[0010] Convolutional architectures, such as dilated convolutional neural networks implemented in embodiments of the present invention described herein, may perform well in genetic sequence modeling tasks and outperform recurrent networks, and reach state-of-the-art accuracy in a broad range of sequence modeling tasks. The training and inference of convolutional neural networks is much faster than for recurrent networks such as long short term memory (LSTM) networks. Dilated convolutional neural networks in particular, may achieve an exponentially large receptive field with fewer parameters and fewer layers.

[0011] A method of automatically ~~basecalling~~ sequencing one or more DNA (deoxyribonucleic acid) molecules of a biological sample is described. The method comprises ~~converting a plurality of fluorescent signals of~~measuring the biological sample, ~~wherein each of the plurality of fluorescent signals is measured by~~using a capillary electrophoresis genetic analyzer, to obtain an ~~at least one~~ input trace comprising digital data corresponding to fluorescent values ~~at~~ comprising a plurality of scans of the biological sample. Scan labelling probabilities for ~~each of~~ the plurality of scans are generated. The scan labelling probabilities are generated using a trained ~~deep~~ artificial neural network comprising a plurality of layers including convolutional layers. A basecall sequence comprising a plurality of basecalls for the one or more DNA molecules based on the ~~one or more~~ scan labelling probabilities for each of the plurality of scans is determined.

[0011a] A non-transitory computer readable medium is also described. The medium comprises a memory storing one or more instructions which, when executed by a one or more processors of at least one computing device, perform automatically sequencing one or more deoxyribonucleic acid (DNA) molecules of a biological sample by obtaining an input trace comprising digital data corresponding to fluorescent values in a plurality of scans of the biological sample conducted by a capillary electrophoresis genetic analyzer; using a trained artificial neural network comprising a plurality of layers including convolutional layers to generate scan labelling probabilities for the plurality of scans; and determining a basecall sequence comprising a plurality of basecalls

for the one or more DNA molecules based on the scan labelling probabilities for each of the plurality of scans.

[0011b] A method of determining scan positions of basecalls in a basecall sequence obtained using scan labelling probabilities for each of a plurality of scans conducted by a capillary electrophoresis instrument on one or more deoxyribonucleic acid (DNA) molecules of a biological sample is also described. The method comprises determining a scan range of the scan range probabilities for each basecall, the scan range comprising scans from a range starting from a first scan corresponding to the basecall up to and including a last scan corresponding to the basecall; searching within the scan range to determine a peak scan labelling probability in the scan range; and using a scan position of the peak scan labelling probability to display, on an electronic display, an indication of the basecall position.

[0012] In some embodiments, a basecall position for each basecall in a basecall sequence is also determined, where the basecall position corresponds to a scan position of a peak scan labelling probability associated with the basecall. In some embodiments, searching for the peak probability associated with a given basecall is made more efficient by first identifying a first and last scan in a scan range of scans corresponding to scan labelling probabilities associated with the given basecall and then only searching within that scan range.