ITALIAN COVID-19 QUARANTINE

USING PUBLIC DATA AND VISUAL AI TO IDENTIFY VIOLATIONS

Powered by LOGOGRAB

Introduction

The research group, Ghost Data, and Visual-AI company, LogoGrab, have combined their technological competencies to provide some support for Italian authorities in fighting the COVID-19 crisis in regards to the current country lockdown.

The goal of our study is to tentatively understand risks, behaviours and trends of the local population in this stage of the crisis through the use of big data analysis and image and video recognition technology.

About The Co-Authors

Andrea Stroppa - Founder

Ghost Data

Ghost Data is a group of researchers based between Italy and US, that performs complex data analysis focused on Instagram and TikTok. Their research on terrorism, luxury counterfeiting, bots, digital propaganda appeared on The New York Times, Washington Post, AP, NBC News, WSJ among others. Recently a report produced by Ghost Data about online counterfeiting was included by the US home office in a report for the President of the United States.

Luca Boschin - CEO
Alessandro Prest - CTO

LogoGrab

LogoGrab is a Visual-AI technology provider for applications in Brand Protection, Authentication, and Monitoring. Established in 2016, LogoGrab now provides Visual-AI solutions to global SaaS platforms through its enterprise Visual-AI API suite.

Methodology

This study harvested images and video from 552,000 profiles and their public Instagram stories (totalling 504,592 stories), covering the period from March 11th to March 18th, 2020.

The data was anonymised as it was gathered and prior to processing. This meant stripping out any personal data, such as profile information, decreasing geo-location precision, blurring faces and removing metadata that could be used to identify individuals.

The final data set contained macro-level geo-tagging and demographic data, which was then stored for processing. Where content did not include geo-information, the Ghost Data technology was able to determine the location to a regional level using other relevant profile data. This technique was perfected during previous research in online counterfeiting, terrorism and digital propaganda.

LogoGrab's Visual-AI technologies were then used to process the data. Object & Scene detection was used to identify images of people in an outdoor environment and public indoor settings (bars, restaurants, etc.) as well as driving or riding vehicles.

The final data set was then compared against the policies released by DCPM (Decree of the President of the Council of Ministers) and other institutional sources defining the rules applied to the current lockdown.[1,2]

[1]"Decreto #IoRestoaCasa, domande frequenti sulle misure adottate dal Governo" , Presidenza del Consiglio dei Ministri, 17 marzo 2020 http://www.governo.it/it/faq-iorestoacasa

[2] "Coronavirus:sì alle passeggiate ma uscite da soli. Cosa si può fare", La Repubblica, Alessandra Ziniti ,13 marzo 2020 https://www.repubblica.it/cronaca/2020/03/13/news/coronavirus_si_alle_passeggiate_ma_uscite_da_soli_cosa_si_puo_fare-251118421/

LOGOGRAB

REVIEW

GHOST DATA TARGETED

552'000

PUBLIC POSTS

Instagram Stories

VISUAL AI

- OBJECT & SCENE DETECTION
- OCR

REVIEW OF IDENTIFIED SUSPICIOUS CASES

GEO TAGGING

DEMOGRAPHICS

Results

The Visual-AI initially identified a pool of 5% of posts that contained suspicious content.

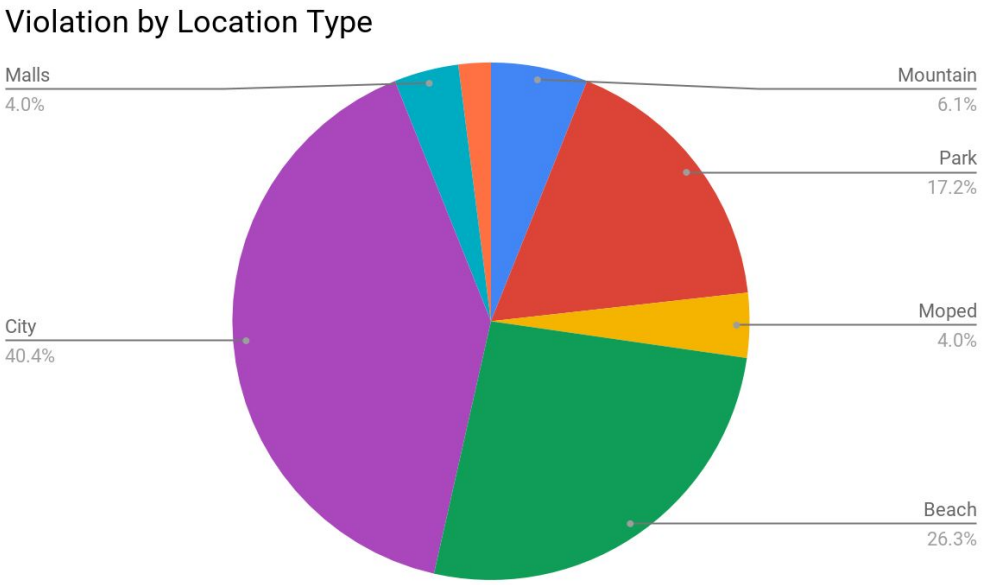
Further filtering was then applied using Optical Character Recognition to identify posts which were tagged with overlay text indicating that the post was not new (e.g., the date of the picture was overlaid in text, or birthday wishes and other popular online contests, identified by specific hashtags). This removed a further 13% from the initial pool of suspicious posts.

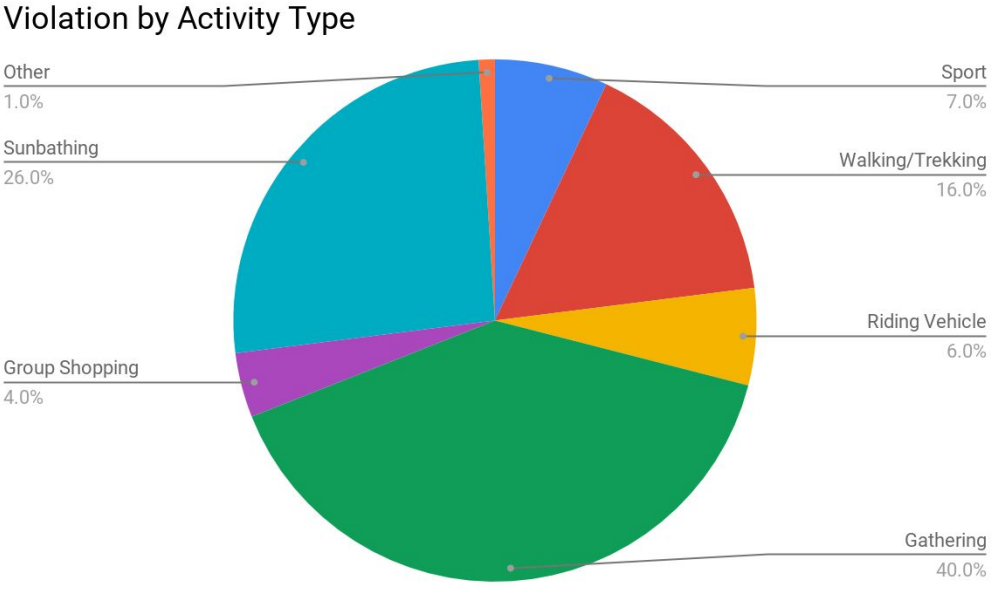
The remaining pool of suspicious media were then manually reviewed by a team of operators trained on the COVID-19 lockdown policy.

The manual review led to a final data set of 6% of the original pool being positively verified as violating lockdown policy rules.

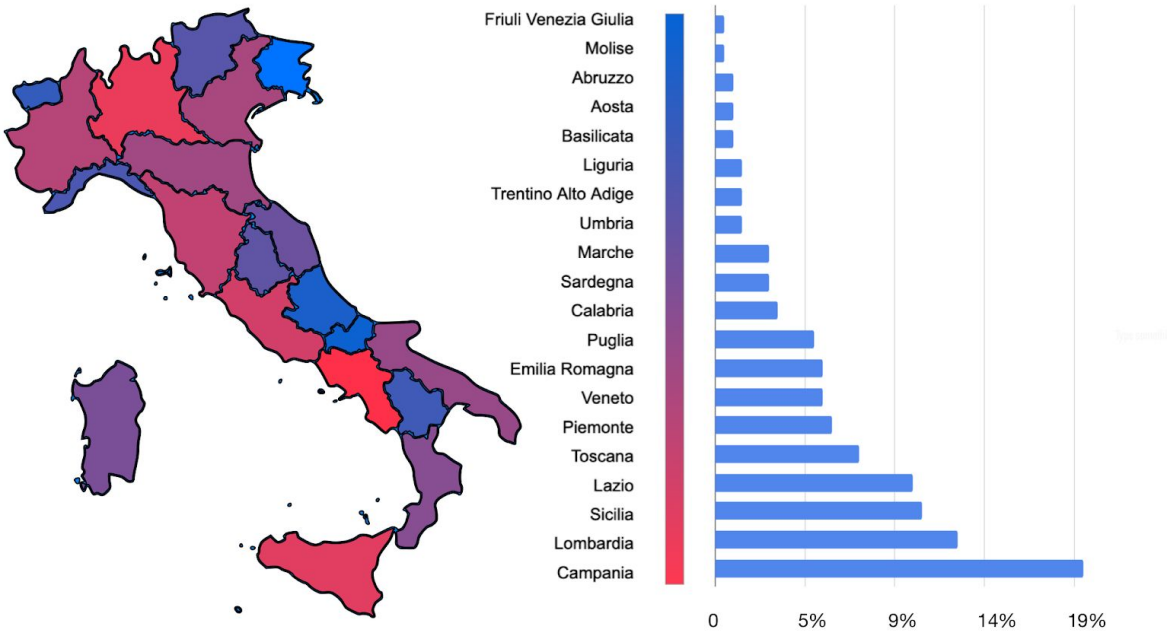
Initial Data

Here are some initial findings:





Regional Analysis Of Violations:



Conclusion

Our initial findings (March 11-18, 2020) show a clear need to increase checking points in urban areas, particularly near public parks and beaches. The most common violations include people gathering in urban areas (40%), sport activities in a group such as walking or hiking (23%), sunbathing on beaches (26%).

Future data analysis could also include time-stamp trends, providing even more granular data to help mitigate these types of violations by focusing the limited resources available to monitor lockdown compliance.

Additionally, a comparison study on current data would also show if the trend of conformance is improving and further, if this correlates with infection rates.

Some Visual Examples