@article{abdel2014convolutional,

title={Convolutional neural networks for speech recognition},

author={Abdel-Hamid, Ossama and Mohamed, Abdel-Rahman and Jiang, Hui and Deng, Li and Penn, Gerald and Yu, Dong},

journal={IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)},

volume={22},

number={10},

pages={1533--1545},

year={2014},

publisher={IEEE Press}

}

@article{lecun1995convolutional,

title={Convolutional networks for images, speech, and time series},

author={LeCun, Yann and Bengio, Yoshua},

journal={The handbook of brain theory and neural networks},

volume={3361},

pages={310},

year={1995},

publisher={Cambridge, MA: MIT Press}

}

@inproceedings{ciresan2011convolutional,

title={Convolutional neural network committees for handwritten character classification},

author={Ciresan, Dan Claudiu and Meier, Ueli and Gambardella, Luca Maria and Schmidhuber, J{\"u}rgen},

booktitle={Document Analysis and Recognition (ICDAR), 2011 International Conference on},

pages={1135--1139},

year={2011},

organization={IEEE}

}

@inproceedings{lecun1994word,

title={Word-level training of a handwritten word recognizer based on convolutional neural networks},

author={LeCun, Yann and Bengio, Yoshua},

booktitle={International Conference on Pattern Recognition},

pages={88--88},

year={1994},

organization={Citeseer}

}

@article{lauer2007trainable,

title={A trainable feature extractor for handwritten digit recognition},

author={Lauer, Fabien and Suen, Ching Y and Bloch, G{\'e}rard},

journal={Pattern Recognition},

volume={40},

number={6},

pages={1816--1824},

year={2007},

publisher={Elsevier}

}

@article{baldi1996hybrid,

title={Hybrid modeling, HMM/NN architectures, and protein applications},

author={Baldi, Pierre and Chauvin, Yves},

journal={Neural Computation},

volume={8},

number={7},

pages={1541--1565},

year={1996},

publisher={MIT Press}

}

@article{zhu2014stock,

title={A stock decision support system based on DBNs},

author={Zhu, Chengzhang and Yin, Jianping and Li, Qian},

journal={Journal of Computational Information Systems},

volume={10},

number={2},

pages={883--893},

year={2014}

}

@article{hochreiter2007fast,

title={Fast model-based protein homology detection without alignment},

author={Hochreiter, Sepp and Heusel, Martin and Obermayer, Klaus},

journal={Bioinformatics},

volume={23},

number={14},

pages={1728--1736},

year={2007},

publisher={Oxford Univ Press}

}

@article{eck2002first,

title={A first look at music composition using lstm recurrent neural networks},

author={Eck, Douglas and Schmidhuber, Juergen},

journal={Istituto Dalle Molle Di Studi Sull Intelligenza Artificiale},

year={2002}

}

@article{hochreiter1991untersuchungen,

title={Untersuchungen zu dynamischen neuronalen Netzen},

author={Hochreiter, Sepp},

journal={Master's thesis, Institut fur Informatik, Technische Universitat, Munchen},

year={1991}

}

@article{hochreiter1997long,

title={Long short-term memory},

author={Hochreiter, Sepp and Schmidhuber, J{\"u}rgen},

journal={Neural computation},

volume={9},

number={8},

pages={1735--1780},

year={1997},

publisher={MIT Press}

}

@incollection{beringer2005classifying,

title={Classifying unprompted speech by retraining LSTM nets},

author={Beringer, Nicole and Graves, Alex and Schiel, Florian and Schmidhuber, J{\"u}rgen},

booktitle={Artificial Neural Networks: Biological Inspirations--ICANN 2005},

pages={575--581},

year={2005},

publisher={Springer}

}

@inproceedings{sak2014long,

title={Long short-term memory recurrent neural network architectures for large scale acoustic modeling},

author={Sak, Hasim and Senior, Andrew and Beaufays, Fran{\c{c}}oise},

booktitle={Proceedings of the Annual Conference of International Speech Communication Association (INTERSPEECH)},

year={2014}

}

@inproceedings{graves2006connectionist,

title={Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks},

author={Graves, Alex and Fern{\'a}ndez, Santiago and Gomez, Faustino and Schmidhuber, J{\"u}rgen},

booktitle={Proceedings of the 23rd international conference on Machine learning},

pages={369--376},

year={2006},

organization={ACM}

}

@incollection{graves2012connectionist,

title={Connectionist Temporal Classification},

author={Graves, Alex},

booktitle={Supervised Sequence Labelling with Recurrent Neural Networks},

pages={61--93},

year={2012},

publisher={Springer}

}

@inproceedings{sutskever2014sequence,

title={Sequence to sequence learning with neural networks},

author={Sutskever, Ilya and Vinyals, Oriol and Le, Quoc VV},

booktitle={Advances in Neural Information Processing Systems},

pages={3104--3112},

year={2014}

}

@article{eddy1996hidden,

title={Hidden markov models},

author={Eddy, Sean R},

journal={Current opinion in structural biology},

volume={6},

number={3},

pages={361--365},

year={1996},

publisher={Elsevier}

}

@inproceedings{bluche2014a2ia,

title={The A2iA Arabic Handwritten Text Recognition System at the Open HaRT2013 Evaluation},

author={Bluche, Th{\'e}odore and Louradour, J{\'e}r{\^o}me and Knibbe, Maxime and Moysset, Bastien and Benzeghiba, Mohamed Faouzi and Kermorvant, Christopher},

booktitle={Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on},

pages={161--165},

year={2014},

organization={IEEE}

}

@article{pearlmutter1995gradient,

title={Gradient calculations for dynamic recurrent neural networks: A survey},

author={Pearlmutter, Barak A},

journal={Neural Networks, IEEE Transactions on},

volume={6},

number={5},

pages={1212--1228},

year={1995},

publisher={IEEE}

}

@article{siegelmann1995computational,

title={On the computational power of neural nets},

author={Siegelmann, Hava T and Sontag, Eduardo D},

journal={Journal of computer and system sciences},

volume={50},

number={1},

pages={132--150},

year={1995},

publisher={Elsevier}

}

@article{bengio1995input,

title={An input output HMM architecture},

author={Bengio, Yoshua and Frasconi, Paolo},

journal={Advances in neural information processing systems},

pages={427--434},

year={1995},

publisher={Morgan Kaufmann Publishers}

}

@article{williams1990efficient,

title={An efficient gradient-based algorithm for on-line training of recurrent network trajectories},

author={Williams, Ronald J and Peng, Jing},

journal={Neural computation},

volume={2},

number={4},

pages={490--501},

year={1990},

publisher={MIT Press}

}

@book{robinson1987utility,

title={The utility driven dynamic error propagation network},

author={Robinson, AJ and Fallside, Frank},

year={1987},

publisher={University of Cambridge Department of Engineering}

}

@article{williams1989learning,

title={A learning algorithm for continually running fully recurrent neural networks},

author={Williams, Ronald J and Zipser, David},

journal={Neural computation},

volume={1},

number={2},

pages={270--280},

year={1989},

publisher={MIT Press}

}

@article{schmidhuber1992fixed,

title={A fixed size storage O (n3) time complexity learning algorithm for fully recurrent continually running networks},

author={Schmidhuber, J{\"u}rgen},

journal={Neural Computation},

volume={4},

number={2},

pages={243--248},

year={1992},

publisher={MIT Press}

}

@article{williams1995gradient,

title={Gradient-based learning algorithms for recurrent networks and their computational complexity},

author={Williams, Ronald J and Zipser, David},

journal={Back-propagation: Theory, architectures and applications},

pages={433--486},

year={1995}

}

@article{bengio1994learning,

title={Learning long-term dependencies with gradient descent is difficult},

author={Bengio, Yoshua and Simard, Patrice and Frasconi, Paolo},

journal={Neural Networks, IEEE Transactions on},

volume={5},

number={2},

pages={157--166},

year={1994},

publisher={IEEE}

}

@article{gers2000learning,

title={Learning to forget: Continual prediction with LSTM},

author={Gers, Felix A and Schmidhuber, J{\"u}rgen and Cummins, Fred},

journal={Neural computation},

volume={12},

number={10},

pages={2451--2471},

year={2000},

publisher={MIT Press}

}

@misc{hochreiter2001gradient,

title={Gradient flow in recurrent nets: the difficulty of learning long-term dependencies},

author={Hochreiter, Sepp and Bengio, Yoshua and Frasconi, Paolo and Schmidhuber, J{\"u}rgen},

year={2001},

publisher={A field guide to dynamical recurrent neural networks. IEEE Press}

}

@inproceedings{bengio2013advances,

title={Advances in optimizing recurrent networks},

author={Bengio, Yoshua and Boulanger-Lewandowski, Nicolas and Pascanu, Razvan},

booktitle={Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on},

pages={8624--8628},

year={2013},

organization={IEEE}

}