Using Audio Processing To Determine Disease Spread

Ajit Narayanan

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Narayanan, Ajit, "Using Audio Processing To Determine Disease Spread", Technical Disclosure Commons, (July 16, 2020)
https://www.tdcommons.org/dpubs_series/3433

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.
ABSTRACT

During epidemics and at other times, it is important for public health officials, individuals, and other entities to make health-related decisions based on sound epidemiological statistics. Such statistics are today generated only when individuals voluntarily seek medical treatment. Diseases that are contagious but initially asymptomatic can spread rapidly through vulnerable populations with little warning. Due to the initial mildness of symptoms, individuals may not realize that they have a dangerous infection and may fail to approach medical professionals for treatment or diagnosis.

This disclosure describes techniques that can leverage sensors, networks, and consumer electronic devices to determine a fine-grained estimate of epidemiological statistics. With user permission, ambient sounds can be analyzed to detect audio symptoms of disease, e.g., cough, changes in voice, etc.; to classify the symptom by disease; and to geographically aggregate disease detections to automatically construct real-time epidemiological maps. Such epidemiological maps can enable accurate public-health decisions and efficient healthcare delivery.

KEYWORDS

- Epidemiology
- Disease tracking
- Public health
- Asymptomatic disease
- Audible symptom
BACKGROUND

During epidemics and at other times, it is important for public health officials, individuals, and other entities to make health-related decisions based on sound epidemiological statistics. Such statistics are today generated only when individuals voluntarily seek medical treatment. Diseases that are contagious but initially asymptomatic (or mildly symptomatic), e.g., the coronavirus disease of 2019, can spread rapidly through vulnerable populations with little warning. Due to the initial mildness of symptoms, patients may not realize that they have a dangerous infection, and may fail to approach medical professionals for treatment or diagnosis.

This problem is compounded by the fact that even if a significant proportion of individuals get tested when they exhibit mild symptoms, the medical infrastructure in a geographical area may not be equipped to conduct large-scale diagnoses. For example, the number of cases requiring testing may overwhelm the number of doctors or nurses that are present in the area, or there may be a shortage of test kits.

In situations where limited medical resources are to be deployed over a widespread area, it is important to direct resources to the most severely affected areas. Similarly, where there are economically impactful recommendations put in place to control the spread of a disease (such as the closing of educational institutions or workplaces, or the requiring individuals to socially distance themselves), there is a substantial societal benefit to micro-targeting such recommendations only to geographies at risk.

Existing techniques of gathering epidemiological data generally have touch points throughout the medical care system. For example, epidemiological maps are constructed from reports generated by doctors or hospitals. Analytics from drug sales and other medical supplies and equipment can serve as epidemiological indicators.
Although applications exist on consumer devices to monitor health, e.g., smartwatches that can detect atrial fibrillation, these generally do not provide epidemiological data for communicable diseases. Some experiments have utilized headphone and environmental sound exposure data to determine the effect of sound on hearing, stress levels, and heart health. Comprehensive expert systems have been built to assist in making health decisions. However, these relate to health, not to the epidemiological mapping of communicable diseases.

DESCRIPTION

This disclosure describes techniques to determine a fine-grained estimate of the prevalence of a medical condition even when the at-risk populations are not formally medically diagnosed. The techniques leverage the nearly ubiquitous presence of sensors, networks, and consumer electronic devices to determine epidemiological statistics. The techniques include one or more of the following procedures, as permitted by users, each described in greater detail below.

- **Disease detection from background audio**, e.g., the use of audio sensors to pick up auditory signals, e.g., coughing, that are potentially indicative of disease.

- **Obtaining medically relevant auditory signals from different geographic locations**, e.g., by incorporating audio-based disease detectors into consumer electronic devices or applications, with the permission of the user.

- **The use of machine learning to detect diseases from audio symptoms**, e.g., by categorizing audio signals into underlying medical conditions.

The described techniques can create an epidemiological map of the incidence of a disease. The epidemiological map can be used to prioritize medical intervention and to provide fine-grained
recommendations for social distancing, closure of businesses, etc. in areas where they are needed while leaving low-risk or no-risk areas unaffected.

Disease detection from background audio

Many diseases, e.g., respiratory diseases, have auditory symptoms. Example auditory symptoms include coughing, sneezing, hoarse voices, a higher incidence of clearing one's throat, an increased frequency of blowing one's nose, etc. It is pertinent to note that the diseases that have these symptoms are often the most contagious ones, since they may be spread by droplets, cough, phlegm, etc.

With user permission, smart devices that include sensors are activated to detect such auditory symptoms. For example, such sensors work in a manner similar to wake-word detection in smart speakers and other smart devices, e.g., with coughs and other disease-symptomatic auditory signals being detected in a manner similar to detection of wake-words. The various techniques to detect coughs or other disease-symptomatic sounds are hereinafter referred to as cough detectors.

A cough detector can be constructed, for example, by using a corpus of cough and similar disease-symptomatic sounds to train a machine learning model (classifier) to detect disease-symptomatic sounds in audio. The machine learning model can be, for example, a recurrent or other type of neural network. Since symptomatic audio signals are situated in voice frequencies, such a neural network can be implemented by adding a head layer to a neural network that is used for automatic speech recognition.

Alternatively, or in addition, audio-processing techniques such as spectrograms or short-term Fourier Transform (STFT) can be used to create audio signatures of coughs or other disease-symptomatic sounds. An audio signature can be generated by subjecting windowed audio
signals to STFT and by analyzing them for positive and negative instances of coughing. Once generated, an audio signature can optimally separate audio that includes a cough from that does not. Depending on a risk level associated with the disease, the threshold for classification on precision-recall curve of the detector that is based on the signature can be tuned: for highly contagious diseases, or for diseases that have a high risk of becoming pandemics or overwhelming medical infrastructure, the threshold on the curve can be chosen to allow a higher rate of false negatives while reducing the number of false positives.

Obtaining medically relevant auditory signals from different geographies

To obtain epidemiologically useful signals for a population at large, cough detectors can be incorporated into consumer electronics and applications that provide audio interfaces to the user, with user permission. Some examples of incorporating cough detectors include:

- **Smart speakers or other smart devices**: Smart speakers and other smart devices generally include an always-on microphone to listen to a user-spoken wake word. Additionally, some smart devices also have onboard global positioning systems or other location sensors that can identify their locations. Thus, such devices can be utilized to implement a cough detector to detect symptoms and construct geographical distributions of incidence of contagious diseases.

- **Smartphones, smartwatches, and smart assistant applications therein**: Many smartphones today have always-on microphones that can detect ambient sounds (such as songs) or can be used to invoke virtual assistants. Techniques for ambient sound detection or assistant invocation can be extended to disease-symptomatic sound detection.

- **Recording of audio by various devices**: With user permission, audio from smartphones and other devices, conference-calling facilities, landlines, etc. can be recorded in certain
contexts, e.g., during conference calls or video meetings. With user permission, a cough
detector can be incorporated into the audio pipeline of such contexts to identify disease-
symptomatic sounds from participants. Advantageously, incorporating a cough detector
into the audio pipeline in a phone call in this manner can detect disease-symptomatic
sounds at either end of the phone call (or any participant in a multi-party conference call).
Even participants that connect via landline telephones or other devices that do not have
cough detectors can thereby be assessed for infection, if they permit recording and
analysis of their audio. With user permission, information pertaining to diseases can be
combined with call metadata, e.g., carrier information, user location, caller ID, etc., to
provide data points that link symptoms to geographical areas.

The cough detector can operate entirely on a user’s device, or, with user permission,
operate partially or completely on a server. In case the cough detector operates entirely out of the
user’s device, cough detection and classification as disease symptomatic (or not) is performed
on-device, e.g., without sending data to another device or server. On-device cough detection and
disease classification can be done by pruning and quantizing a neural network; by using special-
purpose, low-energy consumption, machine-learning ASICs for such processing; etc.

Training of the cough detector (machine learning classifier) can be distributed across
multiple devices, e.g., using federated learning. Such training data, labeled positive or negative,
has no personally identifiable information.

If the task of cough detection/classification in a particular geographical area is performed
partially or completely at a server, data can be sent from client to server in a privacy-preserving
form, using, e.g., differential privacy techniques. Under differential privacy, patterns of groups
within a dataset may be shared publicly while information relating to individuals in the dataset is
withheld. Differential privacy techniques can be used for publicly aggregating geographical information about coughs and other symptoms, thus allowing determination of the geographical distribution of a disease without access to specific information relating to symptoms of individuals. In this manner, an epidemiological map, e.g., a map of disease incidence with associated geographic locations, can be constructed in a privacy-preserving manner.

Machine learning to detect diseases from audio symptoms

There may be multiple diseases or conditions that have similar audio symptoms, e.g., coughs, changes in voice hoarseness, etc., some of which may be of lower concern than others. For example, an individual may cough due to a low-risk condition such as common cold; due to a life-threatening condition such as tuberculosis; due to a condition that may be low-risk for the individual but high-risk for society such as coronavirus disease; etc. An individual may also cough for reasons that are unrelated to disease, e.g., due to food going down the trachea (fluid aspiration).

Fig. 1: Training a machine learning model to identify diseases based on audio symptoms

Per the techniques disclosed herein, illustrated in Fig. 1, a machine learning model can be trained on coughs and other audio symptoms labeled with their underlying causes. During
operation, the machine learning model assigns scores to different medical or non-medical conditions for a given audio symptom. Such a machine learning model can be, e.g., a convolutional or recurrent neural network, and can be trained in various ways, e.g., using the spectrogram of a signal, the mel-frequency cepstrum coefficients (MFCCs) of the signal, etc. Epidemiological information obtained from audio signals can be augmented with the potential probabilities corresponding to different diseases.

When the user permits, the machine learning model can be personalized by training with a user's specific audio, such that the model learns to classify a particular user's coughs into different underlying conditions. The model can be further improved by combining, with user permission, the audio signal with various demographic and health data about the user, that includes factors such as age, gender, past record of respiratory or infectious diseases, location history, history of close contacts with others (contact tracing), etc. Such personalization can be done entirely on the user's device. The probability distribution across different underlying conditions can then be made user-specific rather than symptom-specific. In addition to a single cough instance, a pattern of coughs for a single user can be analyzed to obtain a more accurate signal of a potential health condition. For example, a persistent cough can be distinguished from a one-off cough, or a wet cough from a dry one. Other environmental signals, e.g., weather or smog conditions, can be used to improve the accuracy of this predictor.

By automatically constructing real-time, epidemiological maps of diseases, the described techniques can enable accurate public-health decisions and efficient healthcare delivery. For example, based on the constructed epidemiological maps, medical resources can be allocated to the most severely affected areas, and economically impactful measures such as lockdowns and travel restrictions can be micro-targeted. In addition to generating epidemiological data, to
provide prompt medical care to the user, the user and their medical provider can be informed (with user permission) of any audio symptoms that are indicative of a medical condition.

Further to the descriptions above, a user is provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable the collection of user information (e.g., information about a user’s speech or other audio, user’s social network, social actions or activities, demographic information, a user’s preferences, or a user’s current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user’s identity is treated so that no personally identifiable information can be determined for the user, or a user’s geographic location is generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques that can leverage sensors, networks, and consumer electronic devices to determine a fine-grained estimate of epidemiological statistics. With user permission, ambient sounds can be analyzed to detect audio symptoms of disease, e.g., cough, changes in voice, etc.; to classify the symptom by disease; and to geographically aggregate disease detections to automatically construct real-time epidemiological maps. Such epidemiological maps can enable accurate public-health decisions and efficient healthcare delivery.
REFERENCES


