TITLE OF INVENTION: [DEEP LEARNING-BASED VARIANT CLASSIFIER]

ABSTRACT

The technology disclosed directly operates on sequencing data and derives its own feature filters. It processes a plurality of aligned reads that span a target base position. It combines elegant encoding of the reads with a lightweight analysis to produce good recall and precision using lightweight hardware. For instance, one million training examples of target base variant sites with 50 to 100 reads each can be trained on a single GPU card in less than 10 hours with good recall and precision. A single GPU card is desirable because it a computer with a single GPU is inexpensive, almost universally within reach for users looking at genetic data. It is readily available on could-based platforms.

[FIG. 1]