Genetic Disorder Detection

Diagnosing genetic disorders using facial images

The Alan Turing Institute

About

The Turing Commons is an online platform to support open dialogue and reflection about the responsible design, development, and deployment of data-driven technologies.

This case study was developed by the following students at the Biomedical AI CDT in the University of Edinburgh.

Authors

Nikitas Angeletos Chrysaitis
Passara Chanchotisatien
Ella Davyson
Achille Fraisse
Thibaut Goldsborough
Evgenii Lobzaev
Sebestyén Kamp
Charlotte Merzbacher
Scott Pirrie
Wolf de Wulf

Licensing Information

Download this case study and review others on the Turing Commons case study repository.

Turing Commons

THE UNIVERSITY of EDINBURGH
UKRI Centre for Doctoral Training in Biomedical Artificial Intelligence

Biomedical AI
UKRI CENTRE FOR DOCTORAL TRAINING

Overview

Many genetic disorders are known to be associated to facial features. Recent developments in image analysis have enabled improved facial recognition and analysis software, which can be used to detect genetic disorders from facial features in images taken of a patient. This case study raises issues of algorithmic bias, genetic counselling, and acceptable false-positive rates.

Project Description

This genetic disorder diagnostic app will run on a smartphone and give feedback, in the form of an assessment of the likelihood of the presence of various genetic conditions, to the user based on an analysis of supplied facial photographs. The app will be available to the general public and a user can upload any photograph they wish (of themselves or someone else).

The ultimate goal of the project is to create an accurate and user-friendly mobile application that can be used by healthcare professionals and individuals alike to detect genetic disorders through facial recognition technology. By providing an accessible and affordable means of screening for genetic disorders, the application has the potential to significantly improve early diagnosis and intervention, leading to improved patient outcomes and quality of life.

However, several ethical issues arise. The use of facial recognition technology raises concerns about privacy, especially given the context of use. The project team must also be mindful of potential biases in the dataset and the machine learning model, which may result in inaccurate or unfair diagnoses. Additionally, it is important to consider the potential for harm. While early detection of genetic disorders can lead to better patient outcomes and quality of life, there is also the possibility of negative psychological effects resulting from a diagnosis, as well as potential negative implications for the social and cultural attitudes towards individuals with genetic disorders.

Key Consideration

- **Bias and discrimination**
- **Lack of data across different sub-groups, especially for rare conditions**
- **Feasibility of incorporating multiple disorders into the model**
- **Transparency and interpretability of the model's decisions**
- **Accountability for misdiagnosis**
- **Privacy and unauthorised photo submissions**

Deliberative prompts

1. **How will the collection of data and the model training be performed? How will representativeness across different sub-groups be ensured?**
2. **What level of precision is deemed appropriate for the system?**
3. **How can this algorithm be unbiased, inclusive, and privacy-preserving?**
4. **What is the most effective and ethical way to present the results to users? What are the potential risks of self-diagnosis?**
5. **What safeguards could be put in place to prevent the misuse of this technology? To prevent stigmatisation of individuals based on their genetic makeup?**

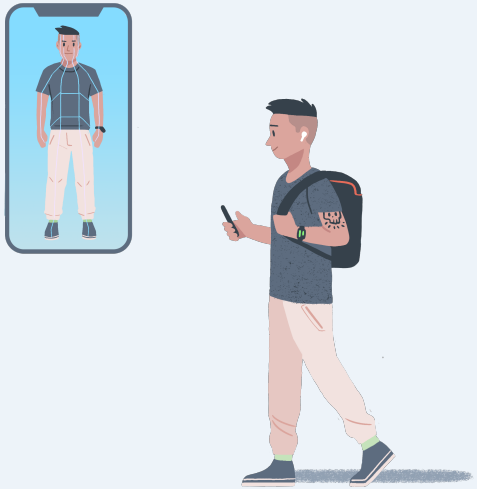
Technology Description

The model will be trained on facial images of individuals with and without the genetic disorders being classified. The model will be optimised to accurately detect the relevant facial features in a user's photo and provide a diagnosis of the potential genetic disorder.

To evaluate the accuracy of the application, the project team will conduct a series of tests using both real and synthetic data. The results of these tests will be used to refine the machine learning model and improve the performance of the application. This process will be quantified using precision and recall measures and possibly associated confidence measures.

Datasheet

Category	Details
Available Data	• Positive genetic test results for a variety of conditions • Datasets of images of people with and without each condition • Prevalence rates of each condition in general population (used to calibrate model)
Analysis Techniques	• Deep neural networks such as Convolutional Neural Networks (CNN), Visual Geometry Group (VGG), Residual Networks (ResNet), Inception, Densely Connected Convolutional Neural Networks and more. • Choice of implementing one large model for all conditions vs. separate models (one per condition studied)



Groups, Organisations and Affected Individuals

1. **Patients and users**
2. **Healthcare professionals**
3. **Researchers and developers**
4. **Regulator bodies and policy makers**
5. **Civil liberties organisations and general public**