ELIS CONSULTING & LABS

Spotting Image Manipulations with Deep Learning

Abstract

The advancements of photo editing tools and techniques to generate artificial images and videos introduce new challenges to detect fake contents. While deep learning and computer vision can help to detect multimedia content alterations, new techniques must be developed to spot doctored contents. This showcase presents some of the open challenges that ELIS Consulting & Labs is facing to address these problems.

elis
Consulting & Labs

Erasmus+
Key competences for an European model of Industry 4.0
i4EU Project

Learn more about us here:
https://www.i4eu-pro.eu

Co-funded by the
Erasmus+ Programme
of the European Union

Introduction

The advancement of photo editing tools allows an increasingly large number of people to easily manipulate images, thus making easier to defraud companies or enabling fake news to spread out. Multimedia forensics solutions and recent deep learning techniques allow us to design neural networks that are suitable for detecting manipulated areas in an image, if any, or to reconstruct the source and history of multimedia content. ELIS Consulting & Labs works on the design of new procedures of automated manipulation detection, in order to offer meaningful support to the process of forensic analysis of images and videos.

The next sections explore challenges that need to be overcome to solve such problems and present two real problems that ELIS Consulting & Labs is solving for large Italian corporates, as well as the skills that one has to master to work in this field.

Challenges

Trust in what we see is increasingly important in a world where image editing becomes easier and accessible for everyone. The recent advancements in machine learning, introduced new important challenges, with the widespread diffusion of the so called *deepfakes*, i.e. fake videos or images artificially generated by neural networks. While these new technologies evolve and become more realistic day by day, distinguishing between real and fake content becomes increasingly difficult.

> “*Trust in what we see is increasingly important in a world where image editing becomes easier and accessible for everyone.*

Photo editing tools allow for a wide range of image manipulations. Image manipulations can be classified by three main basic operations: (i) to cut the content of an image and paste it into another image, (ii) to copy an object and reproduce it into the same image, and (iii) to remove a specific content. All these operations leave some fingerprints that can be identified by an intelligent algorithm. Even

Co-funded by the
Erasmus+ Programme
of the European Union

more, recently, it has been proven in [1] that Generative Adversarial Models (GANs), usually employed to generate deepfakes, leave some traces that can be identified. The spread of such tools is a potential problem for fake news and defraud applications of image forgery. Deep learning can be applied to detect altered content, but large training examples are needed, therefore introducing new challenges to generate sufficiently large training datasets.

How do we solve the problem?

ELIS Consulting & Labs is deeply investing on the development of forgery detection systems able to protect companies from fraud attempts. While fraud attempts increase, multimedia forensic is a flourishing research area with wide range of improvements. Although computer vision and deep learning have shown rapidly increasing performances during the last years on tasks like object detection, pattern matching, or segmentation, all these techniques cannot be easily applied to spot fake images. To solve this much more complex problem, new neural network architectures and design processes are needed.

Recently, an R&D project has been conducted for an insurance company. Doctored images usually show noise inconsistencies that highlight manipulation fingerprints. A two-stream neural network architecture has been designed to analyze RGB and noise incompatibilities; then features learned by the two streams are merged to increase detection accuracy. In order to train the model, an accurate dataset has been produced to provide real-world examples of manipulations to the neural network. Training the network required about two weeks of computation on Azure virtual machines. This system has been released to insurance experts through a web application. When an expert uploads an image on the app, any manipulations are exposed with boxes identifying manipulated areas.

Today, online services usually require verifying the identity of a registering user. Document ID can be counterfeited or downloaded from social networks. In this scenario, manipulation detection is not enough, and also the source of an image suggests possible identity theft. ELIS Consulting & Labs is

Co-funded by the Erasmus+ Programme of the European Union

working on the development of an identity document verification pipeline that is capable to automatically inspect documents verifying the integrity and authenticity. Source identification can be performed training a model to recognize camera fingerprints and visual clues left by every single camera. Even more, downloading and uploading files from social networks leaves traces that can be analyzed. To address this, ELIS Consulting & Labs is building an experimental ensemble model that is trained using signals from multiple detectors.

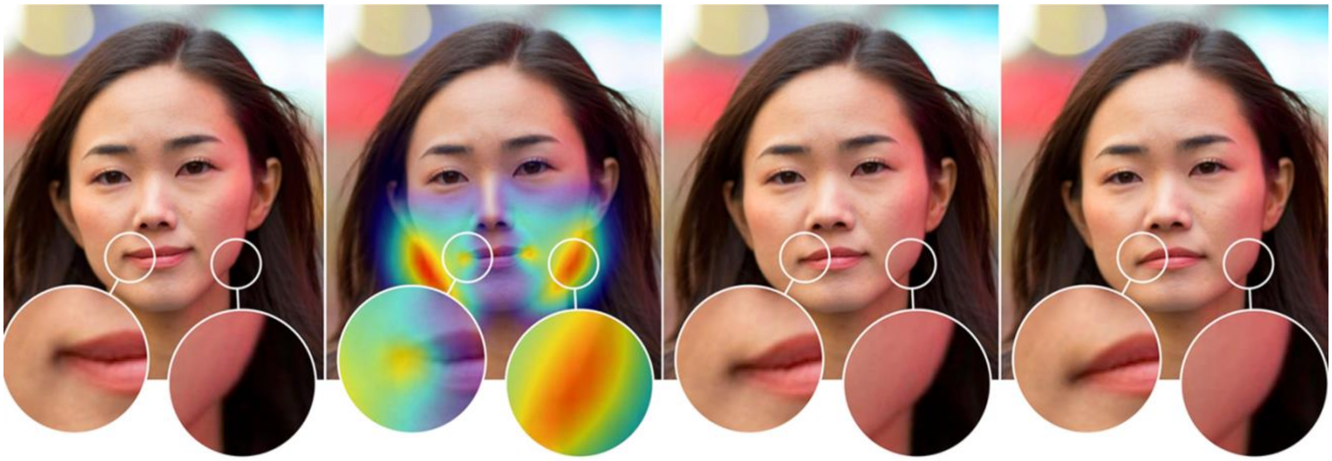

Figure 1 An example of facial manipulation [2].

With the continuous evolution of challenges and technologies involved in this setting, professionals are required to build new skills. Computer vision experts are required to master a wide range of knowledge: machine learning, deep learning, signal processing, linear algebra, multivariate calculus, and programming are the fundamentals of a highly skilled expert. In the world of social media where digital content is produced at incredible high speed, companies and experts leading on the fake content detection technologies will have a cleared road to guarantee trust and verification of multimedia content. For the companies investing in this domain, the main benefit of autonomous systems capable of spotting image manipulations lies in the cost reduction implied by the automatization of critical tasks such as integrity checking and fake detection on images.

Co-funded by the
Erasmus+ Programme
of the European Union

References

1. Francesco Marra and Diego Gragnaniello and Luisa Verdoliva and Giovanni Poggi, *Do GANs leave artificial fingerprints?,* CoRR http://arxiv.org/abs/1812.11842
2. *Adobe Research And Uc Berkeley: Detecting Facial Manipulations In Adobe Photoshop* https://theblog.adobe.com/adobe-research-and-uc-berkeley-detecting-facial-manipulations-in-adobe-photoshop/

This showcase has been collected in the framework of the Erasmus+ project Key competences for an European model of Industry 4.0 (pr. n° 2019-1-FR01-KA202-062965), funded by European Commission.

For more information: www.i4eu-pro.euLegal notice: This publication / communication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.