Coding
Practical AI
Coffee Grinders
llanoajm@gmail.com | lauragomezjurado1@gmail.com | scidroidgames@gmail.com

Smogify

About our project

Under the growing uncertainty of climate change, and the threat it poses to the health of millions of people around the world exposed to the toxic levels of air pollution, *Smogify* is a mobile app powered by computer vision able to identify the pollution level of public routes based on overlaid data consisting of the visible emissions from every individual actor in traffic. Therefore, *Smogify* provides an insight into the distribution of pollution within an urban community.
In order to calculate the visible emissions, the computer vision software is fed with data from video through crowdsourcing. This service is oriented on applying Artificial Intelligence to combat climate change by tracing pollution levels in the most contaminated areas due to transportation. Given the intersection between the climate crisis and other fields, [*Smogify*] yields a wide range of benefits, including improving public health in areas affected by high levels of pollution, and even adding a novel criterion for finding the best route an Uber can take to be more environmentally friendly, get additional input over potential traffic congestion and avoid potential health threats.

The problem that needs to be solved.

In Colombia, every car must undergo a technical inspection periodically to check for excess emissions and to repair the car if necessary. However diligent this inspection might be, we have been stuck behind a public-transport bus, bewildered by its pitch-black smog way too many times not to question: “is this even allowed?” – it is not. What is going wrong, then?

As per *Superfreakonomics*, 1 out of every 140 miles driven are driven drunk. Likewise, “There is just one arrest for every 27,000 miles driven while drunk.” This data, although from the United States, pertains to a great degree to our question: *why are these deprecated vehicles still circulating?* Perhaps, they are simply not being asked for their inspection certificates. After all, in Colombia, we sport not only one license plate, but two, and officers should be able to check for the revision of suspicious vehicles in an instant. The problem is how we identify a vehicle as suspicious. The human eye clearly doesn’t work for identifying drunk drivers, and we believe it

doesn't work for identifying polluting cars either. The smog is visible, but we have grown accustomed to it: currently, our only objective measure for our cars' condition is the inspection. We want to add another objective measure with the help of AI. We want to identify the smog of every individual actor in traffic through crowdsourcing and overlay the data in GPS systems. Identifying a polluting car will draw a yellow trail for as long as the user's camera and the polluting car stay in the same route. This yellow data should get overlaid for every polluting car, getting darker as they get stacked on top of each other. This way we get an estimate of the overall pollution in a given route.

We believe we are very repetitive in the routes we take: we go to school, to our friend's house, or to the supermarket most of the time. Public transport buses follow the same route everyday and so do cargo trucks. Thus, the routes people take tell a story about their routine, and also about their car.
Cars that frequently go through polluted regions are likely polluting themselves.

Collecting this data can be extremely valuable because it provides statistically-prone pools to check for polluting cars. We believe it can also be highly valuable to public-health officials in identifying urban areas that suffer from vehicle contamination. Last but not least, we can add a layer of intelligence to GPS software: our goal shouldn't just be to arrive in the shortest time possible, perhaps we want to avoid cancerogenous gasses as well.

Target Users

We are developing a mobile app that is anchored to our own free-for-use API. This expands our scope from users with hand-held devices to traffic and health authorities and even other data scientists. GPS systems currently offer minimal e.t.a as their main criterion, having a criteria for how clean the air is could be of use to rideshare apps in giving more options to its users.

Benefits

- Provide officers with data on the routes where polluting vehicles are most likely to appear.
- Promote initiatives towards public health in areas suffering from vehicle-generated pollution.
- Add an extra layer of intelligence to GPS systems that allows users to take less contaminated routes: this is especially important in densely populated cities like Bogotá or Delhi.

Current Status

Smogify is a concept currently being developed. The technology and action plan has already been identified and it was concluded the feasibility to start the first prototype. The main area of growth has been in identifying our AI model of choice: VGG-Net to achieve higher rates of accuracy and therefore maximize the reliability and applicability of the app.

Technologies

Since the core service of *Smogify* is based on the recognition of pollution from images (within a video), Smogify will be powered by a deep learning model for image classification; to classify images based on whether there is or not pollution. In order for the users to know the different levels of pollution, the app will receive the verdict from the model (the verdict can be the presence or absence of pollution) and when it identifies presence, the map will show the route with the overlapped verdicts (further explained concluding the Technologies section).

The deep learning model is based on Visual Geometry Group (VGG), which is a deep Convolutional Neural Network (CNN) architecture with multiple layers (Figure 1).

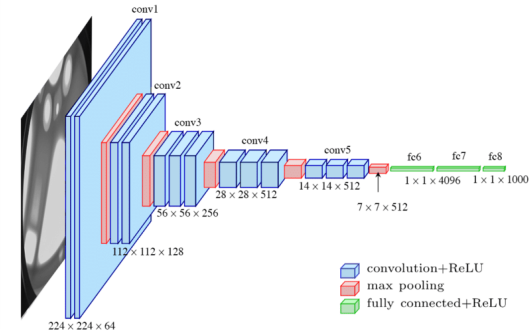


Figure 1. VGG with 16 layers for image processing.

The “deep” refers to the number of layers. In order to feasible its implementation as a mobile app, our architecture is based on a 2-dimensional network that is developed using the libraries; Keras, which is a Python Deep Learning API that follows best practices for reducing cognitive load and widely use for Neural Networks implementations; an Tensorflow, which is an open source library commonly use for different Machine and Deep Learning implementations.
In order to serve the model will use Apple’s CoreML or Tensorflow JS.

Data collection:

- For data manipulation, we will use the Python library Pandas, which facilitates data manipulation when collecting the data before using it in the model. Similarly, the library Beautiful Soup allows us to analyze HTML documents.

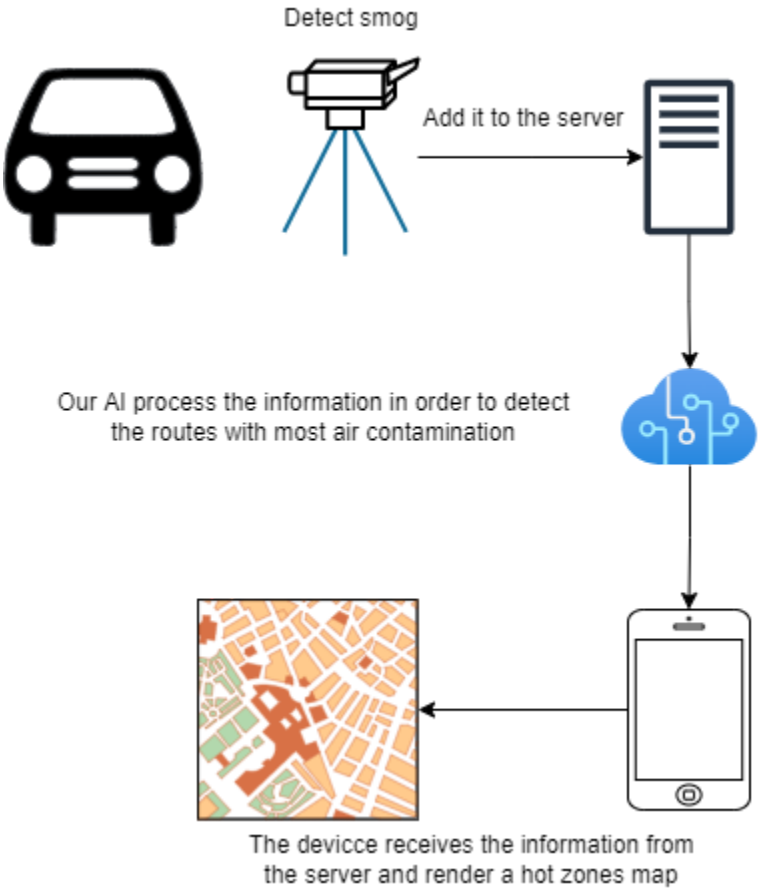
- Data set comprised of videos of smog (TBD)

Front-end

- SwiftUI or UIKit
- React and Next.js

Back-end:

- Prisma ORM (PlanetScale's MySQL)
- REST-based framework (FastAPI (Python) or Fastify (Javascript))



Our mobile app will contain two main tabs: one to scan for pollution or unfiltered exhaust on-the-go, and one with a GPS system rendering hot routes in the map. The smog detection tab will have our computer vision model actively running in real-time, classifying cars' exhaust as filtered or unfiltered. For as long as the user's phone is tracing a car(s) with unfiltered exhaust, it will trace a line along the GPS map in the mapping tab. The lines will get darker and darker as more polluting cars follow the same route because we will overlay the tracings. We will also make sure no two users are tracing the same car: potentially by reading the car's license plate (nice-to-have, not a must).

Ideal Outcome

By the end of the hackathon, we expect to have fully developed *Smogify*, as an app able to build a reliable map shown to the user friendly and clearly, with the routes and their corresponding pollution levels. The outcome will be, not only the app per se, but its usability as a public and accessible tool ready to be implemented in the daily operations of our users. In the long term, we project that *Smogify* works as a tool to provide crucial information for users to reduce traffic pollution and take care of their health.

Estimated Chance of Success

Based on the market research and the background technology available we estimate a chance of 8/10 of success

Bibliographic References

Levitt, S. D., & Dubner, S. J. (2018). *Superfreakonomics*. Debolsillo.

Figure 1. Taken from:

Ferguson, M., Ak, R., Lee, Y. T. T., & Law, K. H. (2017). Automatic localization of casting defects with convolutional neural networks. *2017 IEEE International Conference on Big Data (Big Data)*. https://doi.org/10.1109/bigdata.2017.8258115