Performance Monitoring of Centrifugal Compressor System using LSTM based Deep RNN

Harsh Purohit
Systems and Control
Engineering Department
Indian Institute of Technology Bombay
Mumbai, 400076
Email: harsh.purohit@sc.iitb.ac.in

Karmvir Phogat
Systems and Control
Engineering Department
Indian Institute of Technology Bombay
Mumbai, 400076
Email: karmvir.p@sc.iitb.ac.in

P.S.V.Nataraj
Systems and Control
Engineering Department
Indian Institute of Technology Bombay
Mumbai, 400076
Email: nataraj@sc.iitb.ac.in

***Abstract*—This work presents the results of applying an advanced performance monitoring technique to centrifugal compressor system using deep recurrent neural network (DRNN). In reality, due to different kind of disturbances, the compressor system may lead to catastrophic situations. Therefore, performance monitoring has become an issue of primary importance in modern process engineering automation. Detecting anomalies in such scenarios become challenging using standard statistical approaches. In this article, we discuss a Long Short Term Memory (LSTM) based DRNN technique to predict the faulty behavior of the compressor system. Due to the ability of LSTM to maintain memory, these networks have been proven effective for learning patterns of the time series data with unknown length. This motivates us to propose a performance monitoring schema based on LSTM-DRNN. To validate the proposed approach, we have simulated the compressor model in Simulink and trained the LSTM-DRNN model on the obtained time series data of the compressor system that is running under ideal conditions. Further, the trained network have been used to detect anomalies in the time series data that was generated by introducing disturbance as an inlet temperature changes.**

Keywords—Performance monitoring; LSTM-DRNN; anomaly detection; compressor control system

I. INTRODUCTION

Any automated system or process design is carried out in two important phases: the first phase consists of the design, tuning and implementation of control strategies, and the second phase consists of the supervision of the control loops and early detection of performance deterioration. The later tasks are performed within the frame work of performance monitoring which has got considerable attention from industrial and academic communities [1].

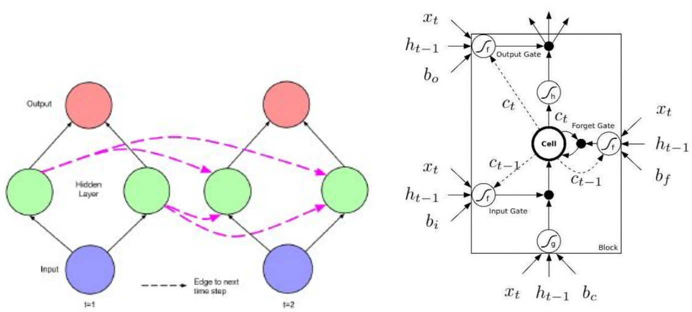
In the past few decades, human intuition has played a big role in the monitoring of control loops for evaluating the performance of the systems and the controllers. This manual monitoring approach is limited in its use because of the fact that a typical operator in the control room is responsible for the entire control system that consists of many control loops. Therefore, automatic performance monitoring techniques that have been developed in recent years are proven far superior than the conventional ones. Some model based conventional techniques used for anomaly detection include observer based approach [2] and parameter identification based methods [3]. On the other hand, if the model identification of a process is not possible under standard techniques then statistical methods are developed to extract the information about the fault from the measurement data [4].

In recent years, neural networks have been employed for performance monitoring applications (for instance see [5], [6]) and have been proven far superior than the classical techniques. Neural networks are valuable alternative to the classical methods because of two reasons: First, they can model a complex phenomenon provided one chooses sufficient number of neurons. Any non-linear function can be approximated with an arbitrary accuracy using proper neural network architecture. Second, training neural networks needs least amount of the process information. In the anomaly detection task, we will apply a special form of recurrent neural network that is know as Long Short Term Memory (LSTM) [7]. These LSTM are widely applied to many sequence learning tasks [8] because this architecture is more robust to training algorithms and it has memory units that are useful in sequence learning. In this work, we use stacked LSTM deep recurrent neural network for performance monitoring from sensor data. LSTM networks can be used to accurately detect deviation from the normal process behavior to a disturbed process behavior. To validate our approach, we needed time series data of the compressor system under normal operations to train the LSTM network, and a set of anomalous data under the influence of disturbance to detect the anomaly. To this end, we have developed a Simulink [9] model of centrifugal compressor system to generate nominal plant data as well as anomalous data under the effect of disturbances.

This article is organized as follows: Section II presents the basic LSTM neural network architecture and data preprocessing steps. Subsequently, Section III explains about the Simulink model of the centrifugal compressor system. Section IV demonstrates the numerical experiments and discuss the results. Finally, in Section V, we conclude the paper with a discussion on the future directions.

II. DEEP RNN AND LSTM

Recently, deep neural networks (DNN) have become very popular in computer vision [10] and speech processing related applications [11]. One type of network that falls into this category of deep networks is recurrent neural networks (RNN). In general, RNN's layers do not provide hierarchical processing



(a) (b)

Fig. 1. (a) The recurrent network unfolded across time steps. (b) LSTM memory cell.

but they actually introduce memory. RNNs have loops at each layers hence when folded out in time it can be considered as a DNN with indefinitely many layers (see Fig. 1(a)). One problem that arises from the unfolding of an RNNs is that the gradient of some of the weights starts to become too small or too large if the network is unfolded for too many time steps [12]. This phenomenon is known as vanishing gradient. Long Short Term Memory (LSTM) [13] neural network overcome the vanishing gradient problem experienced by RNNs. LSTM neural network employ multiplicative gates that enforce constant error flow through the internal states of special units called "memory cell" (see Fig. 1(b)). As shown in the Fig. 1(b), Input gate, output gate and forget gate do not allow memory contents from being perturbed by irrelevant inputs and outputs. Equation for gates are summarized as follows:

$$
\begin{aligned}
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i),\\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f),\\
c_t &= f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_f),\\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o),\\
h_t &= o_t \tanh(c_t).
\end{aligned}
$$

where x_t is input vector, h_t and h_{t-1} are output of current and previous block, c_t and c_{t-1} are memory from current and previous block, σ and $\tanh$ are sigmoid and hyperbolic tangent function, W and b are respective parameters which have to be trained.

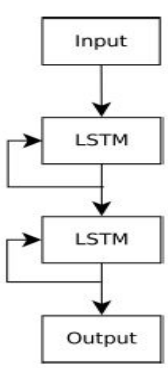


Fig. 2. LSTM network architecture.

1) LSTM neural network architecture: We consider the following LSTM network architecture in our work. We take one unit in the input layer and one unit for output layer. We have two hidden layers of size 50 and 50 respectively. We stack LSTM layer such that each unit in a lower layer LSTM hidden layer is fully connected to each unit in the LSTM hidden layer above it through feed-forward connection. This architecture is illustrated in Fig. 2.

A. Data Preprocessing

We train LSTM neural network model on the normal operating condition dataset. Before feeding this data into the network, we need to pre-process the data for a better accuracy and the fast convergence of the training algorithm. The preprocessing of data is done as follows: first the the overall dataset is divided into sequences with predefined length and then the training and target data is generated in such a way that a sequence of length n will have target value as $(n+1)^{th}$ term of that sequence. Further, to reduce the variance and unify the scale we applied normalization for all sequences as follows:

$$s_i^* = \frac{s_i - \bar{s}_i}{\sigma_i},$$

where $\bar{s}_i$ and σ_i are the mean value and the standard deviation for each sequence.

III. Centrifugal Compressor System

Compressor is a device that increases the pressure of gas by decreasing its volume mechanically. Oil and gas processing industries have several compressors operating either alone or in parallel. To maximize the profit, modeling and control of compressor system is primary concern. In literature, various models can be found that describes the working of centrifugal compressor system. For example, authors in [15] describes a mathematical model for the dynamics of the physical phenomena of compressor system. Depending on the application two important types of models can be considered. First is Control model which can be useful for development of controller. Second is more complex simulation model which can be used for more detailed system dynamic analysis.

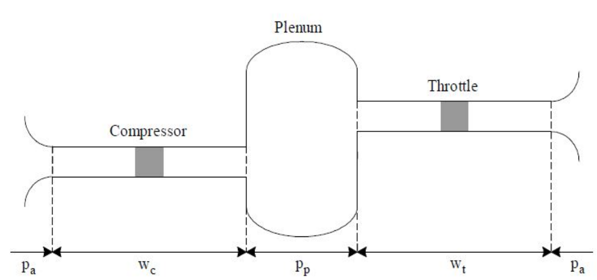


Fig. 4. Simple compressor system's component.

A compressor system for which the main components are a compressor, a plenum volume and a throttle is considered. Fig. 4 shows these components of the system. Dynamic equations

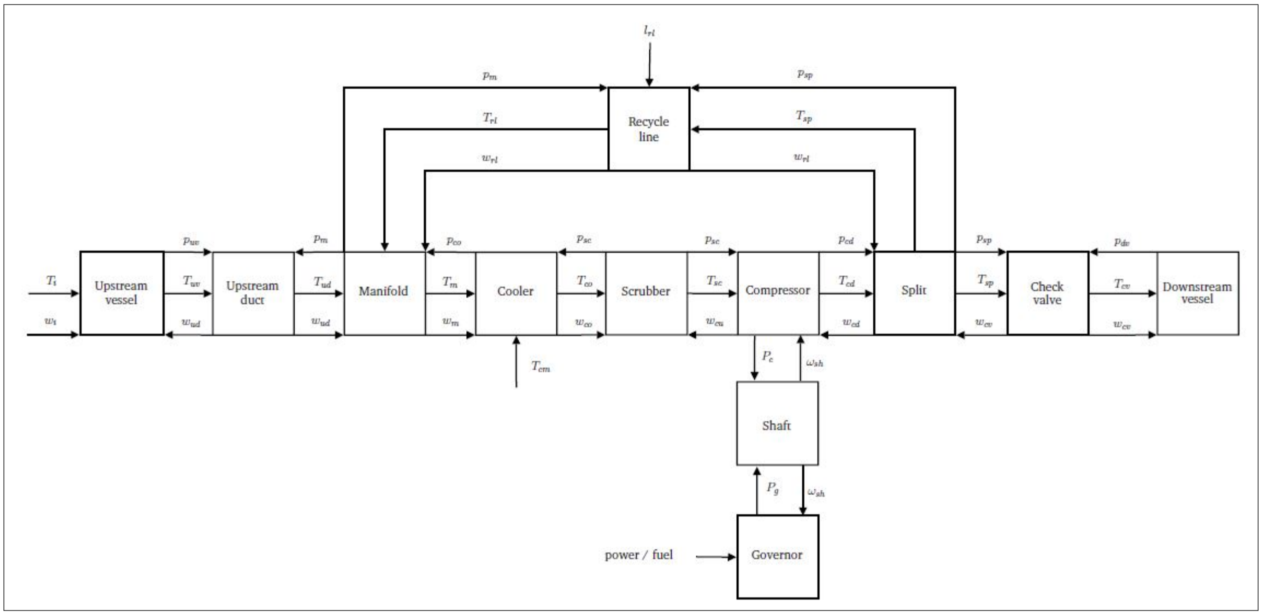


Fig. 3. Interconnection of compressor system modules [14].

for the plenum pressure, mass-flow dynamics and impeller speed dynamics are discussed below:

$$\dot{p_p} = \frac{C_p^2}{V_p}(\omega_c - \omega_t)$$

where p subscript refers plenum for pressure and volume. C_p is the speed of sound in plenum. ω_c and ω_t are mass flow entering the plenum and mass flow leaving the plenum, respectively.

$$\dot{\omega_c} = \frac{A_c}{L_c}(p_a - p_p + F_c)$$

Mass flow in the duct containing the throttling device is modeled in the same manner

$$\dot{\omega_t} = \frac{A_c}{L_c}(p_a - p_p - F_t)$$

where subscript c,t and a refers compressor, throttle and ambient condition. A_c and L_c are area and length of compressor. F_c is working along the flow direction whereas F_t is working against it.

$$\dot{\omega} = \frac{1}{J}(\tau_d - \tau_l)$$

where ω is the impeller speed, J is the moment of inertia of rotating parts. τ_d is torque applied to the impeller by the driving unit and τ_l is the load torque due to fluid flow and friction torque.

In addition to pressure and flow dynamics part, we can have static parts in the model for example, manifold and cooler. Considering this, the overall compressor system can be broadly divided into three categories.

- Part I contains pressure dynamics.
- Part II contains flow dynamics.
- Part III neither pressure nor flow dynamics (also called static module).

The overall compressor system with their interconnections are shown in Fig. 3. Readers can refer [14] for detailed modeling of each module.

A. *Control System*

From control point of view, surge in centrifugal compressor system is a very important phenomenon [16]. Compressor system will surge when forward flow through the compressor can no longer be maintained. This could be occur due to an increase pressure across the compressor and hence, momentary flow reversal occurs. To handle this situation, PID controller has been tuned which typically measures a function of pressure rise versus flow. On the other hand performance controller (i.e. also PID type) is tuned for desired pressure and velocity control in the compressor system. Although theses controllers are in action, disturbance through inlet flow or inlet temperature can affect performance substantially. In our experiment, we try to predict this kind of disturbance through sensor data using LSTM-DRNN. The model has been developed in the Simulink with anti surge and performance controller. The Simulink block diagram is shown in Fig. 5.

IV. Experiment

In the course of developing performance monitoring technique, general approach can be illustrated as in Fig. 6. We take sensor signal as an input signal which is used to train the deep recurrent neural network. Then residual can be generated through measured output signal and predicted output signal.

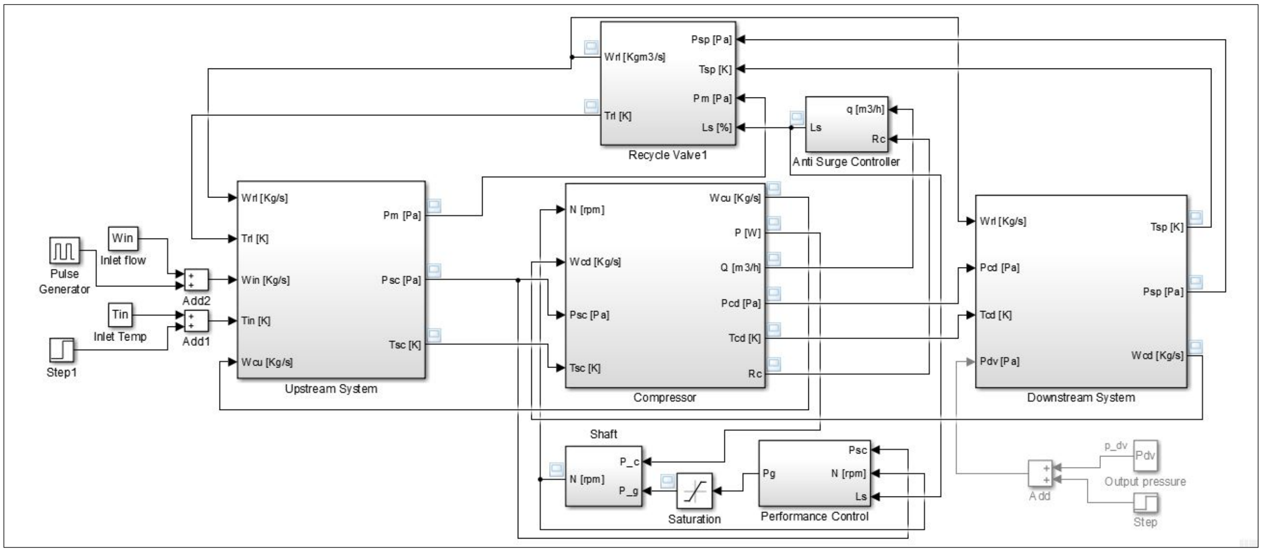


Fig. 5. Simulink block diagram for compressor control system with anti-surge and performance controller.

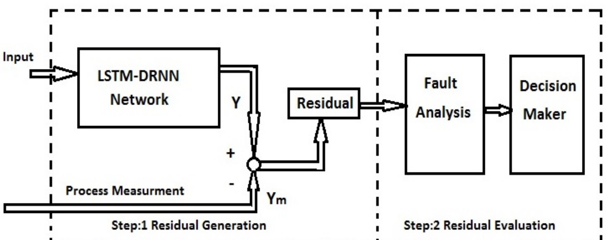


Fig. 6. General performance monitoring technique block diagram.

Further, residual evaluation can be useful for generating the alarm signal.

In our study, we measure the upstream pressure signal for 4500 seconds. Note that, this signal is collected during normal operating condition of centrifugal compressor control system. (i.e. without any disturbance). Fig. 7 illustrates measured pressure signal with normal process behavior. After collecting that, we do preprocessing on data (rearranging in train and target set and normalization) as discussed in Section II-A. Then we feed this processed data to LSTM-DRNN network for training. To train the network, the dropout technique [17] is used for regularization to reduce the over-fitting. The mean square error between training and predicted values is considered as loss function. “RMSprop” [18] optimization algorithm is used for training the network. This LSTM-DRNN neural network is developed with Keras library. [19] (tensorflow [20] backend)

After training, we test the network on disturbed operating condition signal. For example, inlet temperature changes at 1000 second may cause the disturbed operating condition. This disturbed operating condition signal is represented in Fig. 8. Predicted signal and square error (i.e. $(Y - Y_m)^2$) are represented in Fig. 9.

From Fig. 9 it can be observed that squared error is significantly higher than particular threshold value and therefore

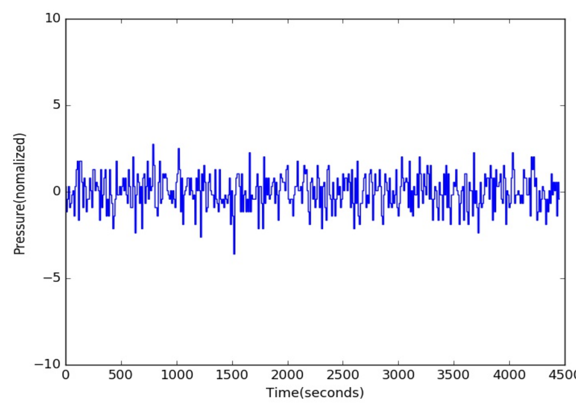


Fig. 7. Normalized upstream pressure signal in normal operating condition.

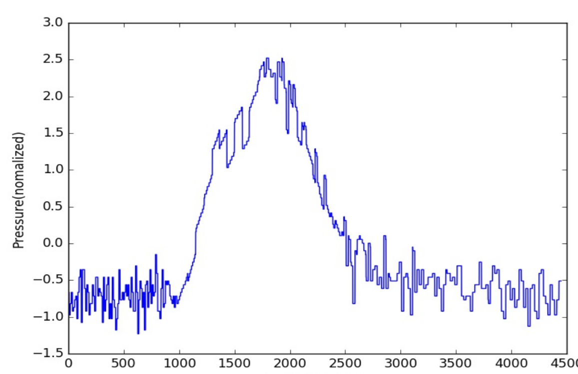


Fig. 8. Normalized upstream pressure signal in disturbed operating condition.

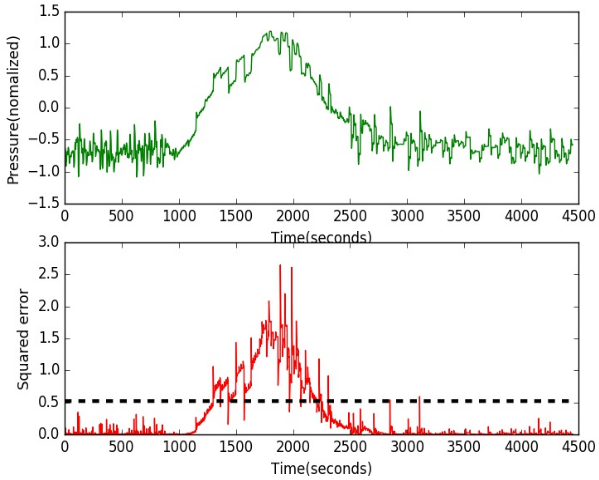


Fig. 9. Predicted signal (in green) and squared error plot with threshold.

the effect of disturbance can be identify in that operating condition. The error threshold level can be introduced as a tunable parameter that allows a user to achieve a satisfactory false positive and false negative detection rule.

V. Conclusion

In this work, we develop a performance monitoring technique for centrifugal compressor system using LSTM-DRNN based network. We developed the network architecture, data preprocessing and training rules for the LSTM-DRNN network. We have also discussed the centrifugal compressor control system. To test our proposed approach, we generated data with inlet temperature disturbance in Simulink model. The applied performance monitoring schema gives satisfactory result for detecting the disturbed process behavior. In future, multi-sensors data can be used to predict anomalous operating conditions and to generate prioritized alert system which does not restricts us to binary decisions.

Acknowledgement

Authors would like to thank the GPU center of excellence (GCOE) and OpenPOWER Research cluster of IIT Bombay to provide infrastructure support for this work.

References

[1] M. Bauer, A. Horch, L. Xie, M. Jelali, and N. Thornhill, "The current state of control loop performance monitoring–a survey of application in industry," *Journal of Process Control*, vol. 38, pp. 1–10, 2016.

[2] Q.-N. Xu, K.-M. Lee, H. Zhou, and H.-Y. Yang, "Model-based fault detection and isolation scheme for a rudder servo system," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2384–2396, 2015.

[3] Y. Wang and F. Ding, "Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model," *Automatica*, vol. 71, pp. 308–313, 2016.

[4] K. Zhang, "Performance assessment for process monitoring and fault detection methods," 2016.

[5] M. Mrugalski, M. Witczak, and J. Korbicz, "Confidence estimation of the multi-layer perceptron and its application in fault detection systems," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 6, pp. 895–906, 2008.

[6] Y. Cheng, R. Wang, and M. Xu, "A combined model-based and intelligent method for small fault detection and isolation of actuators," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2403–2413, 2016.

[7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[8] F. A. Gers, D. Eck, and J. Schmidhuber, "Applying LSTM to time series predictable through time-window approaches," in *Artificial Neural Networks – ICANN 2001 (Proceedings)*, G. Dorffner, Ed. Berlin: Springer, 2001, pp. 669–676.

[9] MATLAB, *version 7.10.0 (R2015a)*. Natick, Massachusetts: The MathWorks Inc., 2015.

[10] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[11] W. Chan, N. Jaitly, Q. Le, and O. Vinyals, "Listen, attend and spell: A neural network for large vocabulary conversational speech recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 4960–4964.

[12] Z. C. Lipton, "A critical review of recurrent neural networks for sequence learning," *CoRR*, vol. abs/1506.00019, 2015. [Online]. Available: http://arxiv.org/abs/1506.00019

[13] J. C. B. Gamboa, "Deep learning for time-series analysis," *arXiv preprint arXiv:1701.01887*, 2017.

[14] P. Aaslid, "Modelling of variable speed centrifugal compressors for anti-surge control," 2009.

[15] F. F. Willems, "Modeling and bounded feedback stabilization of centrifugal compressor surge," 2000.

[16] D.-C. Liaw, C.-C. Song, and J.-T. Huang, "Robust stabilization of a centrifugal compressor with spool dynamics," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 4. IEEE, 2002, pp. 3301–3306.

[17] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting." *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[18] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," *COURSERA: Neural networks for machine learning*, vol. 4, no. 2, 2012.

[19] F. Chollet, "Keras," https://github.com/fchollet/keras, 2015.

[20] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: http://tensorflow.org/