 3.2.1 and in an appendix. Moreover, the comment related to the use of data found in the Disaster Charter is addressed in Section 2.4.4 Finally, the issue regarding the version of the deliverable is resolved by updating the version number.

Abbreviations and Acronyms

AI	Artificial Intelligence
B	Blue
BoW	Bag-of-Word
DCNN	Deep Convolutional Neural Network
ECMWF	European Centre for Medium-Range Weather Forecasts
EMS	Emergency Management System
EMSR	Emergency Management System Response
EO	Earth Observation
FCN	Fully Convolutional Network
FT	Fine-tuning
G	Green
GRD	Ground Range Detected
HPC	High Performance Computing
IoU	Intersection over Union
IW	Interferometric Wide
KLD	Kullback–Leibler Divergence
LDA	Latent Dirichlet Allocation
LR	Logistic Regression
MNDWI	Modified Normalized Difference Water Index
MSE	Mean Squared Error
NCPD	National Civil Protection Department
NIR	Near Infrared
PCA	Principal Component Analysis
R	Red
R-CNN	Region – Convolutional Neural Network
RMSE	Root Mean Squared Error
SAR	Synthetic Aperture Radar
SGD	Stochastic Gradient Descent
SIFT	Scale-Invariant Feature Transform
SIN	Semantic Indexing
SNAP	Sentinel Application Platform
SOM	Self-Organizing Map
SSE	Sum Squared Error
SURF	Speeded Up Robust Features
SVM	Support Vector Machines
SWIR	Shortwave Infrared
t-SNE	t-distributed Stochastic Neighbour Embedding
VHR	Very High Resolution
VLAD	Vector of Locally Aggregated Descriptors

Table of Contents

1 INTRODUCTION

EOPEN extracts knowledge from the collected EO and non-EO data in order to add value in products related to flood monitoring, food security and climate change monitoring for extreme weather conditions. This document presents the developed change detection algorithms and baseline methods for the extraction of water masks in urban areas (PUC1), in agricultural areas (PUC2), as well as to automatically infer whether a road is passable or not due to an extreme weather event, such as floods.

Techniques for change detection are differentiated for each pilot as there are different algorithms and state of the art processes that could be applied. The developed mechanisms include an AI component, which is based on state-of-the-art Deep Learning architectures, allowing for effective decision making, as it is requested by the user requirements.

Apart from the change detection algorithms from EO data, state-of-the-art algorithms for concept and event detection, data clustering and community detection from social media have been developed to support the value-adding chain of EO data analysis. This document briefly introduces baseline approaches which are adapted to the purposes of EOPEN for integration to the 1st prototype of the EOPEN system.

EOPEN extracts high-level concepts from visual content, as provided by the Twitter crawling mechanism of EOPEN. To that end, we present in Section 3 the developed and fine-tuned Deep Convolutional Neural Network (DCNN) descriptors, which are trained in the context of the EOPEN use cases. Event detection is based on burst analysis on the contextual data (e.g. frequency of relevant tweets) which are fused with spatiotemporal information and weather information.

In Section 4, we discuss the techniques for similarity fusion and present a single modality similarity framework. The extracted concepts serve as queries in a search engine, offering the user to retrieve information on demand, which is similar to a given query. During the first year of EOPEN, the information retrieval mechanism of EOPEN platform was able to search by concept.

Grouping the collected EO and non-EO data by content is an additional functionality which is developed to cluster items that contain similar patterns, whether it concerns visual or textual information. Algorithms are parallelised and executed on HPC infrastructure in order to offer a very scalable solution, when needed.

Last but not least, EOPEN detects end user communities in the collected Twitter content. Users who post very often reliable information become influencers and their posts during a crisis event, such as a flooding, are more than useful to the public authorities.

2 CHANGE DETECTION MODULE

The change detection aims at monitoring specific changes in pre-defined Areas of Interest (AoI), including changes in the landscape (flood monitoring, food security) or in the road network. The techniques applied for change detection differ among the different use cases as there are different algorithms and state of the art processes that could be applied. In this deliverable, we present the baseline methods for flood detection and road passability that are the base for change detection for PUC 1 (Flood risk assessment and Prevention) and PUC3 (Climate Change through Earth Observation Flood) respectively. It should be noted that PUC3 covers a wide range of phenomena to be monitored, with road passability being one of them. Finally, as far as PUC2 is concerned, the current deliverable discusses the change detection framework for food security in the context of identifying extreme differences (outliers) between Sentinel-1 backscatter coefficients of different years and thereby appending accordingly the land cover maps to be used for rice paddy supervised classification.

2.1 Architecture of the change detection module

In this section, we discuss the architecture of the change detection module and how it is interconnected with the other EOPEN modules. As depicted in Figure 1, WP4 contains the modules that process the EO and non-EO data retrieved from the different modules of WP3. Change detection considers exclusively EO data and metadata and uses them for detecting changes in the water level as a consequence of floods.

Thus, the Change detection module has as input the EO data and their metadata. It applies change detection techniques and its output is passed to the Clustering EO Imagery module (T4.4) and populates the EOPEN ontology.

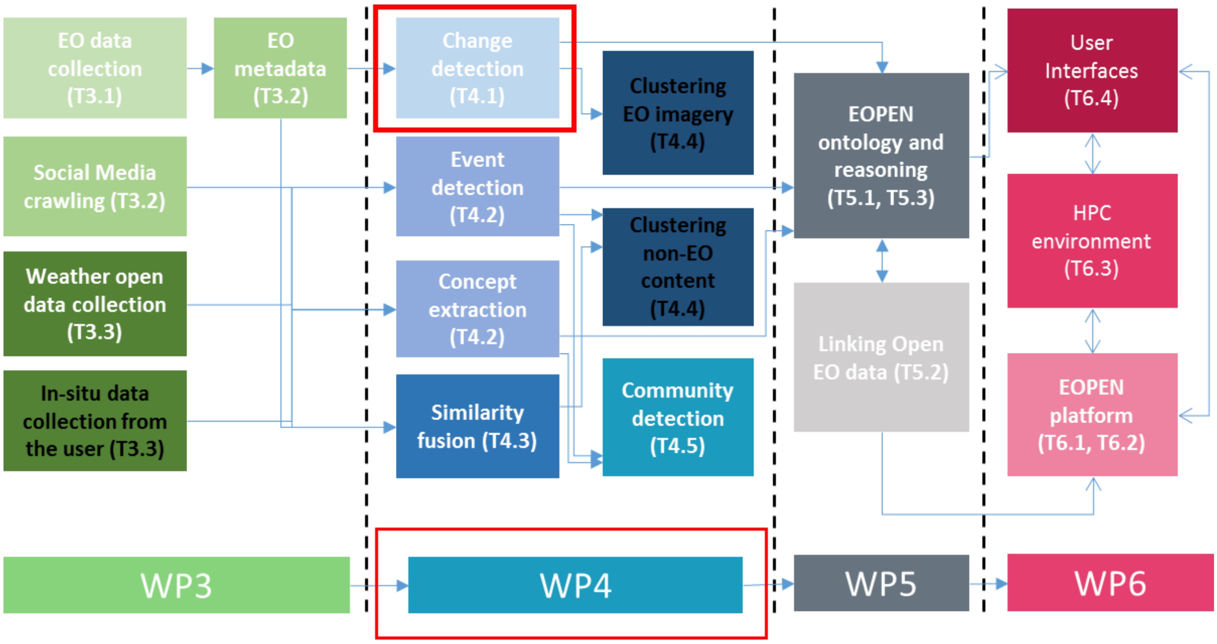

Figure 1: Position and information flow of Change detection module in EOPEN framework.

2.2 Data and use case scenarios

This section contains the data that have been used in change detection and the use case scenarios. It should be noted that our aim is to consider all the available data types for all the use cases. We should note that the data considered are provided by the Copernicus Programme.

Although the provided data do not have the high resolution that data by other satellite providers (such as Airbus DS or DigitalGlobe) have, they have the advantage that they are freely available to all the scientific community compared to the VHR (Very High Resolution) satellite imagery which comes with a very high price.

2.2.1 Data types

Sentinel 1

Sentinel-1 constitutes the first of the Copernicus Programme satellite constellation. Each of its satellites carries an advanced synthetic aperture radar that works in several specialized modes to provide detailed imagery. These data will be used for applications such as monitoring the oceans, sea ice and oil spills. It also provides data to map changing land cover, ground deformation and can be used to help emergency response when disasters such as floods strike and to support humanitarian relief efforts at times of crisis.

The first satellite, Sentinel-1A, launched on 3 April 2014, and Sentinel-1B was launched on 25 April 2016. Two more satellites, Sentinel 1C and 1D, are expected to be launched from 2021 onwards.

Sentinel-1 operates in four exclusive acquisition modes: Stripmap (SM), Interferometric Wide swath (IW), Extra-Wide swath (EW), Wave (WV). Figure 2 depicts the basic characteristics of each acquisition mode.

The primary conflict-free modes are IW, with VV+VH polarisation over land, and WV, with VV polarisation, over open ocean. EW mode is primarily used for wide area coastal monitoring including ship traffic, oil spill and sea-ice monitoring. SM mode is only used for small islands and on request for extraordinary events such as emergency management.

As input for the framework of the change detection approach (see 2.4.2), Sentinel-1 products are required. And more specifically the Ground Range Detected (GRD) products, with data acquired in Interferometric Wide (IW) acquisition mode at High Resolution (HR).

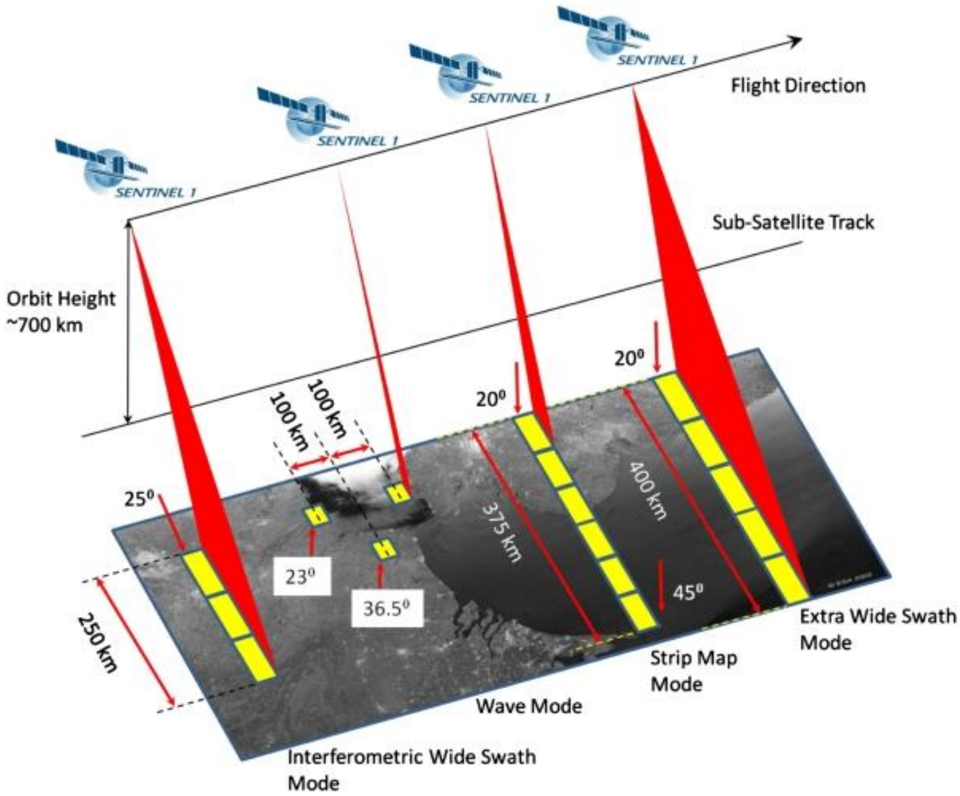

Figure 2: Sentinel-1 Product Modes. Image taken from web site: sentinel.esa.int

Sentinel 2

Where clouds, trees and floating vegetation do not obscure the water surface, high-resolution visible/infrared sensors provide good delineation of inundated areas. Sentinel-2 Earth observation mission data, as part of the EU Copernicus Programme, provides Multi-spectral data with 13 bands in the visible, near infrared, and short wave infrared part of the spectrum. It includes four bands at 10 metres, six bands at 20 metres and three bands at 60 metres spatial resolution, as depicted in Figure 3.

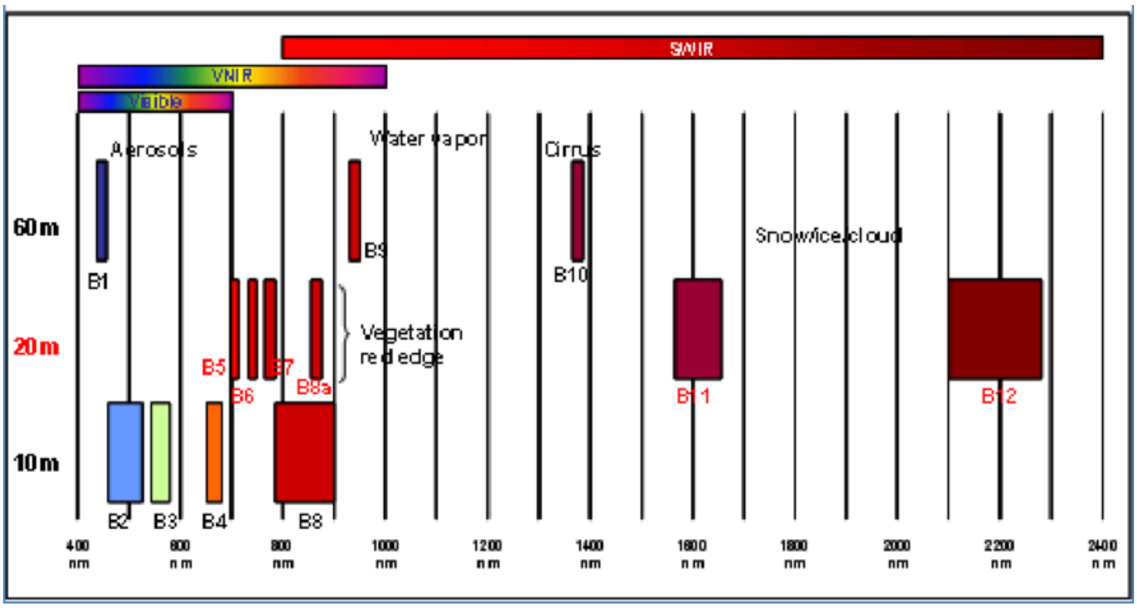

Figure 3: Sentinel-2 spectral bands (Martimort et al., 2007)

Sentinel-2 mission was also selected for its open access through publicly available web APIs and frequent revisit times, currently generating one image per 5-6 days, using two identical satellites (Sentinel 2A and 2B). The first satellite, Sentinel-2A, launched on 23 June 2015, and Sentinel-2B was launched on 7 March 2017. Two more satellites, Sentinel 2C and 2D, are expected to be launched from 2021 onwards.

2.3 Use case scenarios

In this section, we describe briefly the three use cases targeted by EOPEN. A detailed description of the use cases can be found in EOPEN Deliverable 2.1 “"Use Case Design Report".

PUC1 - Flood Risk Assessment and Prevention

This use case deals with flood monitoring with regard to flood risk assessment and prevention. The area that is studied is within the Italian Eastern Alps river District and it is regularly affected by critical flooding from the Bacchiglione River and its tributaries. Due to unfinished planned flood defences, the risk of flooding remains. In case the flood affects the city area, high levels of water are observed in the streets that cause severe problems including drowning of people, building damage and traffic problems. Figure 4 depicts an example of water detection using Sentinel-2 product in the area of Vicenza of Italy.

Figure 4: Example of water detection using Sentinel-2 product.

PUC2 - Food Security through Earth Observation dataset

This use case comprises several different components including food access, distribution, food supply stability and use of food. It discusses the fact that the increase of the earth’s population, will increase the food demand. Hence, in order to tackle of food security, it is necessary to improve nutrition to people. The food crises mainly arise from weather extremes, natural disasters, societal crises, and other reasons such as population growth and cultivation restrictions that require fast and efficient communication for effective and timely decision making. Within the scope of EOPEN’s food security pilot, the focus is on South Korea which has experienced rapid population growth due to urbanization, commercialization of the food chain. In case of an international grain supply failure, the country is exposed to a food security crisis due to its high dependence on major grains and limited exporting countries. Therefore the specific use case will monitor the rice production in South Korea (rice mapping, rice yield estimation) in order to ultimately provide accurate and timely food security related information to the pertinent decision makers. Figure 5 depicts an example of a water mask from the Korea region that delineates both permanent water and rice areas.

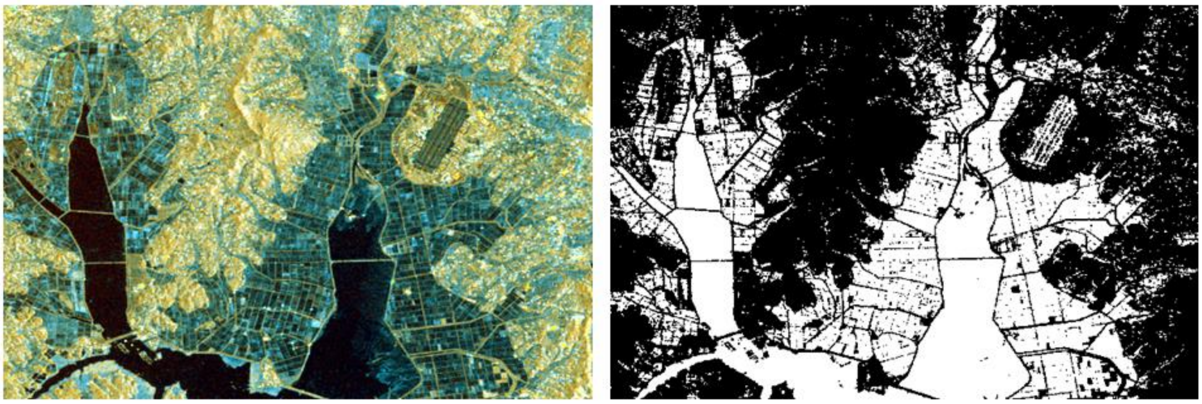

Figure 5: On the left the false-colour visualisation image, combining the processed VV and VH bands. On the right the generated water mask via automatic thresholding of the processed VH band (delineates both permanent water and rice areas).

PUC3 - Monitoring the Climate Change through Earth Observation Flood

This use case deals with the climate change and specifically with the rise of the average temperature and the increase of rate of the increase in precipitation in Finland. Such changes affect mostly winters than summers. Thus, Finland has adapted several mitigation and adaptation plans. Within the same content Finland plans to reduce its greenhouse gas emissions and promote sustainable industry, land use planning and construction. Land-use decisions are critical to mitigate snow recreation industries and the changing pastures of reindeer and their Saami herders. This use case considers historical snow and temperature data, supplemented by EO data, in an effort to support Finnish Transportation Agency's current and future road maintenance for the Finnish drivers and riders as the country shifts into "intelligent traffic", "mobility as a service", and self-driving cars to improve safety, streamline traffic and reduce energy needs. It also involves temperature and snow data support for the Finnish Lapland communities who are experiencing the greatest climate change consequences. In the current deliverable, we focus on the road maintenance issue and we identify whether a road is passable or not due to water. Figure 6 depicts an example of road passability using Sentinel-2 product in the area of Vicenza of Italy.

At this point we should note that Change Detection is inside the scope of another H2020 Project too, i.e. the CANDELA Project[1]. In A.1 Appendix, there is a detailed table, which summarizes the aspects of the Change Detection modules in the EOPEN and CANDELA projects and thus makes clear the similarities and differences of the approaches proposed. In CANDELA project, Sentinel-2 images are exclusively used and the approach proposed is unsupervised and based on a 25-dimensional feature vector, whereas EOPEN considers both

[1] http://candela-h2020.eu/

Sentinel-1 and Sentinel-2 products, uses several features depending on the product type and the approach can be either supervised or unsupervised, according to the problem targeted.

Figure 6: Example of road passability detection using Sentinel-2 product.

2.4 Change detection in PUC1 - "Flood Risk Assessment and Prevention

This section focuses on describing flood monitoring approaches that will be applied on detecting changes of the water level that is attributed to floods or draughts.

2.4.1 State of the art in flood monitoring using EO data

Flooding, as the world's most costly type of natural disaster is responsible for destroying housing, agriculture and communications, costing dozens of lives. Remote sensing systems can provide the required information for delineating the flood-affected areas, assessing the damage or detecting urban areas prone to be flooded.

For water mapping a common approach is developing a threshold methodology that separates water from non-water areas. To avoid a false local threshold, an empirical threshold of -18db can be applied (Twele et al., 2016). The modified difference water index (MNDWI) is more suitable than NDWI for enhancement of water with many built-up land areas in the background than the NDWI because it can efficiently reduce and even remove built-up land noise (Hanqiu Xu, 2006). Improvement in spatial resolution can be achieved by using Sentinel-2 images (Green and SWIR bands) performing pan-sharpening techniques to match the different resolutions of the bands, enhancing image quality (Du et al., 2016).

The current trend in flood detection relies on Neural Networks (Kang et al., 2018), where the Fully-Convolutional Network (FCN), a variant of VGG16 on Gaofen-3 SAR images, is utilized for flood mapping. FCN demonstrates robustness to speckle noise in SAR images. To make the deep learning model more universal, speckle noise is not filtered. In (Kia et al., 2012) the most widely used criteria performances, namely coefficient of determination (R2), sum squared error (SSE), mean squared error (MSE), and root mean squared error (RMSE) are used to optimize the performance of the Artificial Neural Network (ANN). Each method is estimated from the ANN predicted values and the measured discharges (targets). Seven input nodes, each representing flood causative parameters, including rainfall, slope, elevation, soil, geology, flow accumulation, and land use are used during the ANN modeling. Rainfall factor is the main factor in the training of the neural network with the elevation the most important factor for flood susceptibility mapping. The approach in (Skakun, 2012) is based on the segmentation of a single SAR image using self-organizing Kohonen maps

(SOMs) and further image classification using auxiliary information on water bodies that could be derived, from optical satellite images. A moving window is applied to process the image and spatial connection between the image pixels is taken into account. Neural networks weights are adjusted automatically using ground-truth training data.

In contrast, we are investigating a unifying approach to infer whether an area is inundated or not, due to a severe flood event. We report and examine state-of-the-art classification methods with transfer learning based on Deep Convolutional Neural Networks, aiming to develop an effective flood monitoring algorithm.

Finally, it should be noted that, apart from satellite data, other sources could also be considered for addressing the flood detection problem. There exist several options including the use of sensors and flood forecasting models. As far as the sensors are concerned, it is common to consider in situ sensors or river heights within the context of getting an early indication of a flood event. AAWA EWS harvests on a daily basis all the datasets from the regional monitoring networks including the information coming from sensors, and feeds them to their hydrological model. Therefore, inside the AAWA database one can find archive data from 1924 till today and hourly observations from the automatic in situ network from 1994 till now. However, the use of sensors is outside the scope of EOPEN.

Another source could be forecast flood models, such as the ECMWF Global Flood Awareness System[2]. These models cannot be used for monitoring, but solely for forecasting purposes; thus they could be considered as an early warning and their output could trigger the flood monitoring modules.

2.4.2 **Framework**

In this deliverable, we present two frameworks developed for identifying flooded areas using satellite images. Both frameworks are considered as baseline methods. The first framework is a thresholding algorithm that can be applied with slight modification both to Sentinel 1 and Sentinel 2 data, while the second framework is exclusively oriented to handling optical images and thus it can only be applied to Sentinel 2 products.

Sentinel 1

During heavy raining, dense clouds cover the sky making it virtually impossible to map the area below using satellites that are equipped with passive acquisition instruments. Thus, a satellite with the cloud penetrating ability of Synthetic Aperture Radar (SAR) instrument appears as the most suitable candidate for monitoring the floods at real time. Sentinel-1 carries an active system that sends a microwave signal from a sensor platform to the ground and detects backscattered waves that the ground reflects directly back to a receiver on the same platform, allowing them to operate under any condition at day or night times. The received backscatter is depicted in SAR imagery that requires pre-processing to be usable, due to the intense noise. The performed operations are these of calibration, speckle noise removal, terrain correction and conversion to the logarithmic scale (dB). Then a thresholding technique that separates the water from non-water areas is used. Attempting to estimate the threshold on the whole image may fail due to the massive uneven distribution of water

[2] https://www.ecmwf.int/en/about/media-centre/news/2018/major-upgrade-global-flood-awareness-system

and non-water amounts. To solve this, the area of interest is split into multiple smaller tiles where individual thresholds of the VV polarization band are easier to be calculated, eventually forming an average total threshold that is applied to the whole processed image, generating the water-bodies mask. Framework is depicted in Figure 7. The thresholding approach is built with snappy (ESA's python API that provides SNAP (Sentinel Application Platform) (functionality) and runs as a service allowing automatic calculation of water bodies of any Sentinel-1 GRD-IW product that is fed to the model. A false color image combining VV and VH polarization bands is generated for better visualization of the AOI.

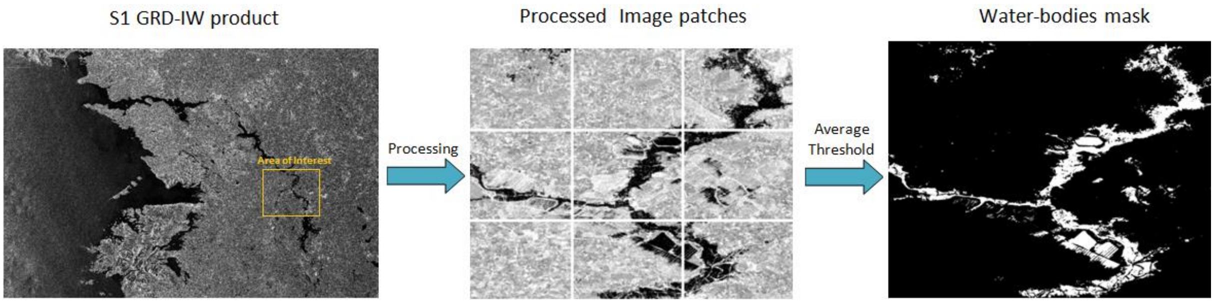

Figure 7: Framework of automatic thresholding water-bodies mask generation of a Sentinel 1 image, Ballinasloe (Ireland).

Sentinel 2

Baseline1

Where clouds and dense vegetation do not obscure the water surface, high-resolution passive sensors provide good delineation of flooded areas. Near-infrared imagers are especially effective, because the near-infrared spectral bands are strongly absorbed by water, but reflected by land.

Figure 8 illustrates the workflow of the generation of the water-bodies mask, using optical satellite data. MNDWI (see 2.4.1) is used combined with a hard threshold to separate inundated from non-inundated areas.

$$MNDWI = \frac{\text{Green} - \text{SWIR}}{\text{Green} + \text{SWIR}}, \qquad MNDWI > 0$$

A binary water-bodies mask is generated. The SCL band of the Level 2A Sentinel-2 product is used to exclude clouds and shadows of the final mask.

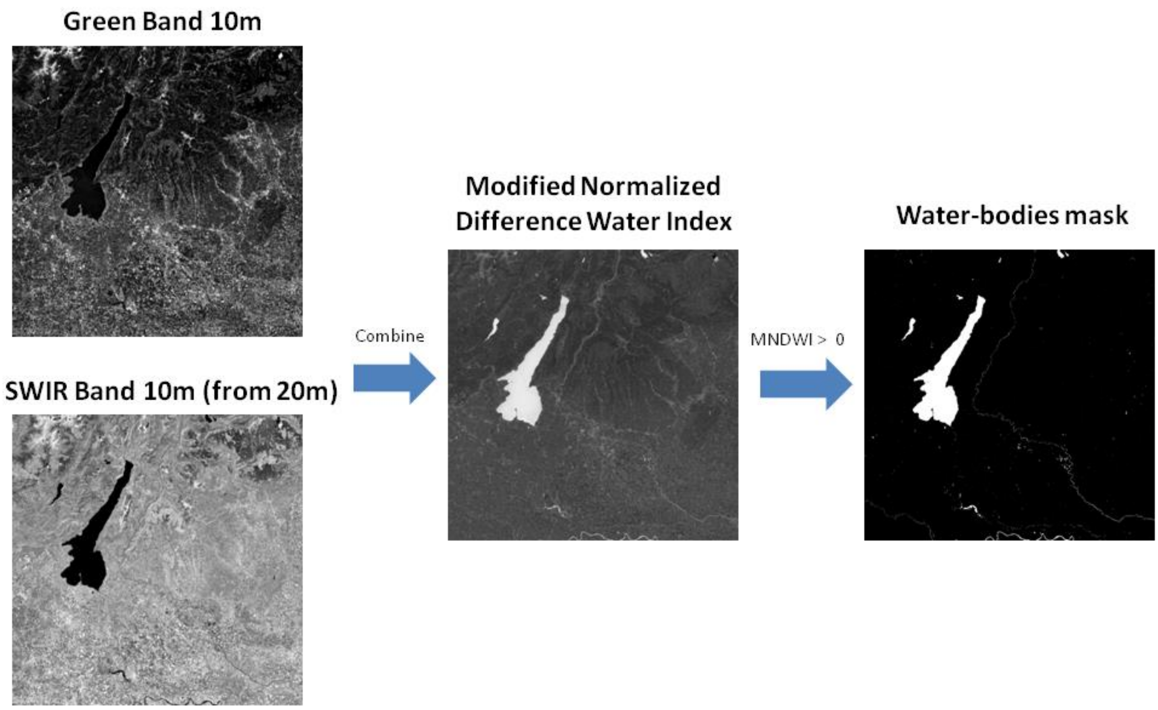

Figure 8: Framework of automatic thresholding water-bodies mask generation of a Sentinel 2 image, Vicenza (Italy).

Baseline2

The baseline framework that is developed for handling Sentinel 2 (optical) products is called discriminant analysis method. This method is a combination of Mahalanobis distance-based classification for flood mask creation and morphological post-processing for flood mask correction which aims at removing erroneous areas (Michail, et al. 2018).

In general discriminant methods split the pattern space into as many regions as the classes encompassed by the training set are and establish bounds that are shared by the spaces. Regarding discriminant analysis, it is one of the most common discriminant methods that relies on prior training of the system with objects known to which class they belong, and allows attributing an unknown sample to a certain class. The bounds established during the analysis, can be based on correlation coefficients and distances. A subcategory of discriminant analysis methods is the distance-based methods that possess superior discriminating power (Morozova, et al. 2013). The most commonly used distance measures are the Euclidean distance and the Mahalanobis distance.

As far as the mask correction step is involved, it comprises the following three processes:

- The first process is considered as a global filter that aims at correcting flood-denoted pixels in image-level. It relies on the assumption that if the percentage of pixels classified as flooded in an image is very small, and then probably these pixels are misclassified samples. So if the percentage of flooded pixels in an image is less than a preset threshold, they are set to non-flooded (the threshold has been empirically set to 5% after examining classification results of a training set).

- The second process functions as a local filter that aims at eliminating small flooded areas, that is groups of a few pixels, usually 1 to 10, classified as flooded in non-flooded areas, which are potentially false positives. It uses connected-component analysis that involves counting the number of pixels in such areas and in case it is less than a threshold (10 pixels), it marks them as non-flooded.
- The third process aims at eliminating small non-flooded areas inside flooded area, which are probably false negatives, while preserving large areas. The operations involved in this process are image dilation and erosion. Image dilation usually uses a structuring element for probing and expanding the shapes contained in an image, while image erosion is to remove the boundaries of regions of foreground pixels which translates to having smaller foreground pixels areas, and larger holes within those areas.

It should be noted that the framework considers the discriminative ability of the variance of the colors Red (R), Green (G), Blue (B) and the Near-Infrared (NIR) values of the satellite image pixels in order to separate flood from non-flood areas. The steps involved in the method are the following:

- Random selection of pixels from the training dataset.
- Representation of each pixel with a 4-dimensional feature vector (R, G, B, NI)
- Classification using discriminant analysis technique. The input is a 4 dimensional feature vector accompanied by an integer label (0, 1) signifying water existence. Regarding the discriminant functions used for classification several options are evaluated including the linear, and Mahalanobis. The output of the classification is a binary mask with 1 for flooded pixels and 0 for non-flooded.
- Validation of classifier developed on testing set of images
- Performing the three aforementioned post-processing morphological operations on masks for removing erroneous areas.

Figure 9 depicts a high-level representation of the framework.

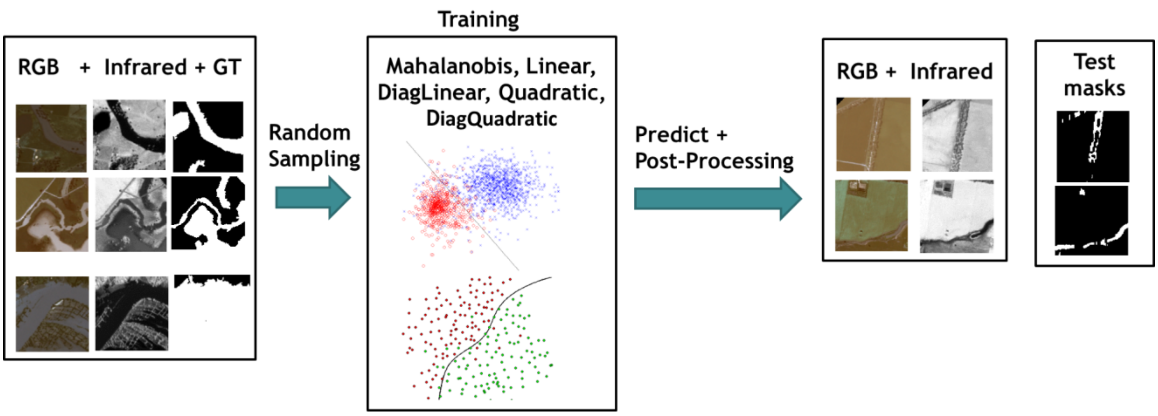

Figure 9: Framework of discriminant analysis method.

2.4.3 **Evaluation**

Sentinel 1 and Sentinel 2 (Baseline 1)

To measure the efficiency of the baseline framework, some significant flood events were selected from Copernicus Emergency Management Service[3] (EMS). The semi-automatic annotation of the flooded areas that is provided is used to compare with our results. As input to our framework the closest timely products to the above EMSR (Emergency Management Service Response) events were retrieved from ESA's Copernicus Open Access Hub[4].

Finally, the evaluation measure used for assessing the performance of the model at successfully detecting water is the Intersection over Union (IoU).

$$\mathrm{IoU} = \frac{Area\ of\ Overlap\ (true\ positive\ pixels)}{Area\ of\ Union\ (true\ positive\ and\ false\ positive\ pixels)}$$

The selected flood events are located to Peru, Greece, Ireland and Norway.

Sentinel 2 (Baseline 2)

The proposed baseline framework is developed on the satellite data of the MediaEval 2017 dataset provided within the context of the Multimedia Satellite Task[5] and is evaluated on satellite data referring to the Moroccan Demonstration Area provided within the context of the H2020 MOSES[6] (Managing crOp water Saving with Enterprise Services) Project.

Specifically, as far as the training set is concerned, the satellite images belonging to the MediaEval 2017 dataset are collected from PlanetLabs[7], and it contains 24,000 pixels (specifically 6000 pixels positive & 18000 negative pixels) randomly selected from 448 images with dimension equal to 320x320 pixels. On the other hand, the testing set comprises one LandSat7 image and four LandSat8 images that capture water volumes data that refer to the Al Massira dam located in Morocco.

Regarding the discriminant functions evaluated during the development of the classifier, the following options were tested: linear, diagonal linear, quadratic, diagonal quadratic, and Mahalanobis.

Finally, the evaluation measure used for assessing the performance of the model is the accuracy.

$$accuracy = \frac{number\ of\ pixels\ recognized\ correctly}{total\ number\ of\ pixels}$$

Accuracy is measured both on full image and solely on the dam and surrounding area. This is realized because the annotation provided refers only to the dam region, while no information on other water areas exists in the LandSat product.

---

[3] https://emergency.copernicus.eu/

[4] https://scihub.copernicus.eu/

[5] http://www.multimediaeval.org/mediaeval2018

[6] http://moses-project.eu/moses_website/

[7] https://www.planet.com/

2.4.4 **Results**

Sentinel 1

Table 1 contains the IoU evaluation method of the generated water-masks compared to EMS annotation areas. The array demonstrates for each EMSR event the automatic threshold of the thresholding algorithm that separated inundated from non-inundated area, as well as the achieved IoU score.

Table 1: Water-masks generation IoU evaluation accuracy.

Published on	Location	EMSR	Threshold (dB)	IoU (%)
2017-03-30	Parachique (Peru)	EMSR199	-18.13	83.5
2016-01-11	Ballinasloe (Ireland)	EMSR149	-16.95	78.4
2018-05-14	Lillestrom (Norway)	EMSR283	-19.59	69.0
2018-06-29	Chrisoupoli (Greece)	EMSR292	-17.49	69.9

Observing the results we conclude that the separating threshold tends to fall between -17 and -19.5db.

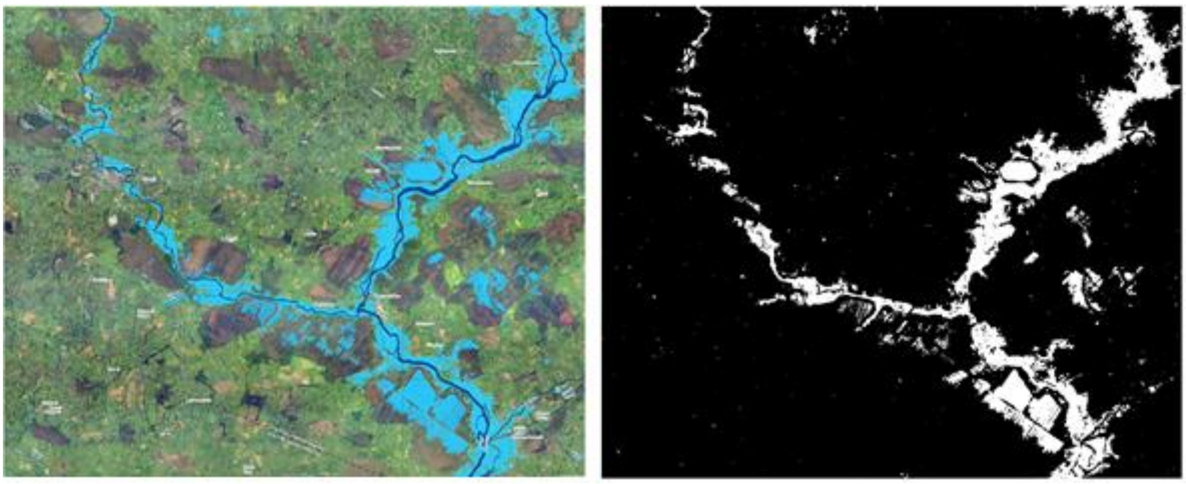

Figure 10. Overview image of Ballinasloe (Ireland) as provided by EMS (left) and its water-bodies mask, calculated by the multi-tiles thresholding technique (right).

Sentinel 2

Baseline1

Table 2 contains a comparison among the IoU performance of the baseline methods of Sentinel-1 Auto Thresholding, Sentinel-2 MNDWI and Sentinel-2 extracted water of the SCL band, compared to the delineation image provided by Copernicus ESA EMS for a specific flood event.

The low performance here is due to the fact that the EMS delineation (annotation) image was not timely close to the used Sentinel-1 and Sentinel-2 satellite images. Especially, the optical Sentinel-2 satellite required a cloud free image.

For this event, the MNDWI method demonstrated better performance compared to the marked as water areas of the included SCL band of the Sentinel-2 product.

Table 2: Compare the IoU of the three different water-bodies mask generation methods.

Published on	Location	EMSR	S1 Auto-Thres	S2 MNDWI	S2 SCL
2018-02-24	Farkadona (Greece)	EMSR271	0,64	0,53	0,45

Baseline2

Table 3 contains the accuracy values of the discriminant analysis method applied on the testing set for several cases including different discriminant functions, different annotation areas and finally with and without applying post-processing.

Table 3: Discriminant analysis accuracy

		Linear	Diagonal Linear	Quadratic	Diagonal Quadratic	Mahalanobis
Full Image	Mask before post-processing	56.2676	45.8455	50.1789	34.8459	51.1556
	Mask after post-processing	56.1781	45.6994	51.8167	36.9312	53.0569
Dam Region	Mask before post-processing	73.4600	73.8431	87.4349	59.5132	87.7580
	Mask after post-processing	73.2535	73.3361	88.9611	62.2060	**89.0873**

After a careful observation of Table 3, we can draw the following conclusions:

- Generally, post-processing operations improve the mean accuracy
- Mean accuracy of full image is low as annotation refers to dam area, which means that we are not informed on the existence of water for the area outside the dam
- Use of Mahalanobis distance in the discriminant analysis produces better results than the other distances

Although the results were quite satisfying, the presented framework is considered the baseline for flood monitoring. Therefore in the next deliverable D4.4 (due on M33), we are planning on evaluating Convolutional Neural Networks that can be used for classifying the image as a whole, R-CNNs that can used for object detection and in particular for recognizing the flooded area and other type of CNNs.

Finally, we should note that in order to evaluate the flood detection algorithm, the need of more annotated data is of critical importance. Thus, in order to tackle this issue and also after receiving the reviewer comments, we evaluated the possibility of accessing the services of the Disaster Charter. Disaster Charter provides satellite data to those affected by natural or man-made disasters through registered organisations, for use in monitoring and response activities. However, a prerequisite for registering to the Disaster Charter is to be a national disaster management authority or its delegated agency in that country. Thus, the only EOPEN partner that can apply for this is AAWA, which has already taken the proceedings for registering. Apart from Disaster Charter, the use of Copernicus specific services is also of interest, such as the EMS. Currently the only entity in Italy that can access these data is the

National Civil Protection Department (NCPD) that receives requests from the regional authorities. As a result, AAWA has to contact the NCPD in order to get these data and this procedure introduces significant delays in receiving the activations.

2.5 Change detection in PUC2 - "Food Security through Earth Observation dataset"

The *PUC2- Food Security through Earth Observation (EO) datasets* seeks to address the monitoring of food security parameters in South Korea, focusing exclusively on rice production. PUC2 has identified three principal target products to be developed, based on the user requirement analysis, as detailed in deliverables D2.1 and D2.2, including rice paddy mapping, rice phenology extraction and rice yield estimation. In comparison to the other two PUCs, this is the most intensive in terms of EO-based developments; with most of the PUC2 implementation effort expected in WP7 under Task 7.1. The identified EO-based target products go beyond the premise of change detection. Nevertheless, within the framework of the envisaged food security monitoring system, the below elaborated change detection methodology integrates as a stepping-stone of the T7.1 developments.

2.5.1 Related Work/ State of the art in food Security using EO data

Despite the fact that the presented methodology does not concern the entirety of EO-based solutions of PUC2, it is deemed necessary to provide a more general background on the state-of-the-art in rice monitoring through EO and place the hereby presented methodology within the processing chain of the PUC2 system.

Manual field visits for survey are costly and time demanding when compared to information gathered through earth observation means. Remote sensing is one of the most effective technologies to map the extent of crops. Rice area mapping at the parcel, regional and national scale has been extensively studied in the past, through several approaches, including mono-temporal and multi-temporal classification schemes that utilize both optical and microwave/SAR data (Qin et al., 2015, Nguyen et al., 2015, Neetu et al., 2014). Tian et al. (2018) have introduced a novel multi-season paddy rice mapping method using Sentinel-1 and Landsat-8 data under a k-means unsupervised classification scheme. Torbick et al. (2017) have produced an updated land cover map, including the rice class, fusing Sentinel-1, Landsat-8 OLI and PALSAR-2 data using a random forest classifier. Finally, Pazhanivelan et al. (2015) have introduced a robust rule-based classification for mapping rice area with multi-temporal, X-band, HH polarized SAR imagery (COSMO Skymed and TerraSAR X), with site-specific parameters.

It becomes apparent that multiple studies have utilized both optical and SAR data for the monitoring of rice, with a lot of research on the usage of Sentinel data, exploiting their unprecedented characteristics in temporal and spatial resolution, but also their open access nature. Studies have used multiple classification techniques including neural networks, supervised and unsupervised machine learning techniques, but also custom rule-based systems. Nevertheless, all studies had to manually collect their training and validation samples, undermining the design of a fully transferable and site independent framework of application for the described systems.

In this regard, and bearing in mind EOPEN's scope wrt scalability, reproducibility and transferability, we decided to design a more dynamic rice monitoring system that is largely independent of hard-to-attain non-EO information. In this deliverable we suggest a land cover map update mechanism based on change detection to produce appropriate training and validation datasets for any given year of inspection via utilizing older ground truth information.

2.5.2 **Framework**

In this deliverable, we present a framework developed for updating through EO assisted change detection the rice map of year 2015 for the South Korean provinces of Dangjin and Seosan to reflect the reality of year 2018 (the selected year of inspection). Detailed land cover maps at the parcel level, distinguishing between rice and non-rice cultivations, are produced every few years in South Korea. However, in order to systematically produce, for any given year, the envisaged products of rice mapping and rice yield estimation, updated land cover maps are required for the appropriate training and validation of the machine learning algorithms utilised. The output product of the below methodology does not attempt to accurately classify rice fields for 2018 but merely delete obvious changes in the land cover map of 2015. Therefore, using only rice pixels from the 2015 land cover map we eliminate outliers for the updated 2018 land cover map. This way the training dataset is refined for the machine learning algorithms to follow for the rice paddy classification product.

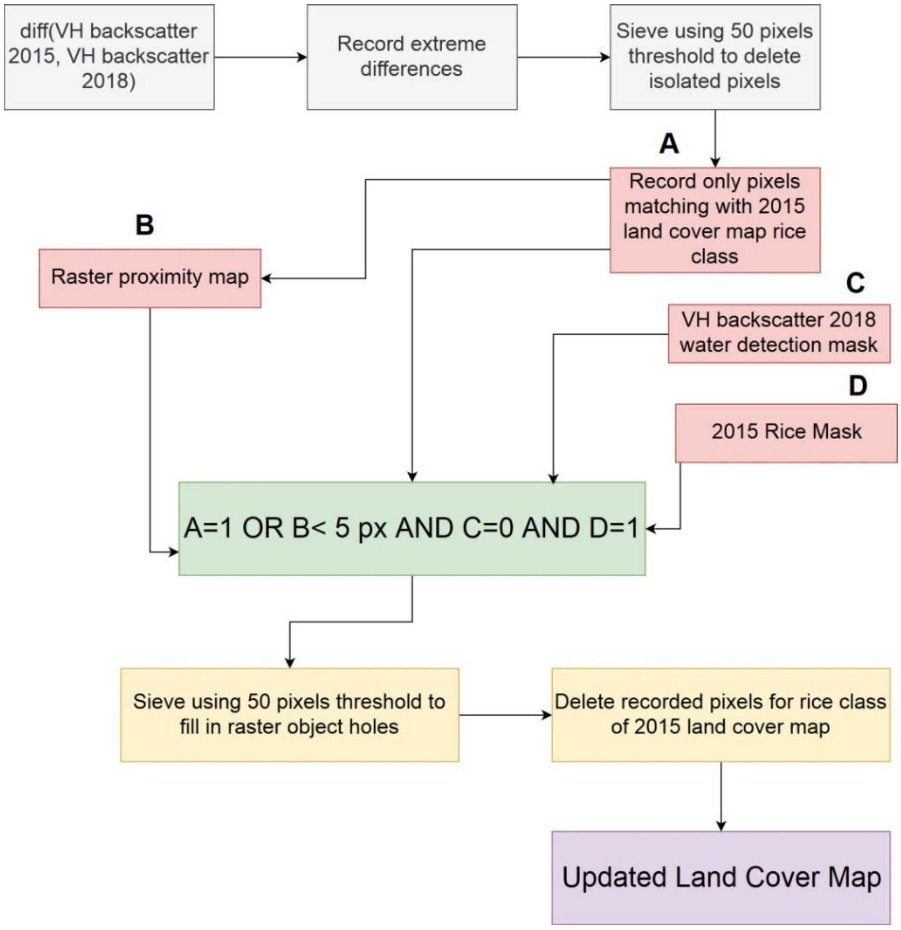

Figure 11: Framework of land cover map update method

Figure 11 illustrates the workflow for the *land cover map update* methodology. The processing was done using GDAL, GRASS and SAGA libraries. The processing chain starts by taking the difference of the VH polarized backscatter of Sentinel-1 images in mid-June of 2015 and 2018 respectively. VH polarization is especially sensitive in the water concentration of inundated rice paddies. Imagery sensed in mid-June was selected as it is the period that paddies have been recently flooded and rice fields are easily detectable. The preprocessing of Sentinel-1 imagery was done in SNAP following the below depicted steps. The two images have been additionally co-registered in order to ensure 1-1 pixel match.

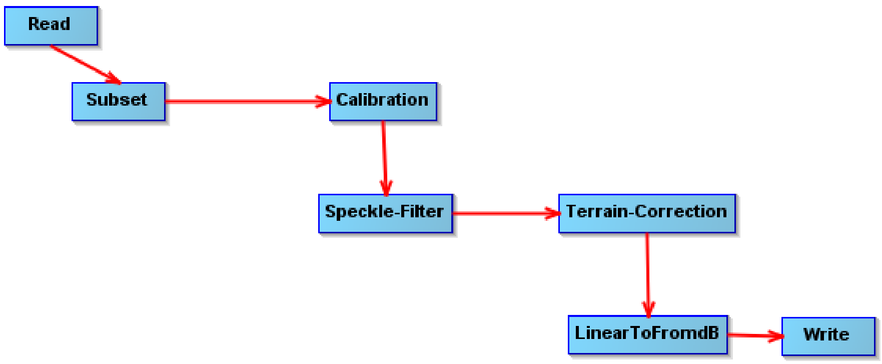

Figure 12: Pre-processing of Sentinel-1 imagery

From the difference product of the two VH backscatter images we record only the pixels of extreme differences, using the values of bottom 2% and top 98% of the cell data range as thresholds. Then the GDAL function *Sieve* is utilised in order to remove raster polygons smaller than 50 pixels and replace them with the pixel value of the largest neighbour polygon, thus eliminating island pixels and filling raster polygon holes. The remaining outlier pixels are superimposed with the 2015 rice mask, keeping only matching records. Hence, we end up with a set of pixels identified as rice in 2015 that appear non-rice for 2018, as illustrated in the following figure.

Figure 13: Unrefined outlier detection product

To further refine the above illustrated product, it was attempted to fill in missing pixels within the identified non-rice parcels. A raster proximity map was generated to indicate the distance from each pixel to the nearest pixels identified as outlier (target pixel). Then pixels that haven't been yet identified as outliers and have a distance of fewer than 5 pixels from outlier pixels are recorded (B in Figure 11). Product C in Figure 11 refers to the automated water detection via thresholding, as described in the PUC1 scenario above, but for the VH backscatter of 2018 this time. Rice is inundated and therefore the algorithm also works for identifying flooded agricultural land. Even though the water mask captures a great deal of the total number of rice parcels in the area of interest, it is not perfect, and for this reason it is simply a part of a more sophisticated change detection methodology.

Figure 14: VH backscatter 2018 water mask via thresholding

Product D in Figure 11 refers to the identified rice pixels in the 2015 land cover map, in other words the 2015 rice mask. Combining products A-D under the Boolean expression in Figure 11, we record all outlier pixels of the unrefined detection product (Figure 13), but also pixels that are concurrently not included in the VH-backscatter 2018 water mask and have a proximity value of fewer than 5 pixels (Figure 14). The final product is illustrated in the Figure 15.

2.5.3 **Evaluation**

The evaluation of the showcased product can be twofold:

1) Evaluate using photointerpretation to determine if identified outliers do not indeed represent land covered by rice. Photointerpretation is performed against a pseudocolor RGB composite VH, VV, VH/VV backscatter products (Sentinel-1) and the true colour composite of 10th September 2018 (Sentinel-2) when differences between rice and non-rice fields are visually maximum.

2) Evaluate if indeed the updated training set offers better accuracy results for the machine-learning based rice paddy classification. Updating a land cover map to function as a training and validation dataset can entail certain pitfalls. For instance prior to the implementation of the presented method, another land cover update methodology was tried out. Unsupervised classification (isocluster-maximum likelihood) identified rice pixels to be used for training the

supervised classification algorithm at the next stage. The identified pixels, though indeed rice in their vast majority, were not random enough to encapsulate all possible instances of rice. Therefore, the supervised classification algorithms were overfit to the training sample of a particular type of rice, resulting in considerable misclassifications (low producer's accuracy, high user's accuracy).

The suggested methodology attempts to circumvent such issues, by not selecting pixels that certainly represent rice, but merely eliminating instances of great land cover change. This way the updated training dataset, although not perfect, is characterized by significantly reduced noise while maintaining its randomness to avoid overfit.

This evaluation method requires overall accuracy estimations for the supervised classifications (at the 2nd stage) trained with both 2015 land cover map and 2018 updated land cover map for 2018, evaluated against the same validation dataset (to be constructed via photointerpretation). The above presented evaluation framework remains to be performed in the coming months, aligning with the timeline of PUC2 related implementations in Task 7.1.

2.5.4 **Results**

The figure below (Figure 15) illustrates the final refined outliers detection output at the top left corner. Rice fields based on the 2015 land cover map and the unrefined outlier detection are also shown for visual comparison.

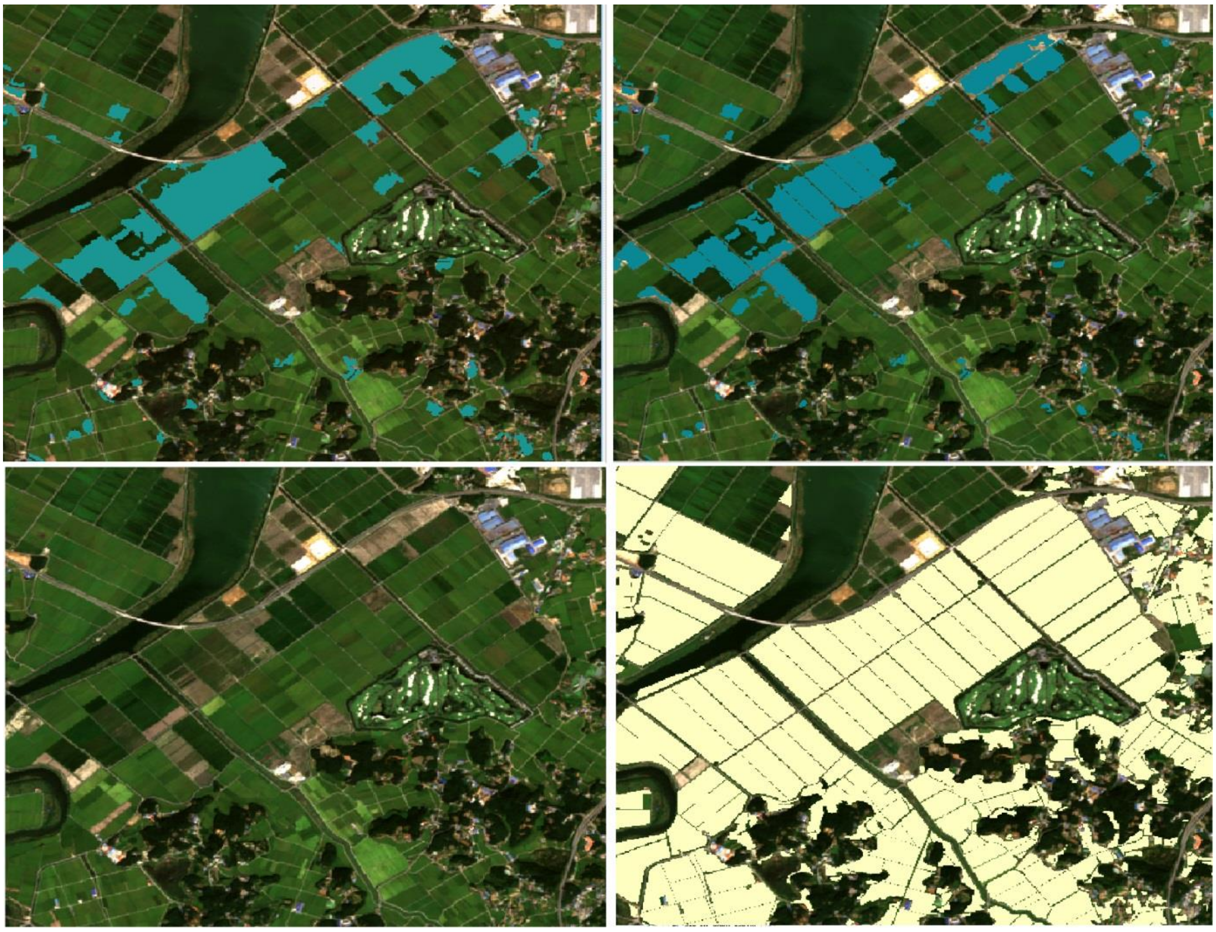

Figure 15: top left - refined outliers detection, top right – unrefined outliers detection, bottom left – true colour composite September 2018, bottom right – rice pixels based on 2018 land cover map

The illustrated figures throughout this section represent an indicative snapshot of the area of interest. The method was tested in a total area of 7,185 ha land that was found to be rice in 2015. The detected outlier pixels amounted to 396 ha, therefore resulting in a 5.5% change. This percentage does not only refer to rice parcels that altered to non-rice cultivations in 2018, but also pixels on the parcel borders that were classified as rice due to less than optimal parcel digitization in the 2015 land cover map. This further refines and removes noise from the training dataset.

2.6 Change detection in PUC3 - "Monitoring the Climate Change through Earth Observation Flood Monitoring"

In general, PUC3 covers a wide range of phenomena to be monitored. However, this section focuses on identifying whether a road is passable or not due to water floods. This scenario is closely related to the road maintenance that involves monitoring of road conditions during extreme weather events. In the current deliverable and similarly to the PUC1, we describe road possibility approaches that will be applied for detecting obstacles and disasters on the road network by considering EO data. The other aspects of PUC3 involve other data types (e.g. weather data) and will be reported within the context of other deliverables (including D3.2Meteorological and climatological data acquisition report).

2.6.1 State of the art in road passability using EO data

A significant issue that occurs during extended rain events is the one of the flooded passages between major waypoints, making the areas unreachable by conventional means of transportation.

The latest approach of the road passability issue involves the usage of deep convolutional neural networks (DCNN). In the work at (Dias et al., 2018) 10 CNNs were trained on ImageNet to predict if it is possible to travel through a flooded area. The impact of transfer learning was demonstrated compared to the limited size of training dataset. For the annotation, oriented squares of the two waypoints were generated, so the model could learn how to find a path between them. The work at (Said et al., 2018) was also based on DCNN (Inception v3) and transfer learning technique. To improve the road patches classification performance, various augmentation operations were performed, including flipping of image and alteration of the brightness. The RMSprop optimizer is used, which allows an adaptive learning rate during the training process.

Alternatively, identifying remotely the passability between two way points can be investigated as a combination of the flood monitoring (2.4.1) and the road extraction topic. The road components can be extracted from Satellite images using Laplacian of Gaussian operator (Babu et al., 2016). The image is pre-processed to identify the color space components. To obtain more details of the image, fusion of a panchromatic and a multispectral image of an area is performed, where objects are identified using the hue-saturation-value (HSY) color models components. Hue and luminance may have similar values at distinguishing roads from sandy regions but can be resolved by using saturation. At the work of (Henry et al., 2018) 3 different Fully-Convolutional Neural Networks (FCNNs) were trained from scratch, where a considerable performance drop is noticed when using weights pre-trained on ImageNet, due to the different nature of SAR images compared to optical ones.

Contrary to these approaches, we aim to infer whether a satellite image patch contains a passable road segment or not, without the need to segment the image patch into "road" and "no-road" regions.

2.6.2 **Framework**

In this section, we present the baseline framework that is developed for identifying whether a road is passable due to water using satellite images. The framework used considers pre-trained Convolutional Neural Networks (CNNs). Several networks were evaluated in order to decide on the model capturing best the "road passability" problem (Moumtzidou, et al. 2019).

This task focuses on detecting the passability of roads due to extreme weather conditions, rather than identifying the actual roads. Nevertheless, OpenStreetMap API has been investigated in order to check whether road information can be extracted and used to potentially improve the road passability framework. The outcome of the study on OpenStreetMap showed that it is technically possible to extract the road network of any country. Thus, we will explore in the future and beyond the EOPEN project how and whether the use of OpenStreetMap can improve the performance in the road passability problem.

In the current version of the road possibility problem, we experimented with the following models: VGG-19 (Simonyan and Zisserman, 2014), Inception-v3 (Szegedy, et al. 2016), and ResNet (He, et al. 2016). VGG is a convolutional neural network model developed by the University of Oxford that was originally developed for the ImageNet dataset. The model involves 19 layers and it has as input images of size 224 x 224 pixels. Inception-v3 is another ImageNet-optimized model. It is developed by Google and has a strong emphasis on making scaling to deep networks computationally efficient, having as input images with size 299 x 299 pixels. Finally, ResNet-50 is a model developed by Microsoft Research using a structure that uses residual functions to help add considerable stability to deep networks, using as input 224 x 224 pixel images. For each of the aforementioned networks, we performed fine-tuning which involved removing the last pooling layers and replacing it with a new pooling layer with a softmax activation function with size 2 given that our aim is to recognize whether there is evidence of road passability or not.

The implementation of the new models is realized with TensorFlow[8] and the opensource neural network Python package Keras[9]. In general, the use of the Keras package aims at simplifying the training procedure by allowing effortless modification of the network structure and the pre-trained weights, freezing the weights in the imported network and eventually training the weights in the newly added layers. Thus it is possible to combine existing knowledge from the imported weights with the gained knowledge from the domain-specific collection of satellite images with ground-truth annotation on road passability. Figure 16 depicts the fine-tuned VGG-19 architecture.

[8] https://www.tensorflow.org/
[9] https://keras.io/

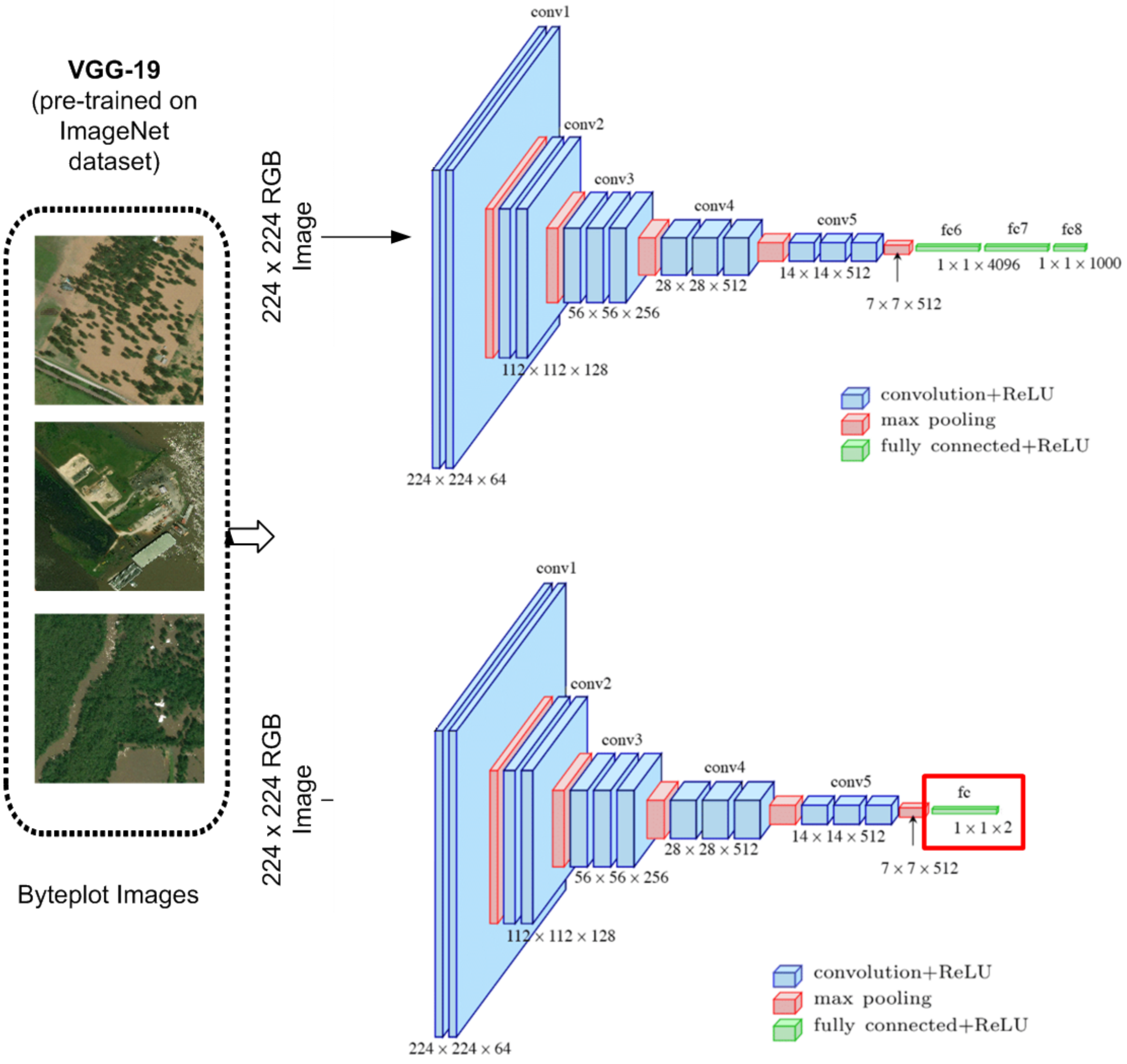

Figure 16: DCNN-based road passability framework.

2.6.3 **Evaluation**

The proposed framework is developed and evaluated on the MediaEval 2018 Satellite Task "Emergency Response for Flooding Events"[10] - data for "Flood detection in satellite images". The satellite imagery was retrieved from the WorldView Satellite archive[11] and captured the event "Hurricane Harvey" in 2017 during the period 2017-08-30 till 2017-09-02. The geographical location was from the Houston area and the ground sample distance 30 cm (1 pixel = 30 cm).

The dataset consists of 1,437 satellite image patches of flooded areas that were manually annotated with a single label to indicate whether the road depicted is passable or not due to

[10] http://www.multimediaeval.org/mediaeval2018/multimediasatellite/

[11] https://www.satimagingcorp.com/satellite-sensors/worldview-3/

floods. The dataset was randomly split into a training a validation set. The training set contained 1,000 images, while the validation set contained the remaining 437 images.
Several experiments were run in order to find the best performing model. The parameters that were tuned included the learning rate, the batch size and the optimizer function. Specifically, the values considered for the aforementioned parameters were the following: learning rate values = 0.001, 0.01, 0.1, batch size values = 32, 64, 128, 256, and the optimizer functions = Adam, Stochastic Gradient Descent (SGD). Finally, the epoch was set to 35 and the loss function considered was the sparse categorical cross-entropy.

Finally, the evaluation measure used for assessing the performance of the model is the accuracy.

$$accuracy = \frac{number\ of\ pixels\ recognized\ correctly}{total\ number\ of\ pixels}$$

2.6.4 **Results**

To evaluate the performance of the different networks we considered accuracy as the evaluation metric. The results of our analysis are shown in Table 4 and Table 5 and they show the accuracy of the train and the validation set for the four networks (i.e. VGG-19, Inception v3, ResNet-50, ResNet-101) for two widely used optimizers, i.e. Adam and SGD. Specifically, Table 4 shows how the learning parameter affects the performance of the networks. After observing carefully the table we can conclude that the networks perform better for the lower values of the learning rate, as they reach an average accuracy of 81.2% and 78.5% for learning rates 0.001 and 0.01 respectively.

In the sequel, we experimented with the batch size parameter and observed the impact on the networks accuracy (Table 5). The conclusion that can be drawn from this experiment is that the increase of the batch size generally improves the accuracy. The best values of accuracy are achieved by ResNet-50 for batch size 256 and Adam optimizer (88.2%) and the Inception v3 for batch size 128 and Adam optimizer (89.9%) (highlighted in bold). However, the accuracy of the validation set for the Inception v3 is significantly lower than that of ResNet-50, probably due to over-fitting reasons. Finally, Figure 17depicts a result of the road passability framework applied on an image belonging to the test set.

Table 4: Neural networks accuracy for different learning rate values.

		Learning rate 0.001		Learning rate 0.01		Learning rate 0.1	
DCNN	Optimizer	Dev. Set Acc.	Valid. Set Acc. Dev.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.
VGG-19	Adam	0,85	0,6911	0,5640	0,4851	0,5700	0,5904
VGG-19	SGD	0,8600	0,7140	0,8380	0,7277	-	-
Inception_v3	Adam	0,796	0,6018	0,8120	0,5973	0,4190	0,4348
Inception_v3	SGD	0,7050	0,6590	0,8100	0,6499	0,8040	0,5995
ResNet-50	Adam	0,872	0,6247	0,8400	0,6453	0,5670	0,5973
ResNet-50	SGD	0,789	0,6865	0,8470	0,5538	0,8060	0,6796
ResNet-101	Adam	0,866	0,5515	0,8450	0,4668	0,7470	0,6590

ResNet-101	SGD	0,799	0,6041	0,8700	0,4668	0,8450	0,5858

Table 5: Neural networks accuracy for different batch size values for best performing learning rates.

			Batch size 32		Batch size 64		Batch size 128		Batch size 256	
DCNN	Learnin g rate	Opti mizer	Dev. Set Acc.	Valid. Set Acc. Dev.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc.	Dev. Set Acc.	Valid. Set Acc. Dev.
VGG-19	0,001	Adam	0,861	0,7666	0,8610	0,7666	0,8610	**0,7667**	-	-
VGG-19	0,001	SGD	0,876	0,7071	0,8630	0,7117	0,8740	0,7162	-	-
Inceptio n_v3	0,01	Adam	0,788	0,6247	0,8610	0,5789	**0,8990**	0,5629	0,8800	0,5378
Inceptio n_v3	0,001	SGD	0,792	0,5950	0,8330	0,6224	0,8480	0,5973	0,8550	0,5995
ResNet-50	0,01	Adam	0,833	0,4943	0,8640	0,6957	0,8720	0,7094	**0,8820**	0,7323
ResNet-50	0,001	SGD	0,804	0,6911	0,8310	0,7094	0,8390	0,7140	0,8390	0,7185
ResNet-101	0,1	Adam	0,86	0,5492	0,8710	0,5126	0,8850	0,5126	0,8890	0,4989
ResNet-101	0,001	SGD	0,789	0,5835	0,8260	0,5995	0,8380	0,5881	0,8390	0,5812

Although the results were quite satisfying, there is room for improvement which translates to higher accuracy values. Therefore in the next deliverable D4.4 (due on M33), we are planning on evaluating the same Convolutional Neural Networks but having a more fine-grained annotation by splitting the images in batches and precisely pointed which part of the image contains evidence that a road is passable or not. We are also planning on addressing the task as a two-phase problem, where there are two networks; one for addressing the issue of road existence and the second for checking whether the existing road is passable or not. Finally, we are planning on evaluating other types of CNN networks, specifically, R-CNNs that can used for object detection and in particular for recognizing the flooded area in a road.

Figure 17: Example of road passability algorithm.

3 CONCEPT AND EVENT DETECTION IN NON EO DATA

This section applies to the social media data (i.e. data retrieved from Twitter) and includes two independent tasks. The first is concept detection and extracts high-level content (i.e. concepts) from textual and visual low-level information in order to be able to retrieve relevant content and to mark multimodal content as relevant or not to a target event (i.e. flood, heatwave). The second is event detection and aims at providing notifications on specific events, as defined by the user scenarios (e.g. flood events) by exploiting contextual data, spatiotemporal and weather information and by applying burst analysis techniques.

3.1 Concept Detection

Concept detection is the problem of identifying to which of a set of categories (i.e. classes) a new observation belongs to, by considering as known the membership of as of observations known as training data. Concept detection is a two-step process that requires at first the construction of a model by using a training set of the target category and then the application of the model for classifying previously unseen data (testing set). The algorithm that implements classification is known as a classifier, i.e. a function that maps one observation to a pre-defined class. Usually, the observations (or instances) are described using a feature vector that is comprised of measurable properties (or features) of the instance. These features may be binary; categorical; integer-valued; or real-valued. Thus, in case of images, the feature values might correspond to the pixels of an image and in case of text, the feature values might be occurrence frequencies of different words.

In the sequel, we present an overview of the relevant work on visual concept detection and the framework applied for extracting visual concepts. The evaluation section follows which includes a short description of the datasets used, the experiments realized, the results produced and finally the conclusions are drawn.

3.1.1 State of the art in Concept Detection

Image classification involves the use of visual concept detection algorithms based on low-level features and classifiers for deciding whether an image shows evidence of high-level concepts such as water, woman, hand, sky that are chosen from a pre-defined concept pool. In this section, we present an overview of the state of the art methods for concept detection in images, and then we present the framework proposed within EOPEN.

As mentioned above, concept detection involves two steps with the first being the extraction of visual features, and the second the training of classifiers for each concept using a ground-truth annotated training set, and the application of these classifiers to unlabeled images. The output of the concept detection is a set of confidence scores that measure the probability that different concepts appear in the image. Thus, the first step is feature extraction and the second is the building of the classification model.

As far as feature extraction is concerned, it refers to the methods that try to capture the visual content of images. Visual descriptors can be divided in two main groups: hand-crafted and DCNN-based descriptors. Hand-crafted features can be divided into categories global and local descriptors based on the whether they capture the global characteristics of an image or detect and capture the local salient points. Common global descriptors are the

MPEG-7 descriptors, and the Grid Color Moments. On the other hand, common local descriptors are the SIFT descriptor (Lowe, 2004), the SURF descriptor (Bay et al., 2008) and their variations. It is quite common to apply in the case of local descriptors a clustering algorithm after the feature extraction in order to form a vocabulary of "visual words" that essentially translates to global descriptor. The best known approaches for visual word assignment are the "bag-of-word" (BoW) representation (Qiu, 2002), the Fisher vector (Perronnin et al., 2010) and the VLAD (Jegou et al., 2010). Regarding DCNN-based features, they are the most recent trend in feature extraction and image representation and they outperform the hand-crafted features in most applications. These features learn directly from the raw image pixels using Deep Convolutional Neural Networks (DCNNs). DCNNs consist of many layers of feature extractors and can be used both as standalone classifiers, i.e., unlabeled images are passed through a pre-trained DCNN that performs the final class label prediction directly, or as generators of image features, i.e., the output of a hidden layer of the pre-trained DCNN can be used as a global image representation (Simonyan, 2014; Markatopoulou, 2015). The latter type of features is referred to as DCNN-based and they are usually preferred due to their high performance both in terms of time and accuracy. Several DCNN software libraries are available, e.g., Caffe (Jia, 2014), MatConvNet (Vedaldi, 2015), and different DCNN architectures have been proposed, e.g., CaffeNet (Krizhevsky, 2012), GoogLeNet (Szegedy, 2015).

Classification step is the second step of the multimedia concept detection process, and it involves the construction of models by using the low-level visual features, and then the application of these models for image labelling. Common classifiers that are used for learning the associations between the image representations and concept labels are the Support Vector Machines (SVM) and Logistic Regression (Markatopoulou, 2015). SVMs are trained separately for each concept, on ground-truth annotated corpora, and when a new unlabeled image arrives, the trained concept detectors will return confidence scores that show the belief of each detector that the corresponding concept appears in the image.

3.1.2 **EOPEN Framework in concept detection**

In the employed framework, we trained a 22-layer GoogleNet network (Szegedy, 2015) on 5055 ImageNet concepts (Pittaras, 2017), which are a subset of the ImageNet "fall" 2011 dataset[12] that was trained originally on 32,326 concepts. The subset of the 5055 concepts was produced by considering the following rules: a) concepts that were very similar were merged, for example all different dog breeds (e.g. Shih-Tzu, Pekinese, Maltese dog) were removed and only the concept dog was kept. The same philosophy was followed for other animals and plants as well, b) concepts that correspond to scientific terms were removed, for example biological terms such as eukaryote, prokaryote, sporozoite etc., and c) concepts with a very few number of positive images were removed. Thus, the classification layer of the trained network which is a fully connected layer had dimension equal to 5055. It should be noted that after the classification layer a *softmax* function was applied following the GoogleNet architecture.

[12] http://academictorrents.com/details/564a77c1e1119da199ff32622a1609431b9f1c47

In the sequel, in order to target the TRECVID Semantic Indexing SIN 2013 task[13], we reduced the number of concepts to the 345 SIN TRECVID concepts[14]. Towards this direction, we performed fine-tuning and replaced the classification layer with dimensionality 5055 with a classification layer with dimension equal to 345. This layer was also a full connected layer. It should be noted that GoogleNet has by default three classification layers. Thus, in order to keep the same architecture, we considered as well three classification layers with dimension equal to 345. Finally, based on research realized on fine-tuning (Pittaras, 2017), we added an extra layer before the classification layers, as it seems to boost its performance. Similarly to the classification layer, this layer was also fully connected. Figure 18 depicts the original GoogleNet architecture and the Fine-Tuned GoogleNet after applying all the changes mentioned above.

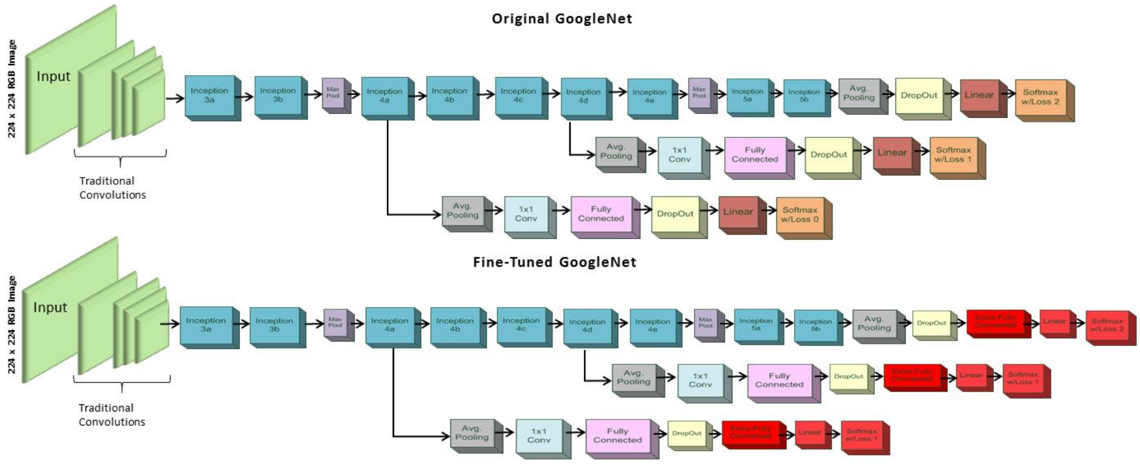

Figure 18. Original GoogleNet architecture and Fine-Tuned GoogleNet (layers in red have been added or are replacing layers existing in the original GoogleNet).

The aforementioned-module DCNN is the principal component of the concept detection framework depicted in Figure 19. Specifically, the concept detection framework receives as input an image. Then, the fine-tuned DCNN is tested on the specific image and a list of concepts along with their probabilities is produced. In the sequel, in order to limit the number of concepts tagging an image, we consider a double threshold that considers both the number of concepts that can appear on an image as well the probability value of the concepts. Specifically, we consider only the top 10 concepts with the higher probability given that their probability is higher than 0.1. Both thresholds emerged after a careful observation of the concepts and their probabilities as well as by the fact that 10 tags are sufficient for describing an image.

Finally, Figure 20 depicts the results of concept detection framework on several images retrieved from Twitter.

[13] https://www-nlpir.nist.gov/projects/tv2013/index.html

[14] http://www-nlpir.nist.gov/projects/tv2012/tv11.sin.500.concepts_ann_v2.xls

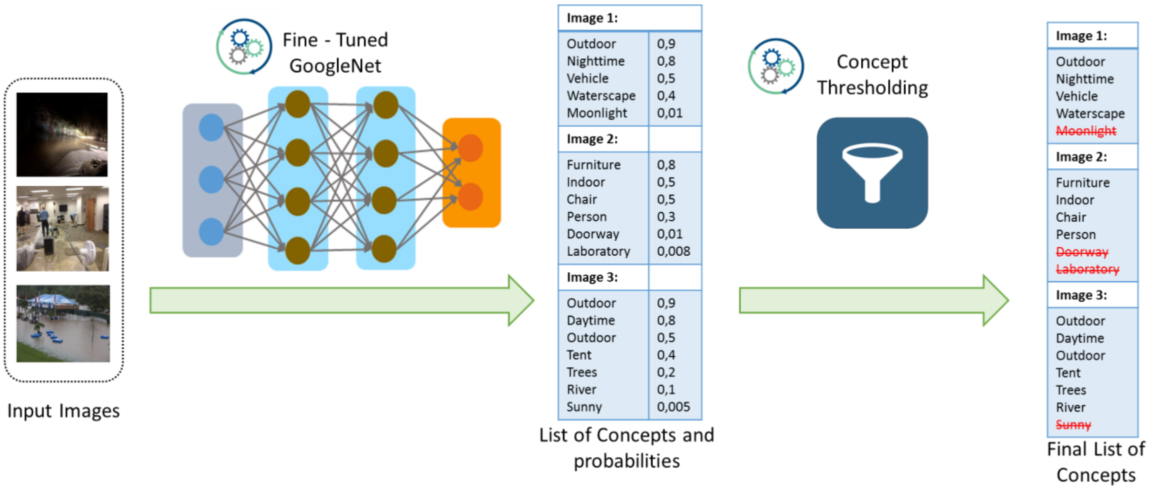

Figure 19. Concept detection framework.

Figure 20. Examples of concept detection on Twitter images (concepts can be found inside the red bounding box)

3.2 Event Detection

Nowadays, water authorities take measures aiming at reducing risks by minimising the possible damage effects and losses that may result from a flood event. Monitoring a flood event requires not only weather, sensor, EO data, and messages from first responders, but also social data from social media platforms. The constantly growing popularity of microblogging, and particularly of the Twitter platform, has led to a collaborative network of news distribution between interested users (Bruns et al., 2012b). At the same time, organizations have developed a new communication channel with their public using Twitter (Saffer et al., 2013). The wide adoption of Twitter by both individuals and authorities can also be reflected in the case of natural disasters (Bruns & Burgess, 2014) and the large amount of posts generated during such events has motivated the research community to investigate on how this data can be proven useful for crisis management. Focusing on real flood incidents, (Bruns et al., 2012a) and (Takahashi et al., 2015) conclude that Twitter has a leading role in crisis communication due to the timely dissemination of critical information.

Regarding the analysis of tweets that are produced during floods, (Saravanou et al., 2015) use geo-tagging and visual analytics tools to discover flood-stricken areas, (Vieweg et al., 2010) employ information extraction strategies to detect the intention of a tweet, i.e. an advice, an evacuation order, etc., and (Cheong & Cheong, 2011) perform social network analysis techniques to identify active players and how they affect the sharing of crisis information. Other works, e.g. (Kongthon et al., 2012) and (Moumtzidou et al., 2018), try to estimate whether text or images from tweets are relevant to floods, while (Reuter & Schröter, 2015) examine the re-tweet ratio to mine related tweets. Contextual information from social media images has also been clustered into events (Schinas et al., 2016), a technique which allows for detecting events from unstructured data streams of social media content. However, we employ spatiotemporal event detection to explore the frequency of collected and relevant tweets in the area of Italy over a one-year period (Andreadis et al., 2018).

3.2.1 **Framework**

Our target is to exploit actual tweets in order to detect if and when a flooding event is occurring, but also to reveal more insights on the event. This will enhance the flood situational awareness and support the authorities' preparedness. The data collection from Twitter content is acknowledged to the beAWARE (H2020-700475) project[15], aiming to demonstrate the availability of relevant content in social media platforms. However, the events (peaks in the time series) and word cloud of locations are generated in the EOPEN project. The relation between EOPEN with beAWARE, including the similarities and the differences of the modules developed in both projects, is presented in detail in the Appendix A.2.

During the first period of EOPEN, we have performed an analysis on the collected tweets over one year in Italy to prepare a suitable set of measurements that involve not only textual and visual data, but also time series that will demonstrate the existence of an event in specific locations and periods of time.

Let $x_1, x_2, \dots, x_T$ be a sample of observations of the number of collected Twitter posts, associated to days $1,2, \dots, T$ respectively. Any new event at day k marks the corresponding value x_k as an outlier point. Popular methods in outlier detection involve Z-Score analysis (given that the sample is normally distributed), statistical modelling to estimate the sample distribution and its associated parameters, Linear Regression Models such as Principal Components Analysis, proximity-based non-parametric models, and Information Theory Models. Deliverable D4.4 on change, event and community detection techniques will examine these event detection techniques in detail, while in this deliverable we report on our data analysis on the collected social media posts.

In order to accumulate a large number of social media data that refer to a specific topic, i.e. floods in Italy in our study case, we have focused on a list of keywords (Table 6) to collect Twitter content. The selected terms focused mainly on flood events in Italy. This procedure lasted from April 01, 2017 until March 31, 2018, resulting in a wide collection of related tweets over a complete year. After one year of crawling tweets that concern flooding

[15] https://beaware-project.eu/

incidents in Italy, the collection counted 43,352 tweets. It was anticipated that part of this data would also include irrelevant posts, thus we proceeded with human annotation, e.g. users that tag tweets as relevant or not. This feedback also serves the development of automatic mechanisms to distinguish related posts, where ground-truth annotation is required for building robust machine learning algorithms that can automatically filter out irrelevant social media posts, as they are developed in WP3 of EOPEN as part of the social media crawling task. Figure 21 displays three time series regarding the number of the crawled tweets per each date of the year; the first refers to the total set, the second to a total of 16,749 annotated tweets by Italian experts and the third to the 4,701 tweets that were marked as relevant. The number of annotated posts is always larger than the relevant ones, which means that there are indeed many irrelevant items in the collection and, therefore, the necessity of a classification method is justified. However, solely by examining the uncharacterized data, it is evident that two important events were detected throughout the year: one on September 10, 2017 and one on November 04, 2017. In fact, it can be confirmed that they relate to the Livorno floods[16] and the anniversary of the 1966 flood of the Arno in Florence[17], respectively.

Table 6: List of terms used to collect relevant tweets

Keywords	English translation
alluvione	flood
alluvionevicenza	flood Vicenza
allagamento	flooding
bacchiglione	Bacchiglione
fiumepiena	full river
allertameteo	weather alert
sottopassoallagato	underpass flooded
alluvione2017	flood 2017
allertameteovicenza	weather alert Vicenza
esondazione	flooding

The content of the tweets was further analysed (e.g. removal of punctuation, URLs, and stop words) in order to discover the words that were most frequently used. The top ten non-location terms and the top ten mentioned locations are gathered in Table 7, together with their number of appearances and their English translation, when needed. Amongst the most repeated non-location words, there is only one term unrelated to floods (the music band Benji & Fede), while the most frequent locations are all places in Italy, including the country itself. These lists are also illustrated as word clouds in Figure 22.

[16] https://en.wikipedia.org/wiki/2017_Livorno_floods

[17] https://en.wikipedia.org/wiki/1966_flood_of_the_Arno

Table 7: Most frequently mentioned terms inside the collected tweets, separated in locations and non-locations

	Non-locations			Locations	
#	**Appearances**	**Word**	**English translation**	**Appearances**	**Word**
1	30820	alluvione	flood	9542	Livorno
2	5418	colpire	to hit	1869	Roma
3	5388	maltempo	bad weather	1532	Firenze
4	4915	vittima	victim	999	Italia
5	4774	allagare	to flood	968	Genova
6	4100	allertameteo	weather alert	932	Valtellina
7	3751	famiglia	family	878	Toscana
8	3484	pensiero	thought	563	Milano
9	3407	tenere	to hold	414	Sardegna
10	3360	benjiefede	Benji & Fede (band)	386	Parma

Using the top five non-location concepts and the top five locations, we have examined their frequency during the complete period of crawling. The time series of the number of appearances of each word are shown in Figure 23 and in Figure 24. The higher usage of the words “bad weather” and “victim” on September 10 compared to November 4 can be interpreted as the difference between an occurring flood and an anniversary. Furthermore, the increase on the appearances of “Livorno” and “Firenze” (Florence) on the same dates agrees with the afore-mentioned events.

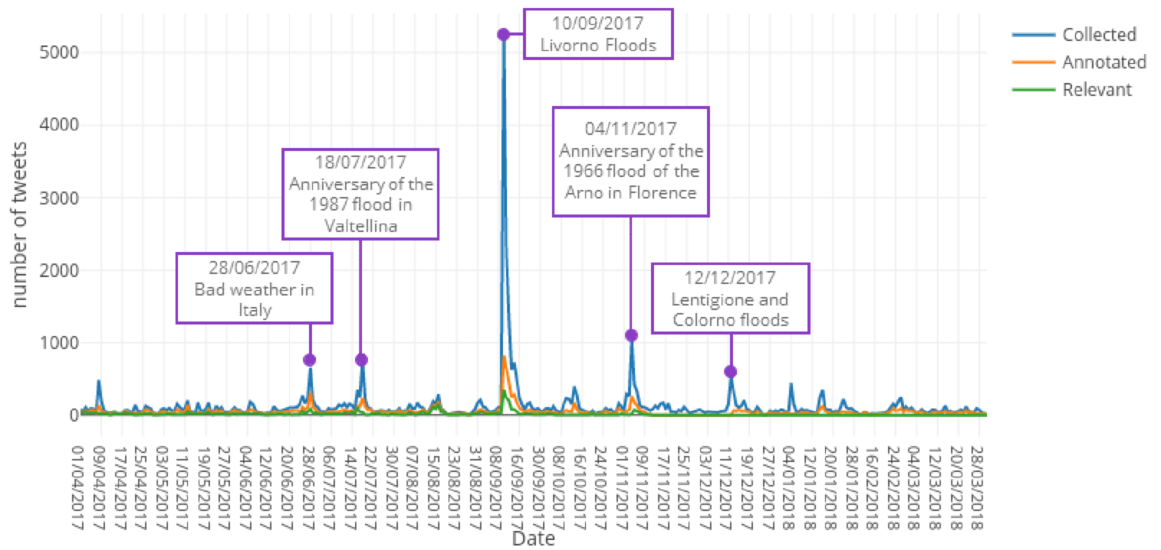

Figure 21. Fluctuation of the number of tweets during last year, grouped as collected, annotated and relevant

Figure 22. Top ten most recurrent words and locations, as word clouds

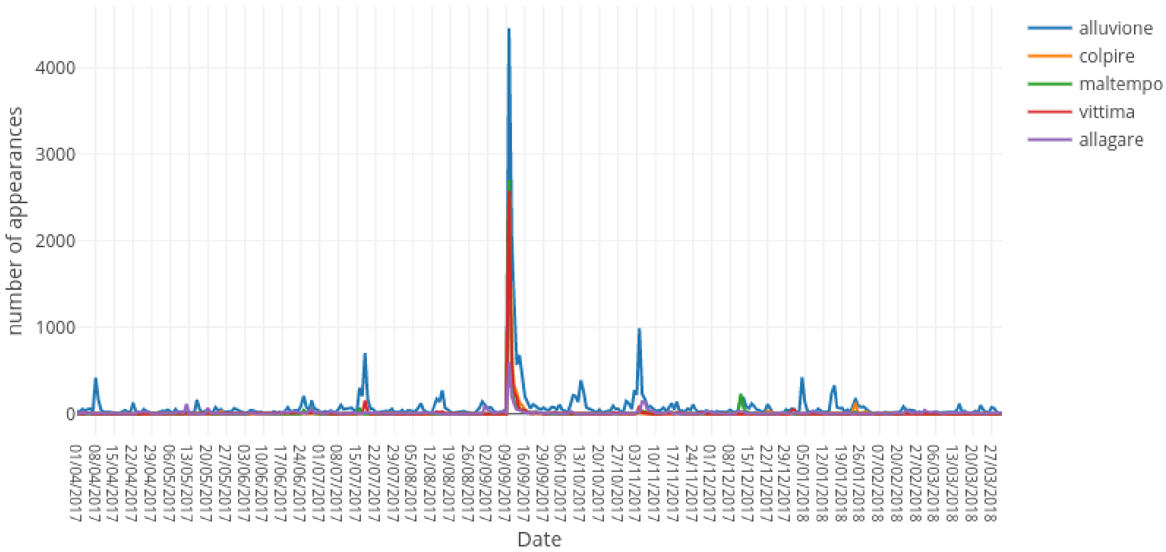

Figure 23. Appearances of the top five most used non-location words during last year

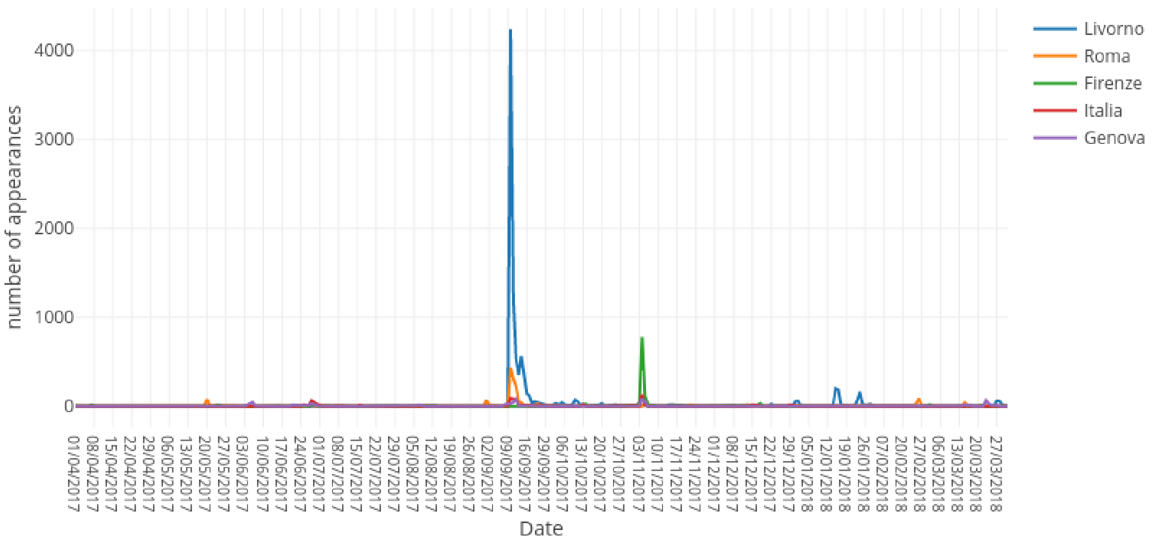

Figure 24. Appearances of the top five most used locations during last year

The results of our analysis indicate that specific flood events can also be detected from the analysis of social media streams using citizen observations. EOPEN is able to use meteorological data to forecast a potential (flood) event before it occurs, social media observations are able to detect this event during its occurrence and analysed satellite imagery can map the flooded areas to support the recovery procedures after the crisis event.

4 SIMILARITY FUSION FROM MULTIPLE SOURCES FOR INFORMATION RETRIEVAL

This section concerns the retrieval of similar observations/content found inside the EOPEN collection. The interesting element in the retrieval is that the content collected is described by several modalities (i.e. visual, textual, spatiotemporal, and concepts). Thus, there is the need to fuse these multimodal objects in a scalable way, taking into account memory and computational complexity, in order to retrieve similar content in response to a query. We present an overview of the existing approaches that realise multimodal fusion and retrieval.

4.1 State of the art in Similarity Fusion

Multimedia retrieval is a challenging problem due to the diversity and size of multimedia data combined with the difficulty of expressing desired queries. Over the years, many different approaches for multimedia retrieval have been introduced and compared. The main challenge seems to be the combination of multiple heterogeneous features (modalities) that can be extracted from collections of multimedia objects (e.g. low-level visual descriptors, high-level textual or visual features, etc.). The aforementioned combination process is known as multimodal fusion.

In some works, users can intervene and provide their feedback to the system. In (Xu, et al. 2015), the authors study the possibility of exploiting user-generated relevance feedback as a way to improve video similarity in video retrieval systems, using multiple modalities.

In general, there are three main strategies for multimodal fusion with respect to the level, at which fusion is accomplished. The first strategy is called early fusion and performs fusion at the feature level (e.g. (Caicedo, et al. 2012), (Magalhaes, et al. 2010)), where features from the considered modalities are combined into a common feature vector. The second strategy is the late fusion that fuses information at the decision level. This means that each modality is first learned separately and the individual results are aggregated into a final common decision (e.g. (Kitanovski, et al. 201), (Younessian, et al. 2012)). An advantage of early fusion inspired approaches (Atrey, et al. 2010) is the fact that it utilises the correlation between multiple features from different modalities at an early stage. However when the number of modalities increases, there is a decrease in their performance due to the fact that this makes it difficult to learn the cross-correlation among the heterogeneous features. On the other hand, late fusion is much more scalable (in terms of the modalities used) and flexible (as it enables the use of the most suitable methods for analysing each single modality) than early fusion, but fails on the other hand to utilize the feature-level correlation among modalities (Atrey, et al. 2010). Finally, the third strategy, called hybrid fusion, aims at exploiting the advantages of both early and late fusion strategies (e.g. (Yu, et al. 2014)).

Another interesting type of fusion is metric fusion (Wang, et al. 2013), an approach that attempts to fuse different “views” of the same modality, e.g. different types of low-level visual features for describing images.

Some multimedia and cross-modal retrieval studies have focused on specific methodologies.

An example is the well-known Latent Dirichlet Allocation (LDA). Wang, et al. (2014) proposed a supervised multimodal mutual topic reinforce modelling approach for cross-modal retrieval, called M3R. Another methodology is the Partial Least Squares (PLS) and a PLS-based framework was proposed by Siddiquie, et al. (2014) that maps queries from multiple modalities to points in a common linear subspace. Correlation matching is another methodology and in (Rasiwasia, et al. 2010), the authors utilized the correlation matching between the textual and visual modalities of multimedia documents in the task of cross-modal document retrieval.

With respect to graph-based methods and random-walk approaches (Ah-Pine, et al. 2015) present a unifying multimedia retrieval framework that incorporates two graph-based methods, namely cross-modal similarities and random-walk based scores. Specifically, the random-walk approach for multimodal fusion was introduced in (Hsu, et al. 2007), where the fusion of textual and visual information leads to improved performance in the video search task. The framework in (Ah-Pine, et al. 2015) includes as special cases all well-known fusion models (e.g. early, late, and diffusion-based) and does not require users' relevance feedback.

Finally, the most recent trend on multimedia retrieval problems is the use of deep learning. In this context, Feng et al. (Feng, et al. 2014) makes use of deep auto-encoders to learn features from different modalities in the task of cross-modal retrieval. Similarly, Wang et al. (2014) proposed a mapping mechanism for multimodal retrieval based on stacked auto-encoders. This mechanism learns one stacked auto-encoder for each modality in order to map the high-dimensional features into a common low-dimensional latent space. Finally, in (Wang, et al. 2015), a model based on Convolutional Neural Networks (CNN) that can be used for modality-specific feature learning was introduced.

4.2 Current EOPEN Framework in similarity search

The current EOPEN framework focuses exclusively on tweets and does not consider EO data. The key characteristics of tweets are that their text is rather short (the upper limit of allowed characters is 140), with non-standard terms, misspellings, "emojis", slang and abbreviations. Apart from the text, tweets may contain an image that is usually semantically related to the text. Therefore, the representation of a tweet involves three main modalities:

- Visual concepts
- Visual features
- Textual representation

During the first year of EOPEN, we have developed a one-modality search tool by considering solely the visual information (i.e. visual concepts or visual features). In both cases, the developed functionality is designed to be triggered by the user, by requesting information relevant to a text-image pair.

As far as visual concepts are concerned, they are extracted by using the concept detection framework presented in Section 3.1.2. Considering the concepts and the probabilities describing each image, it is possible to rank the images by any concept belonging to the concept pool. Fast indexing and retrieval is achieved by inserting in MongoDB at most the top 10 concepts per image.

Regarding the visual features, they are DCNN-based descriptors and they are the output of the last pooling layer of the fine-tuned GoogleNet architecture (Pittaras, 2017), described in Section 3.1.2. The dimension of the last pooling layer is 1024 and it is used as global image representation. The selection of a DCNN-based feature was based on the outcome of several studies that revealed the superiority of such features versus hand-crafted features both in terms of accuracy and time. Figure 25 depicts the layer of the GoogleNet architecture that is used as DCNN-feature.

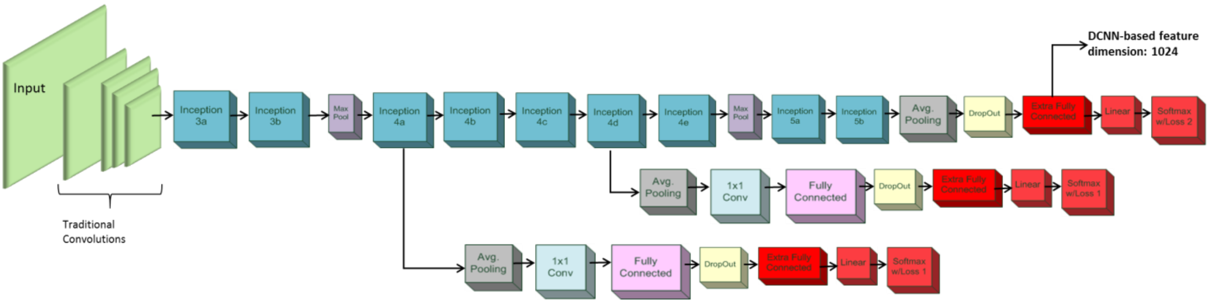

Figure 25. Extraction of DCNN-based feature from the fine-tuned GoogleNet.

For the Nearest Neighbour search, we create an Asymmetric Distance Computation index (Jegou, et al. 2011) and then, we compute the K-Nearest Neighbours from the query image.

Finally, a web service is implemented for accelerating the querying and retrieval process. This is achieved since in order to query the indexing structures, a two-step procedure is realized that involves: a) the loading of the index on the RAM memory, and b) the querying to the index. The first one of these steps is very time-consuming since it requires more than three minutes. Therefore, in order to eliminate the time required for the index loading, a web service is created that loads this indexing structures in RAM, and thus allows instant querying of the structure and eventually fast results retrieval each time a visual query is realized. We illustrate the information retrieval functionality in social media images using visual concepts in Figure 26.

Finally, as far as the textual representation of tweets is concerned after considering the traits of the tweets and based on several studies, we decided on avoiding classic text representation techniques such as BoW or Term Frequency Inversed Document Frequency (TFIDF) as the low number of features doesn't provide enough word co-occurrence. The most recent approaches in text classification that outperform other methods and perform well in cases of short text (when a significantly large corpus is available) are the word embeddings such as word2vec (Mikolov, et al. 2013) and GloVe (Pennington, et al. 2014). Word embeddings stand on the concept that similar words tend to occur together and will have a similar context (e.g. *football* and *basketball* tend to have a similar context around *sports*). It should be noted that both word2vec and GloVe are based on Deep Neural Networks (DNN). Eventually, the Euclidean distance is used for retrieving similar tweets to the query tweet.

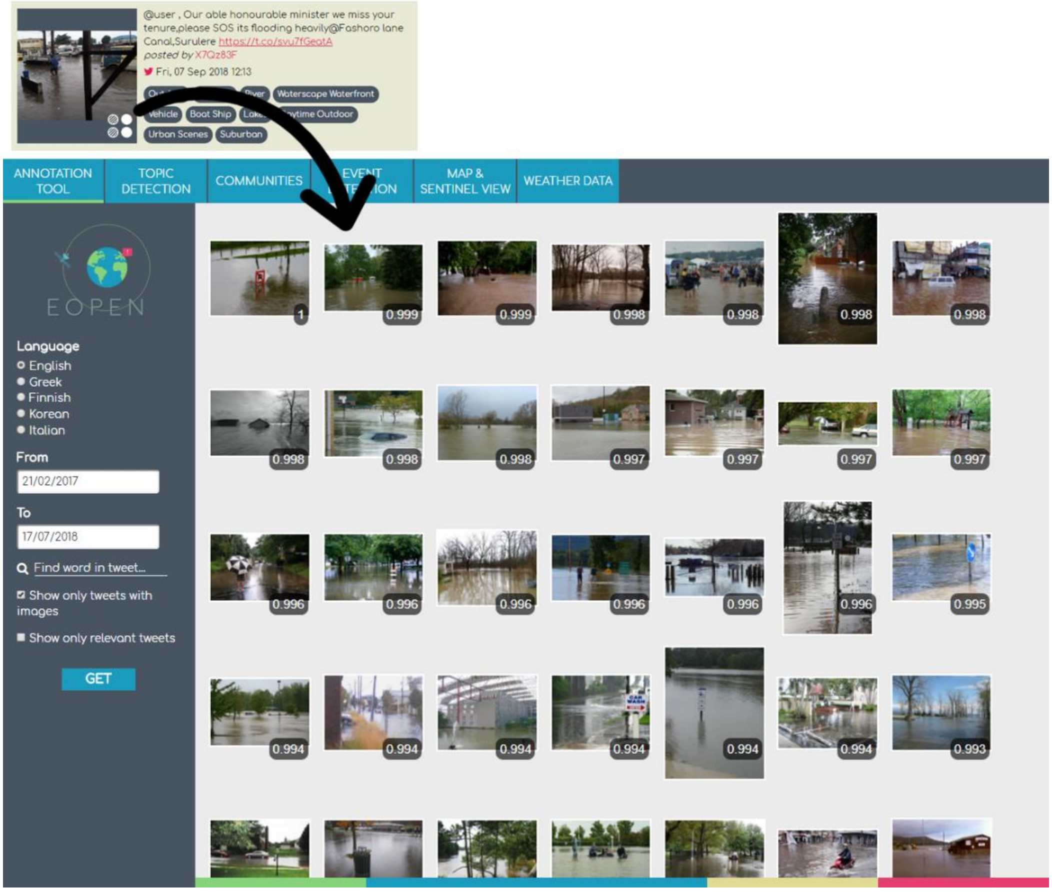

Figure 26. Information retrieval in social media images using visual concepts.

5 DATA CLUSTERING AND EXCHANGE AMONG FEDERATED DATABASES

This section involves the development of clustering algorithms that allow the co-clustering of both non-EO and EO data. As far as the non-EO data are concerned, we discriminate between visual and textual information. In this deliverable we provide an overview of the related work for both data types, and the framework that will be developed both for EO and non-EO data. Here, we focused primarily on enabling clustering of visual and textual information to cover both aspects of EO and non-EO data. The final framework including a co-clustering of both approaches, as well as the experiments on the performance of the algorithms selected for data clustering will be presented in the deliverable 4.2 (“Report on data clustering of EO and non-EO data”) that is due on M25.

5.1 Non-EO Data Clustering

5.1.1 Clustering visual information

Clustering is central to many data-driven application domains and has been studied extensively in terms of distance functions and grouping algorithms (Xie et al., 2016, June). A good clustering method shall have the following properties:

- High intra-class similarity: Cohesive within clusters
- Low inter-class similarity: Distinctive between clusters

To better meet the properties stated above, the algorithm autoencoder was developed. An autoencoder is a type of artificial neural network used to learn efficient data codings in an unsupervised manner (*Liou et al., 2008*). The standard structure of autoencoder is composed of three layers: input layer, hidden layer and output layer. Logically, an autoencoder is divided into two parts: encoder and decoder. The job of the encoder is to compress the input data to lower dimensional features, and the decoder’s job is on the other hand to reconstruct the high dimensional data from the compressed features. Thus, the input and output layer must have the same number of nodes. The similarity between the input data and output data are usually used to judge the performance of an autoencoder model, the higher similarity, the better. However, although autoencoder is good at extract key features from the initial high dimensional data, only utilization of autoencoder cannot meet the requirement for complex problems. Under these circumstances, the combination of autoencoder and other algorithms, for example k-means, are often adopted. It has been proven that the joint use of autoencoder and k-means can highly improve the performance of clustering on images.

Besides unsupervised clustering, the visualization is also a critical problem for high dimensional data. The method “t-distributed Stochastic Neighbour Embedding (t-SNE)” was developed to solve such a problem. t-SNE visualizes high-dimensional data by giving each data point a location in a two or three-dimensional map. The technique is a variation of Stochastic Neighbour Embedding (Hinton and Roweis, 2002), which is more straightforward to optimize, and produces significantly better visualizations by reducing the tendency to crowd points together in the centre of the map (Maaten, L. V. D. and Hinton, G., 2008). The t-SNE algorithm is mainly composed of two main stages. The first stage is to build a probability distribution over pairs of high-dimensional objects. With this distribution, it is highly probable that similar objects are picked, while with little possibility that dissimilar objects being picked. In the second stage, another probability distribution over the points in the low-dimensional map is defined, by which the Kullback–Leibler divergence (KLD) between the two distributions with respect to the locations of the points in the map is minimized. With t-SNE, it is possible to successfully visualize large real-world data sets with limited computational demands. It has been shown that t-SNE outperforms existing state-of-the-art techniques for visualizing a variety of real-world datasets.

Framework

As stated above, autoencoder and other machine learning algorithms will be adopted to do the unsupervised image clustering. To construct the neural network and other models, TensorFlow will be used.

TensorFlow is an open source software library for numerical computation using data flow graphs (Abadi et. al, 2016). In TensorFlow, mathematical operations are represented as the nodes in the graph, and the multidimensional data arrays (tensors) are shown as edges. Keras (TensorFlow backend) is utilized in the project, which is a high level API to build deep learning models. And there exist many advantages of employing Keras TensorFlow. One advantage is that it comes with solid support for machine learning and deep learning, and compared to other frameworks, the numerical computation core of TensorFlow is very flexible and powerful. Another advantage is that Keras TensorFlow, enabling distributed computing, is specialized for deep learning. As a result, users don't need to build the model from scratch, but to call the functionalities defined in TensorFlow.

5.1.2 **Clustering textual information**

Latent Dirichlet Allocation (LDA) (Blei and Jordan, 2003a) has been a model that performs very well when the number of topics is given as input, and has been extended to the hierarchical Dirichlet process (Teh et al., 2012) and DP-means (Kulis and Jordan, 2012), to overcome this issue. F-OPTICS, presented by (Schneider and Vlachos, 2013), has reduced the computational cost of the OPTICS algorithm, originally introduced in (Ankerst et al., 1999), using a probabilistic definition of the reachability distance, without significant accuracy reduction. The OPTICS-ξ algorithm (Ankerst et al., 1999) requires an extra parameter ξ, which has to be manually set in order to find "dents" in the OPTICS reachability plot. We shall use the original definition of the reachability distance (Ester et al., 1996), aiming to limit down the number of parameters. (Gan and Tao, 2015) introduces an approximate version of DBSCAN that could be combined with our proposed DBSCAN-Martingale approach by replacing DBSCAN with its corresponding approximate clustering algorithm. NQ-DBSCAN is a recent algorithm for density-based clustering, (Chen et al., 2018), which aims to reduce the number of unnecessary distance computations in order to provide a faster version of DBSCAN, without significant deviation from the exact output of the original DBSCAN algorithm. In NQ-DBSCAN the selection of the parameters ε and $minPts$) is considered optimal and a priori given. (Schubert et al., 2017) consider more challenging to estimate the density level ε rather than the more intuitive task to determine $minPts$. Recent sequential density-based clustering approaches involve several iterations to obtain the final cluster structure (Louhichi et al., 2014) (Mai et al., 2017).

In EOPEN, we cluster textual streams of data, in a service referred also as "topic detection", which is based on the combination of density-based clustering with LDA, as proposed in (Gialampoukidis et al., 2016d). The module estimates the number of clusters (topics) and the estimation is followed by Latent Dirichlet Allocation (Blei et al., 2003b), so as to assign social media posts to topics. Topic detection takes place either on the most recent (using a timeframe option) posts or on the ones that are provided by a search by keyword.

In Figure 27 we present the current illustration that we have designed for clustering textual data and presenting the groups of tweets using representative terms as the most frequent ones in the language of the corpus.

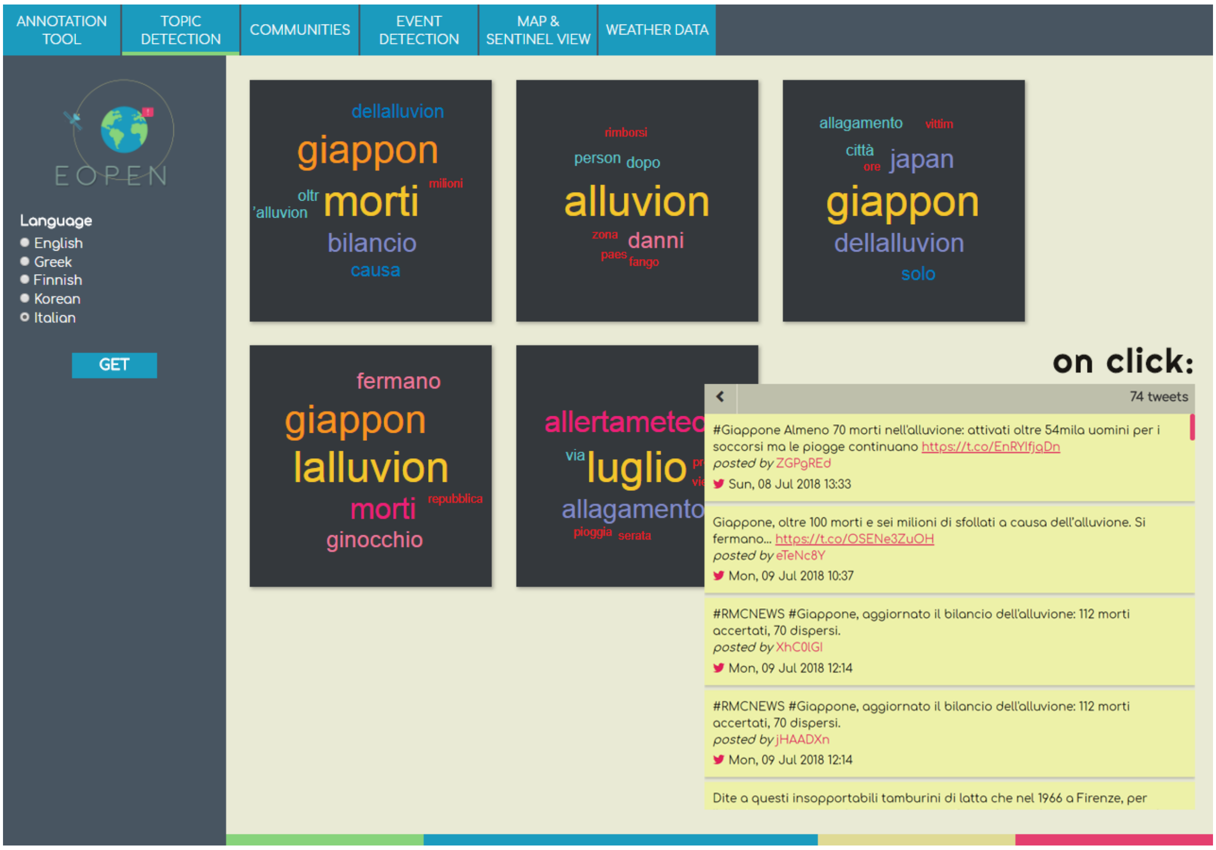

Figure 27. Clusters are presented as word clouds with the most frequent terms and each cluster contains a set of tweets.

5.2 EO Data Clustering

5.2.1 State of the art in clustering EO data

Unsupervised techniques to analyse vast amounts of EO data are nowadays crucial in order to perform complex analytics tasks such as detecting hidden patterns in an automatic manner. Applications of clustering for remote sensing data can be found throughout diverse domains such as agriculture (Abbas et al. 2016; Rahmah and Sitanggang 2016; Pascucci et al. 2018), hydrology (Cucchi et al. 2017; Qiao et al. 2018; Agarwal et al. 2016), earthquake analysis (Savaş et al. 2019; Harris and Anitha 2017), change detection (Leichtle et al. 2017), and anomaly detection (Liu et al. 2017a; Liu et al. 2017b), to name but a few. The most common reason to apply clustering on EO data is the task of image segmentation (Nguyen et al. 2018; Richards et al. 2010; Choubin et al. 2017; Pappas 1992). Image segmentation divides a given image into homogeneous parts, e.g. grouping pixels in an image together if they inhere the same colour; that way, roads or river streams are identifiable using unsupervised algorithms such as clustering (Wagstaff et al. 2001). Unsupervised image segmentation is thus often a pre-processing routine before classification is applied. Although clustering is able to identify patterns, it is not able to label those patterns, meaning clustering successfully detects a river stream, but the algorithm is unaware that the detected surface is a river. Clustering remote sensing data is, however, not limited to images. Further applications include grouping of topographical maps (Kumari et al. 2017) or geophysical data such as surface temperature distributions (Liu et al. 2018). However, the review of the state of the art in clustering visual information (cf. Section 5.1.1) revealed a great number of additional algorithms to be applied for the challenge of image segmentation and classification, which will be considered by EOPEN.

Reviewing the literature of clustering EO data, most applications rely on the well-known k-Means clustering algorithm. k-Means is a partitioning-based, prototype-based clustering algorithm, which divides a given dataset into k clusters, which are represented by their prototype. Normally, the centroid of the cluster is selected as the prototype. In each cluster, objects (e.g. images) are more similar to each other than to objects of other clusters. k-Means is quite versatile, and is therefore applied in diverse scenarios: grouping of text documents based on similarity (Huang 2008), novelty detection (Markou and Singh 2003), outlier detection (Chawla and Gionis 2013), dimension reduction (Ding and Li 2007), image segmentation (Ray and Turi 1999; Pappas 1992), and anomaly detection (Münz et al. 2007). k-Means is parameterized with an appropriate distance measure (e.g. Euclidean distance (Bora and Gupta)) and the number k of desired clusters to detect. The algorithm converges towards a local optimum (Ray and Turi 1999). However, it is unknown in advance, which k yields a global optimum, and thus the algorithm has to be executed multiple items for a given set of pre-defined values of k. Naturally, finding the "optimal" k is not trivial, and thus authors have further developed the algorithm to overcome its inherent limitations: ISODATA clustering (Verma et al. 2017; Abbas et al. 2016), GEO k-Means (Mato and Toulkeridis 2018), Fuzzy c-Means (Choubin et al. 2017), and functional k-Means (Pascucci et al. 2018). Besides k-Means, also other standard algorithms are applied to cluster EO data: hierarchical clustering (Huo et al. 2015), principal component analysis (PCA) (Shahdoosti and Ghassemian 2016), and spectral clustering (Murphy and Maggioni 2017). Spectral clustering performs the clustering in a lower dimension by first performing a dimension reduction using the Eigenvalues of the distance matrix. In (Liu et al. 2017b), spectral clustering is successfully applied to detect anomalistic data as a pre-processing step to improve data quality and remove noisy data.

5.2.2 **Framework**

The module will be implemented as a service, and will offer to users common clustering algorithms such as the mentioned k-Means algorithm, bisecting k-Means (Savaresi and Boley 2001), and Latent Dirichlet Allocation. The objective is that these algorithms can then universally be applied to a diverse set of problems. Since the optimal number of clusters is often unknown in advance, the module will automatically perform multiple iterations of concerned clustering algorithms, evaluate the results based on well-established clustering criteria such as the Silhouette coefficient (Rousseeuw 1987), and finally select the optimal number of clusters. In comparison to other tasks in this work package, the focus is rather on porting existing algorithms to the Big Data infrastructure than to improve existing approaches found in the literature. In this context, the distributed data analytics framework Apache Spark[18] shall be used. Furthermore, the framework developed for clustering of visual content is directly applicable to cluster EO-related imagery as well.

[18] https://spark.apache.org/

6 COMMUNITY DETECTION IN SOCIAL MEDIA

EOPEN will detect end-user communities through their interaction with the EOPEN platform. Any two user accounts who download the same product will be linked, due to their interest in the same area. Moreover, links among users appear in the context of Social Media, where one user follows or mentions another user, leading to a complex network of user-to-user interactions that will be visualised in EOPEN. To detect such online and offline communities, EOPEN will be based on density-based approaches (Gialampoukidis et al., 2016b) to detect communities in complex networks.

6.1 State of the art in Community Detection methods

A large number of community detection algorithms has appeared in the literature, e.g. (Fortunato, 2010), (Malliaros and Vazirgiannis, 2013), (Harenberg et al., 2014), but only few of them are directly applicable to large scale complex networks, such as social media graphs, as reviewed in (Papadopoulos et al., 2012). The Girvan-Newman community detection algorithm (Girvan and Newman, 2002) (Newman and Girvan, 2004) is a divisive hierarchical process based on edge betweenness, a centrality measure that can be quickly calculated (Brandes, 2001). An alternative hierarchical approach for community detection uses the modularity measure as an objective function to optimize (Clauset et al., 2004). Initially, all vertices are separate communities and any two communities are merged if the modularity increases.

The Label Propagation method (Raghavan et al., 2007) initializes every node with a unique label and at each step every node adopts the label that most of its neighbours currently have. Hence, an iterative process is defined, in which densely connected groups of nodes form a consensus on a label and communities are extracted. The Louvain method (Blondel et al., 2008) is also based on the maximisation of modularity and involves two phases that are repeated iteratively. The Walktrap method (Pons and Latapy, 2006) generates random short walks on the graph by simulating transitions between nodes. Since short random walks tend to stay within the same community, it is possible to detect communities using such random walks.

The Infomap method (Rosvall and Bergstrom, 2008; Rosvall et al., 2009; Bohlin et al., 2014) is an information-theoretic approach to community detection which considers the problem of finding a community structure in networks equivalent to a coding problem where the goal is to minimise the Shannon information required for the transmission of a message, by giving a unique name (codeword) to every node in the network.

The graph-analogue of DBSCAN (Ester et al., 1996) is called DBSCAN* (Campello et al., 2013) and defines core objects in a graph in a way similar to the core points of DBSCAN. The transition though from density-based clustering of spatial datasets to community detection in graphs through DBSCAN* does not involve border points, due to an updated definition of reachability (see (Campello et al., 2013) for further details). Both DBSCAN and DBSCAN* are based on two parameters: the desired density level ε and a lower bound for the number of points in a cluster $minPts$. The estimation of the density level is not a trivial task and several approaches that extend DBSCAN have been proposed for extracting clusters without determining a priori the parameter ε, including a DBSCAN-Martingale approach

(Gialampoukidis et al., 2016d) that introduces a variable density level in the clustering algorithm based on a Martingale process. This motivates us to investigate the application of Martingales to DBSCAN* and in particular for introducing an adjustable $minPts$ parameter that plays an important role in community detection.

6.2 EOPEN community detection module

The proposed module offers two novel functionalities, which are based on the key-player identification method of (Gialampoukidis et al., 2016c) and the key-community detection of (Gialampoukidis et al., 2017a). We adopt the Louvain algorithm (Blondel et al., 2008) for key-community detection, as a fast and scalable approach that admits hierarchical and iterative methods to maximise modularity:

$$Q = \frac{1}{2m}\sum_{i=1}^{c}(e_{ii} - a_i^2)$$

where e_{ij} is the fraction of links between a node in community i and a node in community j, a_i is the fraction of links between two members of the community $i, m = \sum_k deg(n_k)$.

On the detected community of nodes, we compute the degree measure per node k, defined as the number of neighbours of the node k. The nodes with the largest degree are known as hubs or "key-players" in the social network interactions.

In Figure 28 communities are visualised as connected graphs of different colour and "key-players" are presented in a horizontal list (at the bottom).

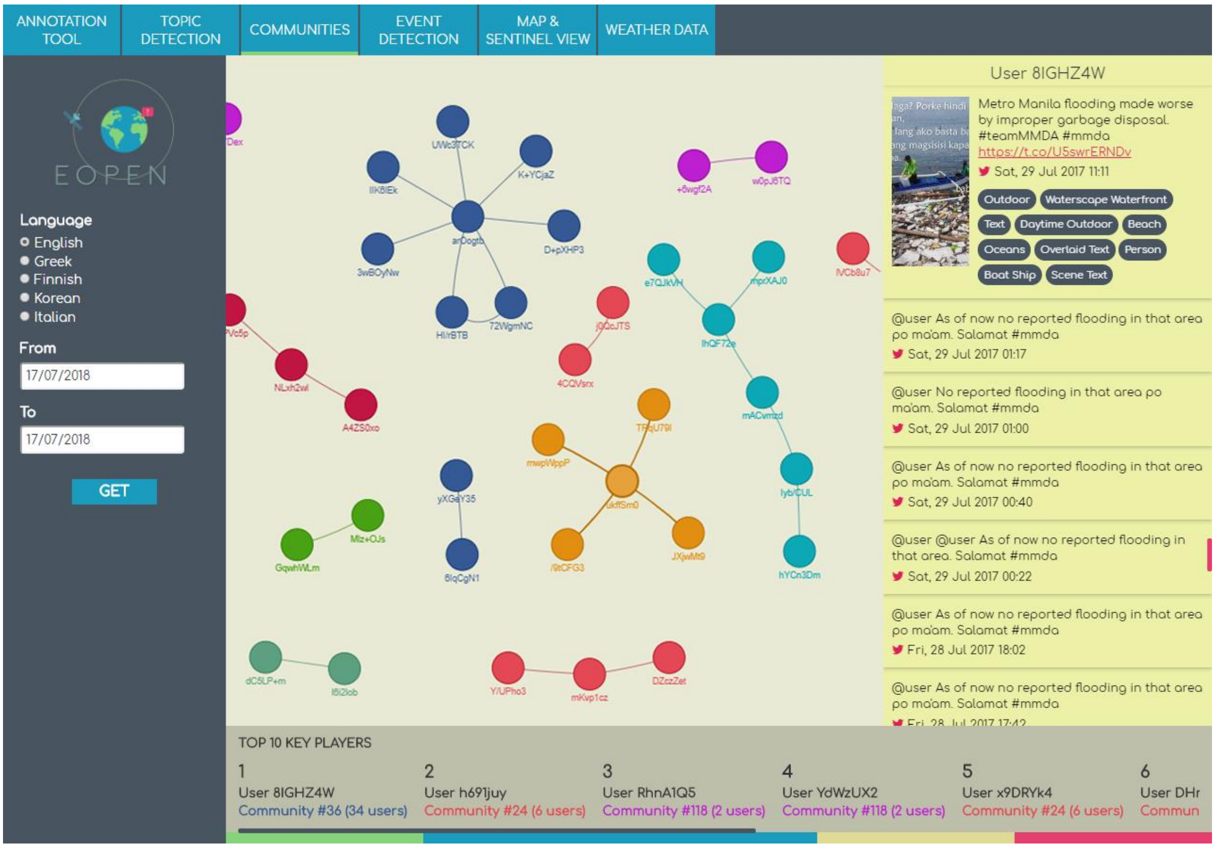

Figure 28. Communities and key-players in social media data collected for EOPEN

7 CONCLUSIONS

In this deliverable, we have presented the first version of change detection techniques. In this context we have proposed baseline approaches for flood monitoring by exploiting both optical and SAR (radar) satellite data. Having two different types of data sources provides us with better flexibility on the generation of the water-bodies masks. SAR satellites are able to provide a constant flow of images, regardless the weather condition, whilst optical data provide more human friendly images. In case of SAR images an automatic thresholding method was applied that processes Sentinel-1 radar data. For radar satellite data an automatic thresholding approach is tested, while for the optical data a modified Modified Normalized Difference Water Index (MNDWI) and a discriminant analysis with post-processing is applied. All approaches showed acceptable performance and are considered as baseline for the further development of the flood monitoring module. Regarding the road passability scenario, we have fine-tuned several pre-trained DCNNs including ResNet-101, VGG-19 and run several experiments in order to tune the learning and batch parameters. The results showed that lower values of learning rate and higher values of batch size improve generally the performance of the network. Finally, regarding the food security scenario visual interpretation of the results showed a rather effective outlier detection method. The framework was designed in a way to respect purpose and needs of the change detection product. The updated land cover map would function as a training dataset that needs to be both clean enough from wrongly labelled entities, but also random and inclusive of heterogeneous rice samples to avoid overfitting of the supervised classifiers. Detailed evaluation would be possible at the second stage (T7.1 rice paddy mapping), where comparisons between classification accuracies using both the 2015 dataset and the updated one can take place.

Regarding change detection, future work will be reported in D4.4 “Change, event and community detection techniques” (M33) for the flood monitoring and road passability scenarios and in D7.1 and D7.2 for food security scenario. Specifically, regarding flood monitoring D4.4 will include training of sophisticated DCNNs based on a time-series of fused images, generated by our baseline algorithms and will be used to predict more accurate water-bodies masks. As far as road passability is concerned, we are planning on evaluating the same DCNNs but have a more fine-grained annotation by splitting the images in batches and also consider the problem a two-stage problem, where there are two networks; one for addressing the issue of road existence and the second for checking whether the existing road is passable or not. Regarding food security, the presented change detection product will be used for the training of supervised classification of rice (Random Forest, Recurrent Neural Networks) to produce accurate rice maps. Then rice yield models will be developed based on the estimated area of rice to estimate the production in kg/ha and compare it with past and expected production values.

Regarding concept detection, a fine-tuned pre-trained GoogleNet was trained on the 345 TRECVID labels using the ImageNet “fall” 2011 dataset and the results were very satisfactory. As future work, we will evaluate other fine-tuning techniques and the results will be presented in D4.4.

Regarding event detection, a preliminary analysis has been carried out on collected social media posts over a one-year period in Italy, which refers to floods. Outliers in the quantity of posts per each day have been interpreted as events and have been successfully linked to real flooding cases, thus proving a correlation. Future plans include a review of state-of-the-art methods in the field of event detection, a real-time implementation of the respective module and the integration of weather data, complementary to the social media data. The aforementioned work will be documented in D4.4.

Concerning the similarity fusion module, an overview of the techniques existing in the bibliography was presented. Furthermore, the modalities describing the non-EO data (i.e. tweets) were described and the approaches for single modality similarity retrieval were discussed. In D.4.4, we plan on fusing all the modalities under a framework that will consider the computational and memory complexity given the large size of data.

Clustering has also been investigated for the case of textual information. After a review of state-of-the-art algorithms, a combination of density-based clustering with LDA has been proposed. A service has been implemented to estimate the number of clusters (also referred to as topics), group all tweets into clusters and find the most frequent words per cluster. Furthermore, a proof-of-concept was developed to enable LDA clustering on the HPDA infrastructure at HLRS. With respect to clustering of EO data, we presented a proof-of-concept for unsupervised image clustering based on TensorFlow. Future work will thus concern both the integration of the service in the EOPEN platform, as well as a stronger exploitation of the HPC and HPDA infrastructure at HLRS for larger datasets and faster outcomes. The implemented solution and results will be documented in deliverable 4.2 "Report on data clustering of EO and non-EO data" (M25).

As far as it concerns the community detection, the Louvain algorithm has been adopted for examining social media relationships and the degree measure has been utilized to find key-players in the communities. In the future, an advanced method to detect key-players will be investigated and other connections beyond social media will be taken into consideration, such as interaction within the EOPEN platform.

Finally, the deliverable includes a comparison of the change detection module developed within EOPEN against the one that has been developed within another European project, i.e. CANDELA, and also a description of EOPEN's connection to the H2020 beAWARE project, since they share common partners. The aforementioned information is found in the Appendix section.

8 REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M. 2016. "Tensorflow: a system for large-scale machine learning", Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation (pp. 265-283).

Abbas, A. W., Minhallh, N., Ahmad, N., Abid, S.A.R. and Khan, M.A.A. 2016. *"K-Means and ISODATA Clustering Algorithms for Landcover Classification Using Remote Sensing"*, Sindh University Research Journal, vol. 48 (2).

Agarwal, A., Maheswaran, R., Sehgal, V., Khosa, R., Sivakumar, B. and Bernhofer, C. 2016. *"Hydrologic regionalization using wavelet-based multiscale entropy method"*, In Journal of Hydrology, vol. 538, p. 22–32.

Ah-Pine, J., Csurka, G. and Clinchant, S. 2015. *"Unsupervised visual and textual information fusion in cbmir using graph-based methods"*, ACM Transactions Information Systems (TOIS) vol. 33(2), p. 9.

Andreadis, S., Gialampoukidis, I., Fiorin R., Lombardo, F., Norbiato, D., Karakostas, A., Ferri M., Vrochidis, S. and Kompatsiaris, I., 2018. *"Social Media Observations for Flood Event Monitoring in Italy over a One-Year Period"*, 2nd International Conference Citizen Observatories for natural hazards and Water Management (COWM 2018), Venezia, Italy.

Ankerst, M., Breunig, M.M., Kriegel, H.P. and Sander, J. 1999. "*Optics: ordering points to identify the clustering structure*", ACM Sigmod Record, ACM, pp. 49–60.

Antoniou, I., Gialampoukidis, I. and Ioannidis, E. 2015. "*Age and time operator of evolutionary processes*", International Symposium on Quantum Interaction, Springer, p. 51–75.

Atrey, P.K., Hossain, M.A., El Saddik, A. and Kankanhalli, M.S. 2010. *"Multimodal fusion for multimedia analysis: a survey"*, Multimedia Systems, vol. 16(6), p. 345–379.

Babu, R.S, Radhakrishnan, B. and PadmaSuresh, L. 2016. *"Detection and extraction of roads from satelliteimages based on laplacian of gaussian operator"*, International Conference on Emerging Technological Trends (ICETT), p. 1–7. IEEE

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., 2008. *"Speeded-Up Robust Features (SURF)"*, Computer Vision and Image Understanding, vol. 110(3), p.346–359.

Blei, D.M. and Jordan, M.I., 2003a. "*Modeling annotated data*", Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, ACM. pp. 127–134.

Blei, D.M., Ng, A.Y. and Jordan, M.I. 2003b. "*Latent dirichlet allocation*", Journal of Machine Learning Research 3(Jan), p. 993–1022.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. 2008. "*Fast unfolding of communities in large networks*", Journal of Statistical Mechanics: Theory and Experiment, vol. 2008 (10), p. 10008.

Bohlin, L., Edler, D., Lancichinetti, A. and Rosvall, M. 2014. "*Community detection and visualization of networks with the map equation framework*", Measuring Scholarly Impact, Springer, pp. 3-34.

Bora, D.J. and Gupta, A. K. 2014. *"Effect of Different Distance Measures on the Performance of K-Means Algorithm: An Experimental Study in Matlab"*, Available online at http://arxiv.org/pdf/1405.7471v1.

Brandes, U. 2001. "*A faster algorithm for betweenness centrality*", Journal of Mathematical Sociology, vol. 25 (2), p. 163-177.

Bruns, A. and Burgess J. 2014. *"Crisis communication in natural disasters: The Queensland floods and Christchurch earthquakes"*, Twitter and society, Peter Lang, p. 373–384.

Caicedo, J.C., BenAbdallah, J., Gonzalez, F.A. and Nasraoui, O. 2012. *"Multimodal representation, indexing, automated annotation and retrieval of image collections via non-negative matrix factorization"*, Neurocomputing, vol. 76(1), p. 50–60.

Campello, R. J., Moulavi, D. and Sander, J. 2013. "*Density-based clustering based on hierarchical density estimates*", Advances in Knowledge Discovery and Data Mining, Proceedings of the 17th Pacific-Asia Conference (PAKDD 2013), Part II. Springer, 2013, p. 160-172.

Chawla, S. and Gionis, A. 2013. *"k -means–: A unified approach to clustering and outlier detection",* In Joydeep Gosh (Ed.): Proceedings of the 2013 SIAM International Conference on Data Mining. Proceedings of the 2013 SIAM International Conference on Data Mining. [Philadelphia, PA]: SIAM, Society for Industrial and Applied Mathematics, pp. 189–197.

Chen, Y., Tang, S., Bouguila, N., Wang, C., Du, J. and Li, H., 2018. "*A fast clustering algorithm based on pruning unnecessary distance computations in dbscan for high-dimensional data*", Pattern Recognition, vol. 83, p. 375-387.

Cheong, F. and Cheong C. 2011. *"Social Media Data Mining: A Social Network Analysis Of Tweets During The 2010-2011 Australian Floods"*, PACIS, vol. 11, p. 46–46.

Choubin, B., Solaimani, K., Habibnejad R., M. and Malekian, A. 2017. *"Watershed classification by remote sensing indices: A fuzzy c-means clustering approach"*, Journal of Mountain Science 14 (10), pp. 2053–2063. DOI: 10.1007/s11629-017-4357-4.

Clauset, A., Newman, M. E. J. and Moore, C. 2004. "*Finding community structure in very large networks*", Physical Review E, vol. 70 (6), p. 066111.

Cover, T. M. and Thomas, J. A. 2012. *"Elements of information theory"*, John Wiley & Sons.

Cucchi, K., Kawa, N., Hesse, F. and Rubin, Y. 2017. "Clustering and Bayesian hierarchical modeling for the definition of informative prior distributions in hydrogeology. In AGU Fall Meeting Abstracts.

Dias, D. and Dias, U, 2018. "*Flood detection from social multimedia and satellite images using ensemble and transfer learning with CNN architectures*", Working Notes Proceedings of the MediaEval 2018 Workshop, Sophia Antipolis, France, 29-31 October 2018.

Ding, C. and Li, T. 2007. *"Adaptive dimension reduction using discriminant analysis and k-means clustering",* Proceedings of the 24th international conference on Machine learning. ACM, pp. 521–528.

Dockhorn, A., Braune, C. and Kruse, R., 2016. "*Variable density based clustering*", , 2016 IEEE Symposium Series on Computational Intelligence (SSCI),IEEE. p. 1–8.

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W. and Li, X. 2016. *"Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band"*, Remote Sensing, vol. 8(4), p. 354.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. 1996. "*A density-based algorithm for discovering clusters in large spatial databases with noise*", Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD 1996). AAAI Press, pp. 226.

Feng, F., Wang, X., Li, R. 2014. *"Cross-modal retrieval with correspondence autoencoder"*, In Proceedings of the 22nd ACM international conference on multimedia. ACM, pp 7–16

Fortunato, S. 2010. "*Community detection in graphs*", Physics Reports, vol. 486 (3), p. 75-174.

Freeman, L. C. 1978. "*Centrality in social networks conceptual clarification*", Social networks, vol. 1 (3), p. 215–239.

Gan, J., and Tao, Y. 2015. "*Dbscan revisited: mis-claim, un-fixability, and approximation*", Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, ACM, p. 519–530.

Gialampoukidis, I., Gustafson, K., and Antoniou, I. 2013. "*Financial time operator for random walk markets*", Chaos, Solitons & Fractals 57, p. 62–72.

Gialampoukidis, I., Kalpakis, G., Tsikrika, T., Papadopoulos, S., Vrochidis, S., and Kompatsiaris, I. 2017a. "*Detection of terrorism-related twitter communities using centrality scores*", 2nd International Workshop on Multimedia Forensics and Security (MFSec2017), ACM.

Gialampoukidis, I., Kalpakis, G., Tsikrika, T., Vrochidis, S., and Kompatsiaris, I. 2016c. "*Key player identification in terrorism-related social media networks using centrality measures*", European Intelligence and Security Informatics Conference (EISIC 2016), August, pp. 17–19.

Gialampoukidis, I., Liparas, D., Vrochidis S., and Kompatsiaris, I. 2016e. "*Query-based topic detection using concepts and named entities*", Proceedings of the 1st International Workshop on Multimodal Media Data Analytics (MMDA 2016), pp. 9–13.

Gialampoukidis, I., Moumtzidou, A., Tsikrika, T., Vrochidis, S., & Kompatsiaris, I. 2016a. *"Retrieval of multimedia objects by fusing multiple modalities"*, In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval (pp. 359-362). ACM.

Gialampoukidis, I., Tsikrika, T., Vrochidis, S., and Kompatsiaris, I. 2016b. "*Community detection in complex networks based on DBSCAN* and a Martingale process*", Semantic and Social Media Adaptation and Personalization (SMAP), 2016 11th International Workshop on (pp. 1-6), IEEE.

Gialampoukidis, I., Vrochidis, S., and Kompatsiaris, I. 2016d. "*A hybrid framework for news clustering based on the DBSCAN-Martingale and LDA*", Machine Learning and Data Mining in Pattern Recognition, Springer International Publishing, pp. 170-184.

Gialampoukidis, I., Vrochidis, S., Kompatsiaris, I., and Antoniou, I. 2017b. "*Topic detection using the DBSCAN-Martingale and the time operator*", 17th Conference of the Applied Stochastic Models and Data Analysis (ASMDA2017).

Girvan, M. and Newman, M. E. J. 2002. "*Community structure in social and biological networks*", Proceedings of the National Academy of Sciences, vol. 99 (12), p. 7821-7826.

Harenberg, S., Bello, G., Gjeltema, L., Ranshous, S., Harlalka, J., Seay, R., Padmanabhan, K., and Samatova, N. 2014. "*Community detection in large-scale networks: a survey and empirical evaluation*", Wiley Interdisciplinary Reviews: Computational Statistics, vol. 6 (6), p. 426-439.

Harris, P., Anitha, J. 2017. *"Post earthquake disaster awareness to emergency task force using crowdsourced data",* 2017 IEEE International Conference on Industrial and Information Systems (ICIIS) (pp. 1-6). IEEE.

He, K., Zhang, X., Ren, S., Sun, J. 2016. *"Deep residual learning for image recognition"*, In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Henry, H., Azimi, S.M. and Merkle, N. 2018. *"Road segmentation in SAR satellite images with deep fully-convolutional neural networks"*, arXiv preprint arXiv:1802.01445,.

Hinton, G.E. and Roweis, S.T. 2002. *"Stochastic Neighbor Embedding"*, Advances in Neural Information Processing Systems, vol. 15, p. 833–840, Cambridge, MA, USA, 2002. The MIT Press.

Hsu, W.H., Kennedy, L.S. and Chang, S.F. 2007. *"Video search reranking through random walk over document level context graph"*, In Proceedings of the 15th international conference on multimedia (pp. 971–980). ACM.

Huang, A. 2008. "*Similarity measures for text document clustering*", Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand, pp. 49–56.

Huo, L.-Z., Tang, P., Zhang, Z. and Tuia, D. 2015. *"Semisupervised classification of remote sensing images with hierarchical spatial similarity",* IEEE Geoscience and Remote Sensing Letters 12 (1), pp. 150–154.

Jegou, H., Douze, M. and Schmid, C. 2011. "*Product Quantization for Nearest Neighbor Search*", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 33(1), p. 117-128.

Jégou, H., Douze, M., Schmid, C. and Pérez, P., 2010. *"Aggregating local descriptors into a compact image representation"*, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 3304-3311). IEEE. San Francisco, CA.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S. and Darrell, T., 2014. *"Caffe: Convolutional architecture for fast feature embedding",* In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675-678). ACM.

Kang, W., Xiang, Y., Wang, F., Wan, L. and You, H. 2018. *"Flood Detection in Gaofen-3 SAR Images via Fully Convolutional Networks"*, Sensors, 18(9), p. 2915.

Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., and Moradi, A. 2012. *"An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia",* Environmental Earth Sciences, 67(1), 251-264.

Kitanovski, I., Trojacanec, K., Dimitrovski, I., Loskovska, S. 2013. *"Multimodal medical image retrieval",* In ICT Innovations 2012. Springer, p. 81–89 16.

Kongthon, A., Haruechaiyasak, C., Pailai, J. and Kongyoung, S., 2014. *"The role of Twitter during a natural disaster: Case study of 2011 Thai Flood"*, Technology Management for Emerging Technologies (PICMET), 2012 Proceedings of PICMET'12: IEEE, pp. 2227–2232.

Krizhevsky, A., Sutskever, I. and Hinton, G. E., 2012. *"Imagenet classification with deep convolutional neural networks"*, In Advances in neural information processing systems (pp. 1097–1105).

Kulis, B., and Jordan, M.I. 2012. "*Revisiting k-means: New algorithms via Bayesian nonparametrics*", Proceedings of the 29th International Conference on Machine Learning (ICML-12), pp. 513–520.

Kumari, M., Singh, C. K. and Basistha, A. 2017. *"Clustering Data and Incorporating Topographical Variables for Improving Spatial Interpolation of Rainfall in Mountainous Region"*, Water Resources Management, vol. 31 (1), pp. 425–442.

Lan, Z.Z., Bao, L., Yu, S.I., Liu, W. and Hauptmann, A.G. 2014. *"Multimedia classification and event detection using double fusion"*, Multimedia Tools Applications, 71(1):333–347.

Leichtle, T., Geiß, C., Wurm, M., Lakes, T. and Taubenböck, H. 2017. *"Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment",* In International Journal of Applied Earth Observation and Geoinformation 54, pp. 15–27. DOI: 10.1016/j.jag.2016.08.010.

Liou, C.-Y., Huang, J.-C. and Yang, W.-C. 2008. *"Modeling word perception using the Elman network"*, Neurocomputing, vol. 71(16–18): 3150.

Liu, H., Zhan, Q., Yang, C. and Wang, J. 2018. *"Characterizing the Spatio-Temporal Pattern of Land Surface Temperature through Time Series Clustering: Based on the Latent Pattern and Morphology",* Remote Sensing 10 (4), p. 654. DOI: 10.3390/rs10040654.

Liu, Q., Klucik, R., Chen, C., Grant, G., Gallaher, D., Lv, Q. and Shang, L. 2017a. *"Unsupervised detection of contextual anomaly in remotely sensed data",* Remote Sensing of Environment, vol. 202, p. 75–87.

Liu, Q., Zhang, Y., Lv, Q. and Shang, L. 2017b. *"Data quality screening for high-resolution satellite imagery via spectral clustering",* 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 4742-4745). IEEE.

Louhichi, S., Gzara, M., and Abdallah, H.B. 2014. "*A density based algorithm for discovering clusters with varied density*", Computer Applications and Information Systems (WCCAIS), 2014 World Congress on, IEEE. pp. 1–6.

Lowe, D. G., 2004. *"Distinctive image features from scale-invariant keypoints"*, International Journal of Computer Vision, 60, 91–110.

Maaten, L. V. D. and Hinton, G. 2008. "*Visualizing data using t-SNE*", Journal of machine learning research, vol.9 , p. 2579-2605.

Magalhaes, J. and Rüger, S. 2010. *“An information-theoretic framework for semantic-multimedia retrieval”*, ACM Trans Inf Syst (TOIS) 28(4):19

Mai, S.T., Dieu, M.S., Assent, I., Jacobsen, J., Kristensen, J., and Birk, M. 2017. “*Scalable and interactive graph clustering algorithm on multicore cpus*”, 2017 IEEE 33rd International Conference on Data Engineering (ICDE) (pp. 349–360). IEEE.

Malliaros, F. D. and Vazirgiannis, M. 2013. “*Clustering and community detection in directed networks: A survey*”, Physics Reports, vol. 533 (4), p. 95-142.

Markatopoulou, F., Mezaris, V., & Patras, I., 2015. *“Cascade of classifiers based on binary, non-binary and deep convolutional network descriptors for video concept detection”*, In Proc. of the IEEE International Conference on Image Processing 2015 (pp. 1786–1790).

Markatopoulou, F., Moumtzidou, A., Galanopoulos, D., Mironidis, T., Kaltsa, V., Ioannidou, A., Symeonidis, S., Avgerinakis, K., Andreadis, S., Gialampoukidis, I., Vrochidis, S., Briassouli, A., Mezaris, V., Kompatsiaris, I. and Patras, I. 2016. *“ITI-CERTH participation in TRECVID 2016”*, Proc. TRECVID 2016 Workshop, Gaithersburg, MD, USA

Markou, M. and Singh, S. 2003. *“Novelty detection: a review—part 1: statistical approaches”*, Signal Processing 83 (12), pp. 2481–2497.

Martimort, P., Berger, M., Carnicero, B., Del Bello, U., Fernandez, V., Gascon, F., Silvestrin, P., Spoto, F., Sy , O., Arino, O., Biasutti, R. and Greco, B. 2007, *“The optical high-resolution mission for GMES operational services”,* 2007 IEEE International Geoscience and Remote Sensing Symposium (pp. 2677-2680). IEEE.

Michail, E., Moumtzidou, A., Gialampoukidis, I., Avgerinakis, K., Scarpino, M. G., Vrochidis, S., Vingione, G., Kompatsiaris, I., Labbassi, K., Menenti, M., Elghandour, F.-E. 2018. *“Testing a Flood Mask Correction Method Of Optical Satellite Imagery Over Irrigated Agricultural Areas”*, 2nd Mapping Water Bodies from Space Conference (MWBS2018), ESA-ESRIN Frascati, Italy.

Mikolov, T., et al. 2013. *“Distributed representations of words and phrases and their compositionality”*, Advances in neural information processing systems.

Morozova, M., Elizarova, T., & Pleteneva, T. 2013. *“Discriminant analysis and Mahalanobis distance (NIR diffuse reflectance spectra) in the assessment of drug’s batch-to-batch dispersion and quality threshold establishment”*, European Scientific Journal, ESJ, vol. 9(27).

Moumtzidou, A. Bakratsas, M., Andreadis, S., Gialampoukidis, I., Vrochidis S., and Kompatsiaris, I. *“Road passability estimation using deep neural networks and satellite image patches”*, Conference On Big Data From Space (BiDS’19), February 2019, Munich, Germany (accepted for publication)

Moumtzidou, A., Andreadis, S., Gialampoukidis, I., Karakostas, A., Vrochidis, S. and Kompatsiaris, I., 2018. *“Flood Relevance Estimation from Visual and Textual Content in Social Media Streams”*, Companion of the The Web Conference 2018 on The Web Conference 2018, International World Wide Web Conferences Steering Committee (pp. 1621–1627).

Münz, G., Li, S. and Carle, G. 2007. *“Traffic anomaly detection using k-means clustering”*, GI/ITG Workshop MMBnet, pp. 13–14.

Murphy, J. M. and Maggioni, M. 2017. *"Unsupervised Clustering and Active Learning of Hyperspectral Images with Nonlinear Diffusion"*, available online at http://arxiv.org/pdf/1704.07961v5.

N. Pittaras, F. Markatopoulou, V. Mezaris, I. Patras, *"Comparison of Fine-tuning and Extension Strategies for Deep Convolutional Neural Networks"*, Proc. 23rd Int. Conf. on MultiMedia Modeling (MMM'17), Reykjavik, Iceland, Springer LNCS vol. 10132, pp. 102-114.

Neetu, Prashnani, M., Singh, D. K., Joshi, R. and Ray, S. S., 2014. *"Understanding crop growing pattern in barddhaman district of west bengal using multi-date risat 1 mrs data"*, The Inter-national Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Technical Commission VIII Symposium.

Newman, M. E. J. and Girvan, M. 2004. "*Finding and evaluating community structure in networks*", Physical Review E, vol. 69, p. 026113.

Nguyen, D., Clauss, K., Cao, S., Naeimi, V., Kuenzer, C. and Wagner, W., 2015. *"Mapping rice seasonality in the mekong delta with multi-year envisat asar wsm data"*, Remote Sensing7(12),pp. 15868–15893.

Nguyen, C., Starek, M., Tissot, P., Gibeaut, J. 2018. *"Unsupervised Clustering Method for Complexity Reduction of Terrestrial Lidar Data in Marshes",* Remote Sensing 10 (1), p. 133. DOI: 10.3390/rs10010133.

Papadopoulos, S., Kompatsiaris, Y., Vakali, A., and Spyridonos, P. 2012. "*Community detection in social media*", Data Mining and Knowledge Discovery, vol. 24 (3), pp. 515-554.

Pappas, T. N. 1992. *"An adaptive clustering algorithm for image segmentation"*, IEEE Trans. Signal Process. 40 (4), pp. 901–914. DOI: 10.1109/78.127962.

Pascucci, S., Carfora, M., Palombo, A., Pignatti, S., Casa, R., Pepe, M. and Castaldi, F. 2018. *"A Comparison between Standard and Functional Clustering Methodologies: Application to Agricultural Fields for Yield Pattern Assessment"*, Remote Sensing, vol. 10 (4), p. 585.

Pazhanivelan, S., Pandian, K., Christy, P. & Mary, N., Elangovan, S., Jeyaraman, S., et al. 2015. *"Rice crop monitoring and yield estimation through COSMO Skymed and TerraSAR-X: A SAR-based experience in India"*, ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-7/W3. 10.5194/isprsarchives-XL-7-W3-85-2015.

Pennington, J., Socher, R. and Manning, C.D. 2014. "*GloVe: Global Vectors for Word Representation"*.

Perronnin, F., Sanchez, J., Mensink, T., 2010. *"Improving the Fisher kernel for large-scale image classification"*, In: 11th Eur. Conf. on Computer Vision: Part IV (pp. 143-156). Springer-Verlag.

Pons, P. and Latapy, M. 2006. "*Computing communities in large networks using random walks*", Journal of Graph Algorithms and Applications, vol. 10 (2), pp. 191-218.

Qiao, J., Zhao, X., Pan, L., Zuo, X., Zhang, X., Zhang, Q. and Huang, S. (Eds.) 2018. *"Bio-inspired Computing: Theories and Applications",* Singapore: Springer Singapore.

Qin, J., Xu, J. J., Hu, D., Sageman, M. and Chen, H. 2005. “*Analyzing terrorist networks: A case study of the global salafi jihad network*”, Intelligence and security informatics (pp. 287–304). Springer.

Qin, Y., Xiao, X., Dong, J., Zhou, Y., Zhu, Z., Zhang, G., Du, G., Jin, C., Kou, W., Wang, J. and Li, X., 2015. *“Mapping paddy rice planting area in cold temperate climate region through analysis of time series landsat 8 (oli), landsat 7 (etm+) and modisimagery”*, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 105, p. 220–233.

Qiu, G., 2002. *“Indexing chromatic and achromatic patterns for content-based colour image retrieval”*, Pattern Recognition, vol. 35, p. 1675-1686.

Raghavan, U. N., Albert, R., and Kumara, S. 2007. “*Near linear time algorithm to detect community structures in large-scale networks*”, Physical Review E, vol. 76 (3), p. 036106.

Rahmah, N. and Sitanggang, I. S. 2016. *“Determination of Optimal Epsilon (Eps) Value on DBSCAN Algorithm to Clustering Data on Peatland Hotspots in Sumatra”*, IOP Conference Series: Earth and Environmental Science 31, p. 12012. DOI: 10.1088/1755-1315/31/1/012012.

Rasiwasia, N., Costa Pereira, J., Coviello, E., Doyle, G., Lanckriet, G.R., Levy, R. and Vasconcelos, N. 2010. *“A new approach to cross-modal multimedia retrieval”*, In Proceedings of the international conference on Multimedia (pp 251–260). ACM.

Ray, S. and Turi, R. H. 1999. *“Determination of number of clusters in k-means clustering and application in colour image segmentation”*, Proceedings of the 4th international conference on advances in pattern recognition and digital techniques. Calcutta, India, pp. 137–143.

Reuter, C. and Schröter J. 2015. *“Microblogging during the European floods 2013: What Twitter may contribute in German emergencies”*, International Journal of Information Systems for Crisis Response and Management, vol. 7, p. 22.

Richards, J. W., Hardin, J., Grosfils and E. B. 2010. *“Weighted model-based clustering for remote sensing image analysis”*, Computational Geosciences, vol. 14 (1), pp. 125–136.

Rosvall, M. and Bergstrom, C. T. 2008. “*Maps of random walks on complex networks reveal community structure*”, Proceedings of the National Academy of Sciences, vol. 105 (4), p. 1118-1123.

Rosvall, M., Axelsson, D. and Bergstrom, C. T. 2009. “*The map equation*”, The European Physical Journal Special Topics, vol. 178, p. 13-23.

Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. In Journal of computational and applied mathematics 20, pp. 53–65.

Saffer, A.J., Sommerfeldt, E.J. and Taylor, M., 2013. *“The effects of organizational Twitter interactivity on organization–public relationships”*, Public Relations Review, vol. 39, p. 213–215.

Said, N., Pogorelov, K., Ahmad, K., Riegler, M., Ahmad, N., Ostroukhova, O., Halvorsen, P. and Concin, N. 2018. *“Deep Learning Approaches for Flood Classification and Flood Aftermath Detection”*, MediaEval 2018, Multimedia Benchmark Workshop

Saravanou, A., Valkanas, G., Gunopulos, D. and Andrienko, G., 2015. *"Twitter Floods when it Rains: A Case Study of the UK Floods in early 2014"*, Proceedings of the 24th International Conference on World Wide Web, ACM, pp. 1233–1238.

Savaresi, S. M. and Boley, D. L. 2001. *"On the performance of bisecting K-means and PDDP"*, Proceedings of the 2001 SIAM International Conference on Data Mining. SIAM, pp. 1–14.

Savaş, C., Yıldız, M., S., Eken, S., İkibaş, C. and Sayar, A. 2019. *"Clustering Earthquake Data"*, In Albert Gyamfi, Idongesit Williams (Eds.): Big data and knowledge sharing in virtual organizations, vol. 56. Hershey PA: Engineering Science Reference (Advances in Knowledge Acquisition, Transfer, and Management), pp. 224–239.

Schinas, M., Papadopoulos, S., Petkos, G., Kompatsiaris, Y. and Mitkas, P.A., 2016. *"Multimodal event detection and summarization in large scale image collections"*, Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, ACM, pp. 421-422.

Schneider, J. and Vlachos, M. 2013. "*Fast parameterless density-based clustering via random projections*", Proceedings of the 22nd ACM international conference on Conference on information & knowledge management, ACM, pp.861–866.

Schubert, E., Sander, J., Ester, M., Kriegel, H.P. and Xu, X., 2017. "*dbscan revisited, revisited: why and how you should (still) use dbscan*", ACM Transactions on Database Systems (TODS) vol. 42, p. 19.

Setiyono, T., Quicho, E., Gatti, L., Campos-Taberner, M., Busetto, L., Collivignarelli, F., Javier García Haro, F., Boschetti, M., Islam Khan, N. and Holecz, F. 2018. *"Remote Sensing Spatial Rice Yield Estimation Based on MODIS and Sentinel-1 SAR Data and ORYZA Crop Growth Model"*, Remote Sensing, vol. 10.

Shahdoosti, H. R. and Ghassemian, H. 2016. *"Combining the spectral PCA and spatial PCA fusion methods by an optimal filter"*, Information Fusion, vol. 27, pp. 150–160.

Siddiquie, B., White, B., Sharma, A. and Davis, L.S. 2014. *"Multi-modal image retrieval for complex queries using small codes"*, In Proceedings of international conference on multimedia retrieval. ACM, p 321.

Simonyan, K. and Zisserman, A., 2014. *"Very deep convolutional networks for large-scale image recognition"*, CORR, arXiv technical report.

Skakun, S. 2012. *"A neural network approach to flood mapping using satellite imagery"*, Computing and Informatics, vol. 29(6), 1013-1024.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. *"Going deeper with convolutions"*, In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. 2016. *"Rethinking the inception architecture for computer vision",* In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826.

Takahashi, B., Tandoc Jr, E.C. and Carmichael, C., 2015. *"Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines"*, Computers in Human Behavior, vol. 50, p. 392–398.

Teh, Y.W., Jordan, M.I., Beal, M.J. and Blei, D.M., 2012. "*Hierarchical dirichlet processes*", Journal of the American Statistical Association.

Tian, H., Wu, M., Wang, L. and Niu, Z. 2018. *"Mapping Early, Middle and Late Rice Extent Using Sentinel-1A and Landsat-8 Data in the Poyang Lake Plain, China"*, Sensors. 18. 185. 10.3390/s18010185.

Torbick, N., Chowdhury, D., Salas, W. and Qi, J. 2017. *"Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2"*, Remote Sensing, vol. 9. 119.

Twele, A., Cao, W., Plank, S. and Martinis, S. 2016. *"Sentinel-1-based flood mapping: a fully automated processing chain"*, *International Journal of Remote Sensing*, vol. 37(13), p. 2990-3004.

Vedaldi, A. and Lenc, K., 2015. *"Matconvnet – convolutional neural networks for matlab"*, In Proceeding of the ACM International Conference on Multimedia.

Verma, A. K., Garg, P. K. and Prasad, K. H. 2017. *"Sugarcane crop identification from LISS IV data using ISODATA, MLC, and indices based decision tree approach"*, Arabian Journal of Geosciences 10 (1), p. 16.

Vieweg, S., Hughes, A.L., Starbird, K. and Palen, L., 2010. *"Microblogging during two natural hazards events: what twitter may contribute to situational awareness"*, Proceedings of the SIGCHI conference on human factors in computing systems, ACM, pp. 1079–1088.

Wagstaff, Kiri, Cardie, Claire, Rogers, Seth, Schrödl, Stefan, others (2001): Constrained k-means clustering with background knowledge. In : ICML, vol. 1, pp. 577–584.

Wang, J., He, Y., Kang, C., Xiang, S. and Pan, C. 2015. *"Image-text cross-modal retrieval via modality-specific feature learning"*, In: Proceedings of the 5th ACM on international conference on multimedia retrieval. ACM, pp 347–354

Wang, W., Ooi, B.C., Yang, X., Zhang, D. and Zhuang, Y. 2014. *"Effective multi-modal retrieval based on stacked auto-encoders"*, Proc VLDB Endow, vol. 7(8), p. 649–660.

Wang, Y., Lin, X. and Zhang, Q. 2013. *"Towards metric fusion on multi-view data: a cross-view based graph random walk approach"*, In Proceedings of the 22nd ACM international conference on conference on information & knowledge management. ACM, pp 805–810

Xie, J., Girshick, R., and Farhadi, A. 2016. *"Unsupervised deep embedding for clustering analysis"*, International conference on machine learning (pp. 478-487).

Xu, H. 2006. *"Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery"*, *International journal of remote sensing*, vol. *27*(14), p. 3025-3033.

Xu, S., Li, H., Chang, X., Yu, S.I., Du, X., Li, X., Jiang, L., Mao, Z., Lan, Z., Burger, S. et al 2015. *"Incremental multimodal query construction for video search"*, In Proceedings of the 5th ACM on international conference on multimedia retrieval. ACM, pp 675–678

Younessian, E., Mitamura, T. and Hauptmann, A. 2012. *"Multimodal knowledge-based analysis in multimedia event detection"*, In Proceedings of the 2nd ACM international conference on multimedia retrieval. ACM, p 51.

A Appendix

A.1. Change Detection in EOPEN and CANDELA

Change detection	EOPEN	CANDELA
Input data	**Problem 1: Water body detection** • Sentinel-1 + DEM for slope reduction • Sentinel-2 for water delineation **Problem 2: Change detection for flood/no-flood estimation** • Sentinel-2 for water relative changes RS and ML **Problem 3: Rice paddy mapping** • Sentinel-1 images **Problem 4: Snow depth estimation** • Sentinel-1 • Twitter data **Problem 5: Road passability estimation** • VHR images (WorldView3 from DigitalGlobe (0,3 m resolution))	All Sentinel-2 data products
Input data format	**Problem 1: Water body detection** • Sentinel-1 product in .zip format as they become available from Copernicus. • DEM in TIFF format. • Sentinel-2 product in .zip format as they become available from Copernicus with JPEG2000 files holding spectral band BOA reflectance. **Problem 2: Change detection for flood/no-flood estimation** • Sentinel-2 products in patches as they became available from	GeoTIFF *Sentinel-2 data products are converted into GeoTiff images thanks to a function implemented by TAS FR on CANDELA platform*

<table>
<tr>
<td></td>
<td>Multimedia Satellite task 2019 - http://www.multimediaeval.org/mediaeval2019/multimediasatellite/

Problem 3: Rice paddy mapping
• VH polarized backscatter of Sentinel-1 images

Problem 4: Snow depth estimation
• Sentinel-1 VV and VH
• Twitter posts in JSON

Problem 5: Road passability estimation
• Satellite image patches as they became available during the Multimedia Satellite task 2018 - http://www.multimediaeval.org/mediaeval2018/multimediasatellite/</td>
<td></td>
</tr>
<tr>
<td>Features extracted</td>
<td>Problem 1: Water body detection
• VV + VH from Sentinel-1 image
• DEM as an array from the area of interest
• MNDWI index from Sentinel-2 image

Problem 2: Change detection for flood/no-flood estimation
• DCNN-based: ResNet101, VGG-16, VGG-19, Inception, Inception_ResNet_v2

Problem 3: Rice paddy mapping
• VV + VH from Sentinel-1

Problem 4: Snow depth estimation
• VV + VH from Sentinel-1
• number of relevant-to-snow tweets
• number of Twitter images that contain the concept “snow”</td>
<td>An encoded representation

More concretely, a 25-dimensional vector of each input patch</td>
</tr>
</table>

	Problem 5: Road passability estimation • DCNN-based: ResNet101, VGG-16, **VGG-19**, Inception, Inception_ResNet_v2	
Supervised/Unsupervised	**Problem 1: Water body detection** • Supervised (slope reduction with Sentinel-1 and DEM using Deep Neural Networks) • Unsupervised (MNDWI thresholding with Sentinel-2 images) **Problem 2: Change detection for flood/no-flood estimation** • Supervised with pretrained differences of RGB images • Unsupervised with differences of MNDWI and outlier detection **Problem 3: Rice paddy mapping** • Unsupervised and supervised approaches examined (paper in press: DOI: 10.1109/TGRS.2020.2981671) **Problem 4: Snow depth estimation** • Unsupervised estimation of snow depth and fusion with the number of relevant-to-snow posts and Twitter images **Problem 5: Road passability estimation** • Supervised on Sentinel-2 patches using Deep Learning	Unsupervised
Dataset for experiments	**Problem 1: Water body detection** For Sentinel-1 + DEM annotated data of lakes provided by the EOPEN partner: AAWA (Italian water authority) For Sentinel-2: satellite images near Vicenza with ground truth about water bodies **Problem 2: Change detection for flood/no-flood estimation** MediaEval 2019 Satellite Task 4: "City-centered satellite sequences" with	*Dataset used by TAS FR* **Urbanization:** Pairs of Sentinel-2 images over several countries, like United Arab Emirates, Turkmenistan, Angola and Vietnam *(another dataset will be proposed by TerraNIS)*

	timeseries of flood events in urban areas. **Problem 3: Rice paddy mapping** Dangjin city, South Korea, EOPEN partners: Korea University and Sundosoft LtD. Paper in press: DOI: 10.1109/TGRS.2020.2981671 **Problem 4: Snow depth estimation** Sentinel-1 images of Finland of year 2018-2019. 3,200 annotated tweets by the EOPEN partner: Finnish Meteorological Institute within this period and area. **Problem 5: Road passability estimation** 1,437 VHR satellite image patches, provided by MediaEval 2018 Satellite Task: "Emergency Response for Flooding Events"	**Forest fire:** Pairs of Sentinel-2 images over an area of the Amazon forest affected by the wildfires in 2019 *(another dataset will be proposed by SmallGIS)* **Vineyards:** Pairs of Sentinel-2 images over a region near Bordeaux affected by a frost in April 2017 **Maritime traffic:** Time-series of Sentinel-2 images over a harbor zone close to Marseille from December 2017 to March 2018
Open/closed source	Open source for all cases	Open source for all use cases
API or GUI available	**Problem 1: Water body detection** GUI (EOPEN) **Problem 2: Change detection for flood/no-flood estimation** Available on demand from CERTH **Problem 3: Rice paddy mapping** No **Problem 4: Snow depth estimation** No	API is available on Candela platform *It is the same API for all the use cases*

	Problem 5: Road passability estimation Available on demand from CERTH	
Use of Docker	**Problem 1: Water body detection** Yes **Problem 2: Change detection for flood/no-flood estimation** No **Problem 3: Rice paddy mapping** Yes **Problem 4: Snow depth estimation** No **Problem 5: Road passability estimation** No	Yes *It is the same Dockers for all the use cases*

A.2. Connection to the beAWARE project

The H2020 beAWARE project proposes an integrated solution to support forecasting, early warnings, transmission and routing of the emergency data, aggregated analysis of multimodal data and management of the coordination between the first responders and the authorities. beAWARE started on 01/01/2017 (and concluded on 31/12/2019), while EOPEN followed on 01/11/2017. The two projects share some tasks that involve similar types of data, such as social media data, meteorological data, and a hydrological model, while there is also a common use case, i.e. floods in north-eastern Italy. However, the goals and the approaches of each project are completely different and are described in this appendix (please also see Table 8).

Regarding the acquisition of social media data, the beAWARE Crawler has served as a basis for EOPEN, but the search criteria had to change. In beAWARE they referred to floods, fires and heatwaves, whereas in EOPEN to floods, snow coverage and food security. Also, the criteria in EOPEN have been extended to include not only keyword-based queries, but user-based and location-based queries, too. The task of relevance estimation has been part of both projects, but in EOPEN updated models have been developed, based on recent annotation and more evolved text classification technologies. Moreover, beAWARE included a relevance estimation technique based solely on images. Clustering was another common task, but in beAWARE grouping was taking into consideration the location and the time of tweets, whereas in EOPEN the similarity between the texts/images of the tweets. Finally, tasks that concern social media data, such as concept extraction, community and event detection, and localisation, have been introduced in the EOPEN project.

As far as it concerns the meteorological data, the aim in beAWARE was to create real-time warnings based on HIRLAM forecasts. Furthermore, the use cases simulated extreme weather events by using historical meteorological data. In EOPEN, the use cases are more general and the data extends more broadly in both geographical and temporal scales.

In regards with the hydrological model, AAWA improved it by ingesting the HIRLAM forecast in the context of the beAWARE project, while in EOPEN they improved the entire EWS with data from Copernicus services (LAND), HIRLAM forecast and the new hydraulic model Basement (2D model), useful also for mountain rivers due to the solid transport availability. The role of AAWA in beAWARE was more oriented to field operations with volunteers; beAWARE provided an interface for the management of an emergency (civil protection volunteers). Based mainly on alerts from the beAWARE mobile app and the forecast of the hydraulic model (only 1D), an algorithm was developed to calculate the risk in real time.

In addition, the EOPEN project involves the acquisition and analysis of Earth Observation data, a source of information that was completely excluded in beAWARE.

Table 8: Comparison between beAWARE and EOPEN projects

	beAWARE H2020 DRS, 2017-2020	EOPEN H2020 EO RIA, 2017-2020
Social media data		
Twitter crawling	●	●
Relevancy estimation	●	●
Concept extraction		●
Clustering	●	●
Community detection		●
Event detection		●
Localisation		●
Satellite data		
Water mask generation		●
Change detection		●
Road passability		●
Meteorological data		
NWP forecasts (HIRLAM)	●	●
Extreme weather time series	●	
Weather observations		●
Climatological trends		●
Climatological model output		●
AAWA EWS		
AAWA hydrological model with HIRALM forecast	●	
AAWA hydraulic model (Basement)		●
Satellite data into the		●

hydrological module		
AAWA Risk calculation based on alerts and forecast	●	
AAWA event detection based on forecast		●