CNN models – how their structures, algorithms and training datasets affect their performance?

1 Trained 108 CNN models of different structures, algorithms and datasets

2 Tested their accuracy levels with a target dataset of 3,000 images with 256 signal sample rate and 20 units of the time lag parameter

- The evaluation metric used is the marco-F1 score
- Equation = True Positives/ (True Positives + 0.5 (False Positives + False Negatives))
- Each category of the 10 PQD types is considered equally important in this calculation

Table 1: Macro-F1 score of each of the trained CNN models

Datasets		C) CNN structures					
A) Signal sample rate	B) Time lag parameter	Fundamental structure	Increased kernel size	Increased 1 layer	Increased 2 layers	Used max pooling	Reversed order of layers
100	10	57.0%	46.0%	52.2%	58.0%	24.2%	49.3%
	20	15.4%	11.0%	9.5%	22.9%	18.6%	11.7%
	30	14.4%	9.1%	13.2%	10.1%	7.3%	8.9%
200	10	32.1%	32.4%	1.9%	56.6%	32.5%	31.0%
	20	50.8%	50.7%	50.5%	62.6%	53.8%	45.1%
	30	15.4%	16.9%	20.4%	24.4%	12.6%	15.1%
256	10	22.5%	30.0%	25.2%	34.5%	20.3%	25.2%
	20	96.4%	97.8%	97.7%	97.8%	97.1%	93.8%
	30	42.0%	36.9%	37.8%	49.2%	24.0%	29.5%
300	10	20.2%	24.3%	24.7%	23.9%	20.8%	21.7%
	20	67.5%	71.0%	74.6%	77.1%	72.5%	57.0%
	30	62.3%	51.5%	58.7%	60.6%	2.2%	49.4%
400	10	24.0%	16.8%	12.4%	26.3%	20.0%	15.7%
	20	32.2%	38.7%	45.8%	56.2%	34.8%	33.1%
	30	96.8%	96.7%	96.7%	96.9%	95.4%	94.2%
500	10	16.8%	14.0%	17.3%	23.0%	18.1%	19.2%
	20	22.6%	31.7%	31.6%	34.0%	20.6%	23.2%
	30	46.3%	45.9%	61.4%	71.5%	49.9%	59.7%

Source: Team analysis