Appendix A: Different architectures considered in this work

The metric scores for the six additional models considered are presented below. These models were part of the initial trials, and the parameters used are outlined below. Modifications to these parameters were made to improve the final DNN model discussed in the main text. The following results provide insight into the challenges faced during model development and the rationale behind further optimization efforts. The architectures of the six models are listed in Table V.

TABLE V: Model list and their corresponding architecture

Model	Architecture
1	$[100 - 200]$
2	$[200 - 100]$
3	$[200 - 200]$
4	$[100 - 200 - 300]$
5	$[300 - 200 - 100]$
6	$[300 - 300 - 300]$

All models were trained using the `SMORMS3` optimizer with a batch size of 64 for $2,000$ epochs. The training dataset for these models was smaller than that used for the final model, comprising $2,500$ instances of each pole configuration for each curriculum. A split of 80% for training and 20% for testing was used.

The curriculum setup for these models differs from that shown in Table II. Only three curricula were used, as detailed in Table VI.

TABLE VI: Curricula arrangement of the eight classes.

Curr. label	Included classes	Config. presented
1	Classes $0, 1, 2$	At most 1 pole
2	Classes $0, 1, 2, 3, 4$	At most 2 poles
3	All classes	At most 3 poles

Each model was trained for 300 epochs on curriculum 1, followed by another 300 epochs on curriculum 2, and a final $1,400$ epochs on curriculum 3, for a total of $2,000$ epochs. Table VII summarizes the training and testing accuracy scores.

TABLE VII: Training and testing accuracy of each model.

Model	Training accu.	Testing accu.
1	83%	79%
2	83%	79%
3	87%	81%
4	79%	75%
5	88%	82%
6	89%	83%

The validation dataset for these models, like the training dataset, consisted of $2,500$ instances per class. Precision and recall scores for each class of each model (M1, M2, etc.) are provided in Table VIII. The accuracy and macro F1 scores are summarized in Table IX.

TABLE VIII: Classwise performance of our models using a validation dataset.

M1	Precision	Recall
0	0.23	0.00
1	0.83	0.92
2	0.79	0.88
3	0.32	0.49
4	0.53	0.52
5	0.60	0.72
6	0.56	0.56
7	0.56	0.6

M2	Precision	Recall
0	0.89	0.41
1	0.84	0.94
2	0.89	0.70
3	0.39	0.50
4	0.48	0.52
5	0.63	0.71
6	0.54	0.53
7	0.56	0.63

M3	Precision	Recall
0	0.13	0.00
1	0.93	0.85
2	0.92	0.62
3	0.30	0.54
4	0.52	0.50
5	0.57	0.74
6	0.50	0.58
7	0.54	0.62

M4	Precision	Recall
0	0.90	0.21
1	0.82	0.94
2	0.82	0.83
3	0.32	0.47
4	0.47	0.46
5	0.64	0.69
6	0.54	0.48
7	0.54	0.65

M5	Precision	Recall
0	0.98	0.27
1	0.97	0.98
2	0.81	0.49
3	0.39	0.60
4	0.44	0.60
5	0.64	0.77
6	0.58	0.57
7	0.60	0.60

M6	Precision	Recall
0	0.73	0.07
1	0.94	0.81
2	0.94	0.50
3	0.34	0.64
4	0.44	0.59
5	0.64	0.74
6	0.52	0.56
7	0.61	0.65

TABLE IX: Average performance metrics versus validation dataset across models.

Model	Accu.	Macro F1
1	58.57%	0.57
2	61.92%	0.63
3	55.85%	0.55
4	58.99%	0.61
5	61.01%	0.64
6	57.02%	0.61