VANGUARD: An Opensource Approach to Predict Cancer Generation and Evolution by the Use of Nanoparticle SERS, Probabilistic Analysis, and Collaborative Data Processing

Romero Hernandez, Joel (School: Institut F.X. Lluch i Rafecas)

Cancer is one of the most important problems for public health. This project tries to explore a new way to reduce its danger: VANGUARD is a multidisciplinary platform for the generation of universal preventive diagnostics by combining the SERS technic (to detect single-nucleotide mutations in short strings using Raman spectra) with pattern recognizing, collaborative data extraction mechanisms and the design of a probabilistic calculation structure, all in an opensource approach. The first phase consisted in obtaining basic research data to work with. I decided to study a commonly mutated fragment of the DNA keeper TP53 gene (codons 243 to 251) related to the most dangerous cancers: lung, pancreatic and colorectal. Four synthetic copies of this fragment were combined with a simple gold substrate and subjected to Raman microscopy, obtaining 7 samples (1 wild, and 2 for each cancer): using only half of the references needed, and making comparations by eye to reduce cost, samples were recognised properly 83.3% of tests, and the percentage composition of the DNA strings was approximated with an average error of 2.23%. The second phase was developed at home, and it consisted in approaching a diagnostic platform divided into three specific frameworks: data extraction and processing (LIBRA), probability mapping and clustering (ATLAS), and sample identification (ARGOS). After creating a sample recognition software based on example-learning (Retina), thousands of procedurally samples were analyzed, giving an accuracy from 90 to 100%. With web crawlers and a collaborative work app I designed, I used data about dozens of genes to design a calculation to know the cancer potential risk, implemented in scalable approaches for Arduino, Windows, Linux, and Android.