InVID

Deliverable D3.3: Final verification framework (confidential document - only a summary is publicly available)

Markos Zampoglou, Symeon Papadopoulos, Giorgos Kordopatis-Zilos,
Olga Papadopoulou, Vasileios Mezaris, Alexandros Metsai,
Konstantinos Apostolidis, Yiannis Kompatsiaris, Lazaros Apostolidis,
Christos Koutlis, Dimitrios Giomelakis / CERTH
Lyndon Nixon, Adrian M.P. Brasoveanu / MODUL
Roger Cozien, Gregoire Mercier / EXO MAKINA

24/12/2018

Work Package 3: Content Verification

InVID - In Video Veritas: Verification of Social Media Video Content for the News Industry

Innovation Action

Horizon 2020, Research and Innovation Programme

Grant Agreement Number 687786

Dissemination level	CO
Contractual date of delivery	31/12/2018
Actual date of delivery	24/12/2018
Deliverable number	D3.3
Deliverable name	Final verification framework (confidential document - only a summary is publicly available)
File	`InVID_D3.3.tex`
Nature	Report
Status & version	Final & V1.0
Number of pages	3
WP contributing to the deliverable	3
Task responsible	CERTH
Other contributors	MODUL, EXO MAKINA
Author(s)	Markos Zampoglou, Symeon Papadopoulos, Giorgos Kordopatis-Zilos, Olga Papadopoulou, Vasileios Mezaris, Alexandros Metsai, Konstantinos Apostolidis, Yiannis Kompatsiaris, Lazaros Apostolidis, Christos Koutlis, Dimitrios Giomelakis / CERTH Lyndon Nixon, Adrian M.P. Brasoveanu / MODUL Roger Cozien, Gregoire Mercier / EXO MAKINA
Quality Assessors	Denis Teyssou / AFP, Gregoire Mercier / EXO
EC Project Officer	Alberto Rabbachin
Keywords	Content verification, video forensic analysis, near-duplicate detection, logo detection, context aggregation, location detection

Abstract

The InVID project has come to its conclusion. The tools and technologies that we have been developing, integrating, and testing during the previous years are now in their final form. Work Package 3 has provided an array of tools, integrated in a powerful platform, which aims to provide journalists and investigators with enhanced capabilities in verifying news-related user-generated videos. The previous two deliverables of the Work Package, D3.1 and D3.2, described the tools that were developed, the state-of-the-art algorithms and technologies that were implemented, adapted, and improved, the user feedback that was received and the way it shaped the component development, and the status of integration at the end of the first and second year respectively. Here, we present the final versions of the components and services, the achieved improvements in algorithmic and service performance, and the final integration status. Through their integration with the Verification Plugin and the Verification Application, these components of InVID have been seeing increasing real-world usage since the second project year, and an increased uptake in the third year. Compared to the previous year, all components have seen substantial improvements.

- Our work in *Video Forensics* was geared towards automated or semi-automated video analysis. The video forensics filters designed during Year 2 and described in D3.2 were used to train and evaluate a system based on deep convolutional neural networks, designed to take the filter outputs and return a binary tampering localization mask that can be easily interpreted by a human user. In parallel, our previous year's work into convolutional neural networks for tampering detection (i.e. taking the filter outputs and returning a single-value result on the probability that the video was tampered) was significantly extended with further models, datasets, and experiments. Thus, our video forensics work in year 3 included a two-pronged approach into deep learning for tampering detection and (semi-) automation of the interpretation of filter outputs.

- In *Near-Duplicate Detection*, the algorithm developed during the previous years was further improved, leading to an approach that further surpasses the state of the art in accuracy. This is achieved by combining the proposed Deep Metric Learning approach with a Chamfer Distance metric to exploit the distances between video frames during video similarity calculation. We also completed the development of a very large-scale dataset which allows for realistic evaluations of Near-Duplicate Video Retrieval algorithms, and also enables evaluations in Fine-grained Video Retrieval tasks. Furthermore, several improvements and extensions were made in the service functionalities according to the obtained feedback.

- The *Logo Detection* module was improved by replacing it with a more reliable deep learning framework, extending its coverage with user submitted contributions, and improving its performance by adapting and extending the synthetic training data generation process with further training data augmentation steps.

- The *Location Detection* module was further improved in terms of accuracy through the inclusion of several disambiguation steps that reduce the number of errors and lead to increased performance. Furthermore, our efforts to provide a more reliable evaluation dataset and methodology have led to the development of an entire ecosystem of tools for the integration of the Recognyze tool but also for performance evaluations.

- Finally, the *Context Aggregation and Analysis* module was extended with new functionalities, and underwent improvements with respect to speed, reliability, and the structure of the provided information. In parallel, the increasing user base that has been developed during the second and third year of InVID has allowed us to use this component to significantly expand the Fake Video Corpus dataset. Combined with the Near-Duplicate Detection algorithm, a large dataset including well-established cases of fake and real videos was created, including their reposts and near-duplicates. In the context of the CAA component, the characteristic patterns of this dataset were explored, with the aim of gathering insights for contextual video verification.

The integration of these components is now complete, having reached a level of seamless interaction with the InVID platform. Their constant use in operational conditions guarantees that, besides their achievements with respect to evaluations and benchmarks, these tools are also ready for large-scale real-world use, providing state-of-the-art performance for real-world video verification.

