Deep temporal detection - A machine learning approach to multiple-dwell target detection

Daniel Gusland, Sigmund Rolfsjord, Børge Torvik
Norwegian Defence Research Establishment (FFI)
Kjeller, Norway
Email: Daniel.Gusland@ffi.no, Sigmund-Johannes-Ljosvoll.Rolfsjord@ffi.no,Borge.Torvik@ffi.no

***Abstract*—Detecting small targets, such as an Unmanned Aerial Vehicle (UAV) in high clutter and non-homogeneous environments is challenging for a radar system. Traditional Constant False Alarm Rate (CFAR) detectors have suboptimal performance in many scenarios. In this paper, we attempt a new approach to radar detection, based on machine learning, to increase the P_D while retaining a low P_{FA}. We propose two approaches, using a Convolutional Neural Network (CNN) on the range-Doppler images and stacking multiple range-Doppler images as layers, called the Temporal CNN detector. The models are trained and tested solely on measured radar data by using the estimated position and velocity from a collaborative target UAV. It is shown that training a model based solely on measured data is achievable and performance metrics calculated from the testing data shows that both models outperform the Cell-Averaging Constant False Alarm Rate (CA-CFAR) by having higher P_D with the same P_{FA} The current test results indicate that the temporal CNN is able to increase the detection distance close to 30%, while retaining the same P_{FA} as the CA-CFAR.**

***Index Terms*—Radar, detection, UAV, deep learning, CNN, CFAR, non-homogeneous**

I. Introduction

In this paper we attempt a new approach to radar detection based on machine learning, aiming at decreasing the P_{FA} when detecting small targets in high clutter environments, trained exclusively on measured radar data. Detection is the task of investigating the received signal and deciding whether there is a target present or if the signal contains only interference and noise. There is in essence two hypotheses in the detection scenario. H_0, the null hypothesis, states that there is no target present, only interference and noise. H_1 states that the signal is a combination of a target response, interference and noise. The decision made by the detection algorithm is not always correct, and some detections might contain just interference and noise, this is known as a false alarm. Detection algorithms are usually designed according to the Neyman-Pearson criterion. This essentially fixes the false alarm rate, P_{FA}, and then maximises the probability of detection, P_D, for a given Signal to Noise Ratio (SNR) [1]. Setting a hard threshold requires detailed knowledge of the radar system, the operation environment and the observed target. All of these variables might change during radar operation, altering the P_{FA} or P_D. To alleviate this, floating-level detectors, called CFAR detectors, were developed. CFAR detectors analyse the signal from cell to cell and use a number of adjacent cells to estimate the interference level. This estimate is then used to adjust the threshold for detection to meet the desired P_{FA} on average. As an example, the CA-CFAR calculates the threshold based on the average interference power of the adjacent cells and has close to optimal performance in homogeneous interference environments. If a target or clutter appears in the measurement, the interference measurement will be skewed and can lead to excessive false alarms or target masking. This is widely known and several CFAR algorithms have been suggested to mitigate these effects such as the Ordered Statistic-CFAR (OS-CFAR) and the Greatest Of-CFAR (GO-CFAR). Ghandi and Kassam [2] presents a good analysis of several different CFAR detectors and their performance in non-homogeneous environments. They conclude that no CFAR algorithm performs well in all combinations of homogeneous and non-homogeneous background noise and suggest that an adaptive approach may be the best solution if the operational environments change.

Several approaches for target detection using machine learning approaches have been attempted. One attempt at an adaptive detector is presented by Qi and Hu [3], where they utilize a Neural Network to assess whether the background contains a clutter transition, multiple targets or homogeneous noise and then use the CFAR most appropriate for the estimated environment. Other researchers have attempted to use an Artificial Neural Network (ANN) as a detector [4] [5] [6] [7]. Amoozegar and Sundareshan [8] show that an ANN can achieve a higher P_D with the same P_{FA}, particularly when the window size is small. Kück [9] proved that an ANN can function as a universal detector in mixed non-Gaussian environments. Cheikh and Faozi [10] show that an ANN-CFAR performed better than CA-CFAR in K-Distributed clutter. Recently, Akhtar and Olsen [11] presented an approach to train an ANN with a CA-CFAR and correcting the mistakes of the CA-CFAR, yielding a decreased P_{FA}. Building on the work of Akhtar and Olsen, Carreto et. al. [12] presents the Smart-CFAR. Where they prove that an ANN can mimic a CA-CFAR, but also improve on CA-CFAR in specific scenarios, such as clutter edges. Common for all of these efforts is that they are based on simulated data and use conventional ANNs.

Dette er en postprint-versjon/This is a postprint version. Publisert versjon/Publised version: http://dx.doi.org/10.1109/RADAR42522.2020.9114828

In this paper, we propose a new machine learning approach for target detection in radar, based solely on measured radar data. By solely using measured radar data, we remove any assumptions on both target and background that would be introduced by simulated data thereby enabling the machine learning algorithm to model the target and noise more accurately. This could in turn help reduce the P_{FA} so that the system in effect can achieve a longer detection range with the same P_{FA}. This paper starts by explaining how the data collection and generation is achieved in Section II before continuing with an explanation of the machine learning detectors in Section III. The experimental results and analysis are presented in Section IV. These results are discussed in Section V before the conclusion of the paper i Section VI.

II. Data collection

To generate data for training and testing, a series of UAV flights were conducted at two separate locations, one for training and one for testing.

A. Radar system

The radar system used for the trials is an X-band Ubiquitous Frequency-Modulated Continuous Wave (FMCW) radar with digital beamforming. A ubiquitous radar system transmits a broad beam and the receive beams are generated in signal processing. Using a ubiquitous system for these measurements is very beneficial as all the data are available for post-processing. Raw data was recorded during the measurement trials, enabling the generation of training and test data after the trial. The radar system has a horizontal array, so it is not able to separate targets or clutter in elevation. During the trials, the system was configured with a 30 MHz bandwidth at 9.2 GHz center frequency and a Pulse Repetition Frequency (PRF) of 2.5 kHz. For each location, the position and heading of the radar system were accurately measured.

B. UAV flights

During the trials, the UAV was flown at several velocities and ranges from the radar to generate a diverse data set for both training and testing. The flights for the test data sets were chosen so that they included non-homogeneous clutter such as clutter ridges. The UAV was also flown beyond the maximum detection range of the system, using conventional CA-CFAR, to be able to have test data for an increase in detection range. The position of the UAV was logged using its internal navigation system at a rate of 10 Hz. The data was collected at remote locations to minimize the presence of other targets of interest within the surveillance area.

C. Generating data set

To generate the data set for training and testing, the raw radar data is matched filtered, Doppler processed and beamformed. Using the position and heading of the radar and the navigation data from the UAV, the beam containing the UAV is selected and the rest of the beams are discarded to remove any risk of side lobes from the target and minimize the risk of interfering targets. The range and radial speed of the UAV relative to the radar is estimated and the range-Doppler image is labelled based on this position estimate. As previously mentioned the UAV flights contain diverse clutter conditions. The training data set is comprised of 13369 range-Doppler images where each pixel is evaluated. The test data set contains 3425 range-Doppler images collected at a different time and location from the training data set.

III. Machine learning detectors

This section introduces the two machine learning approaches proposed in this paper, the CNN-CFAR and the temporal CNN-CFAR. The former uses a patch similar to that of the CA-CFAR on one range-Doppler image. The latter utilizes patches from three concurrent range-Doppler images by stacking them over each other.

A. CNN-CFAR

The previous attempts on machine learning detectors in radar has been based on fully-connected conventional ANNs. Using a CNN we take advantage of the spatial relations of the range-Doppler image, by assuming that features shifted in range or Doppler can be processed in a similar fashion. Since CNNs reuse weights for different spatial locations they dramatically reduce the need for training data. For efficiency and practical purposes we employ a simple fully-convolutional architecture similar to the segmentation setup described in [13]. The CNN-CFAR takes one range-Doppler image as input as shown in Figure 3 with the CUT in the middle. The guard cells used for the CA-CFAR in Figure 2 are not necessary for the CNN-CFAR. With a fully-convolutional network our setup works for arbitrary sized input. Our target output is a grid of the same size as the input image, where cells corresponding to UAV positions and velocities are labelled 1 and all other cells are labelled 0. The network is trained by minimizing a weighted channel-wise softmax cross-entropy loss. We restrict the theoretical field-of-view of the network, meaning the number of input cells that contribute to each output value, by using a small network with small convolutional kernels. This means that input patch for each output value is similar to that of our CA-CFAR.

B. Model description

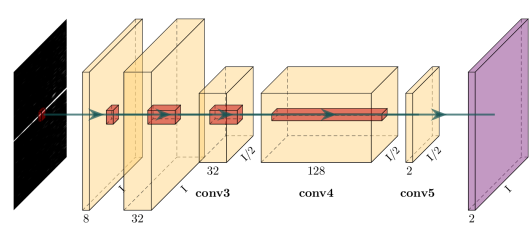

Fig. 1: Overview of network architecture for 1 image CNN. Red boxes indicate size of convolutional kernels, yellow boxes show the resulting output data of a layer. Numbers indicate the number of convolutional kernels for each layer.

Restricting the field-of-view keeps the model from overfitting to the training data. We applied randomized value scaling for augmentation and additionally used a small random crop to avoid overtraining border patches. Primarily for speed and simplicity in training, we chose to use a small and simple network architecture. With Scaled Exponential Rectified Unit (SELU) as activation function [14], we avoided the need for batch normalization and made our network even simpler. Our test architecture was a small 5-layer network with no batch-normalization (see Figure 1). Larger networks may yield better results, but exploring network architecture was deemed beyond the scope of this experiment. Run-time performance was not a concern for choosing the network architecture, as we can easily achieve real-time performance.

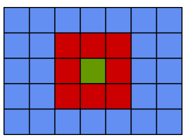

Fig. 2: Illustration of the 2D CFAR with CUT, Guard cells and noise estimate cells.

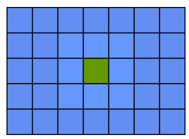

Fig. 3: Illustration of the 2D DNN detector patch, guard cells are not necessary.

C. Temporal CNN-CFAR

The Deep Temporal Detector proposed in this paper is inspired by the way a trained radar operator recognizes a target in a range-Doppler image. When observing a range-Doppler display an operator typically tracks real targets based on their persistence, not just their power. False detections are observed as flickering noise, whereas targets are consistent points. This is not possible for a CA-CFAR detector to assess, as it does not evaluate from dwell-to-dwell. Incoherent binary detections strategies such as the m-of-n detection criterion [1] take this into account, but only on a binary level. Other techniques for incoherent integration might also be utilized, but calculating a proper threshold for incoherent detection requires assumptions on target and noise. The Temporal CNN-CFAR attempts to exploit the local spatio-temporal information in the detection process by stacking three range-Doppler images on top of each other as separate channels. This is illustrated in Figure 4, where the three layers are combined. Three images were selected as a proof of concept, with the assumption that the target will have moved slightly, but not extensively between the first and last image. The idea can be expanded depending on target assumptions and radar configuration. Comparing the temporal CNN-CFAR with a CA-CFAR on a single range-Doppler image is therefore not completely fair, because they rely on different dwell times. This is a technique of stacking layers previously used in video classification and several approaches on how to fuse temporal information has been investigated by Karpathy et. al. [15]. In accordance with Karpathy [15], we found late fusion slightly better than early fusion. We implemented the slow-fusion regime, by using 3D-

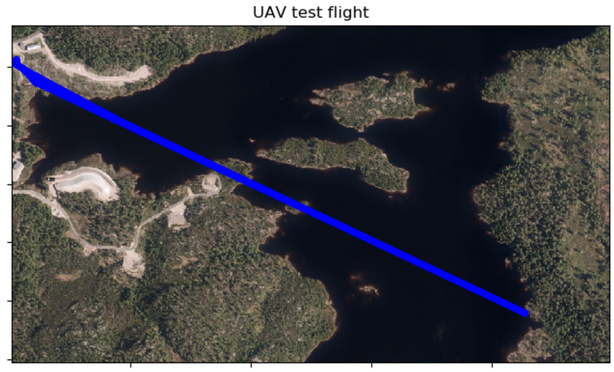

Fig. 5: UAV flight plotted over a satellite image.

convolutions with kernel size 1 in the temporal dimension before the fusion stage.

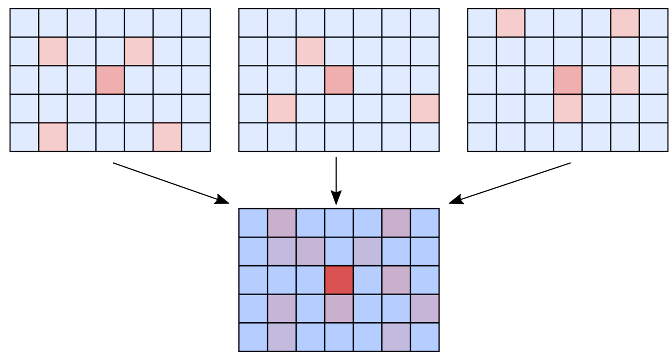

Fig. 4: Illustration of the stacking of range-Doppler images performed before the Temporal CNN-CFAR

IV. Experimental results and analysis

After training on a number of data sets, the network is run on the test set. The test set was recorded at a different location and under different conditions than the training set. It can be seen from the plot of the UAV test flight overlayed a satellite image that the test set contains a mix of homogenous background noise and clutter transitions. The test flight outbound and inbound from the radar at a distance of 0-1200 meters. The data has been processed with a 200 ms Coherent Processing Interval (CPI), which should yield an approximate detection distance of 800 meters using conventional CA-CFAR with the RF power setting during the trial.

A. Comparing the performance

There are no analytical terms for the P_{FA} and P_D of the CNN detections, the metrics are therefore extracted from the test data set using the following definitions:

$$P_{FA} = \frac{\text{false detections}}{\text{tested cells}}, \quad P_D = \frac{\text{correct detections}}{\text{targets}} \qquad (1)$$

Both the CA-CFAR and the CNN detectors output a float value for each tested cell. Using the test set, a large amount of these values are tested and the metrics for P_D and P_{FA}

Dette er en postprint-versjon/This is a postprint version. Publisert versjon/Publised version: http://dx.doi.org/10.1109/RADAR42522.2020.9114828

are calculated. This enables us to find the thresholds for a given P_{FA}. These metrics are calculated for both CFAR and CNN detectors, where all non-target cells are considered to be clutter cells. The CNN detectors effectively works as both detectors and clutter filter. To prevent the results from being skewed from slow-moving tree clutter, a simple clutter filter is applied to the CFAR output before calculating the metrics.

B. Performance analysis

Calculating the metrics for the full flight yields the Receiver Operating Characteristic (ROC) curve plotted in Figure 6. The curve reveals that the CNN and Temporal CNN detectors increased P_D with the same P_{FA} as the CA-CFAR. As an example, for a P_D of 0.8 for this full test set, the CA-CFAR would have a P_{FA} of close to 10^{-1}, whereas the P_{FA} for the Temporal CNN and the CNN is close to 10^{-4} and 10^{-2} respectively.

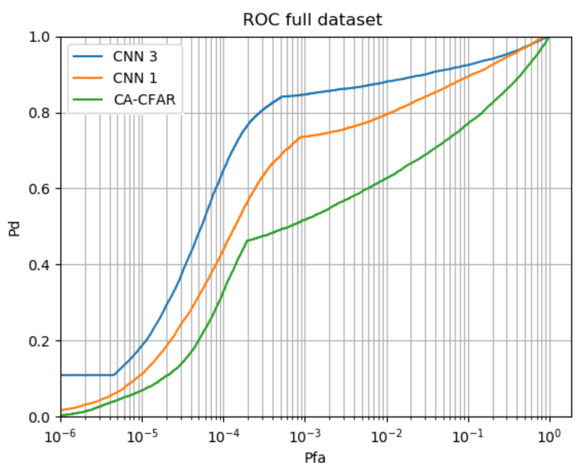

Fig. 6: P_D vs P_{FA} for CA-CFAR, CNN-detector and Temporal CNN-detector.

If we set the desired P_{FA} to be 10^{-3}, we can analyse how the P_D changes with respect to range. Figure 7 shows a histogram of the P_D at different ranges. Firstly, it is interesting to notice that the performance of the detectors is similar below 450 meters. The second thing to notice is the drastic effect of the two land-patches the UAV passed at 500-750 meters in range on all the detectors, but particularly the CA-CFAR. The clutter ridges are likely the cause of lowered P_D at approximately 480, 550 and 650 meters. The two CNN detectors increase the P_D approximately 0.25-0.35 in the clutter transitions. The highest clutter ridge improvement is seen in the area from 550-650 meters, which is directly between the two islands. In this section, the P_D of the CA-CFAR shows clear signs of masking with a P_D of close to 0.5, whereas the CNN detectors maintain a P_D close to 0.9.

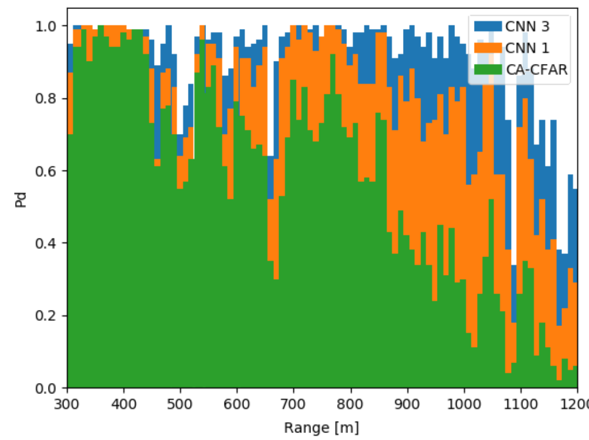

Fig. 7: Histogram of the P_D of several detectors as a function of range. P_{FA} set to 10^{-3}

The last interesting section is the range beyond 700 meters. From the flight path in Figure 5, we can observe that this area is mostly homogeneous, although both the transmit and receive beam are broad and will include a significant amount of clutter. The ROC for the ranges beyond 700 meters is plotted in Figure 8. This ROC confirms the findings in the previous histogram, the CNN detectors have significantly increased performance compared to the CA-CFAR in the low SNR scenario. These results indicate that the background in this area is not homogenous, given the poor performance of the CA-CFAR. Further analysis are needed to find the reason for this inhomogeneity and whether it is caused by external or internal noise.

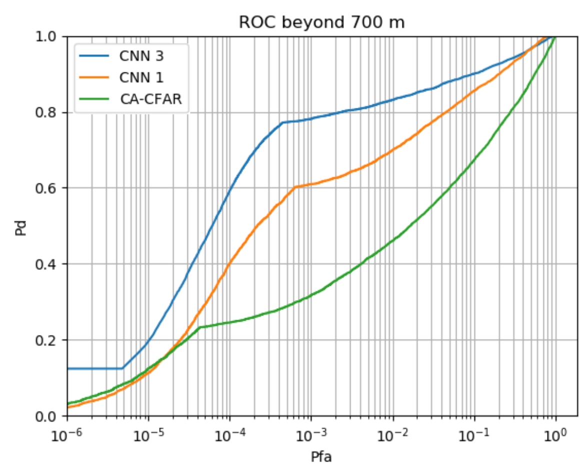

Fig. 8: P_D vs P_{FA} for CA-CFAR, CNN-detector and Temporal CNN-detector for ranges beyond 700 meters.

Throughout the last section, the CNN detectors show drastically increased P_D. As an example, the P_D using CA-CFAR at 1000 meters is approximately 0.15, whereas the corresponding P_D using the Temporal CNN detector is 0.9.

Dette er en postprint-versjon/This is a postprint version. Publisert versjon/Publised version: http://dx.doi.org/10.1109/RADAR42522.2020.9114828

This improvement results in a significant increase in detection distance. The CA-CFAR has a P_D of 0.8 at about 770 meters, whereas the Temporal CNN detector retains the same P_D to about 1080 meters, which is close to a 40% increase. According to the m-of-n detection criterion [1], in this case the 2-of-3 criterion, we would expect a maximum of 0.1 increase in P_D for the same P_{FA}. For ranges beyond 800 meters, the Temporal CNN has a 0.1-0.3 increase in P_D compared to the CNN detector, which is above the expected 0.1 increase of the 2-of-3 criterion [1].

V. Discussion

Both the CNN detectors outperform the CA-CFAR, with an increased P_D while retaining a low P_{FA} in both homogeneous and non-homogeneous environments. The CNN detectors can increase the detection range around 30% under these conditions while retaining the same P_{FA}. Investigations into the nature and distribution of the background noise under the test conditions remain. The accuracy in the position and velocity estimates of the UAV can lead to wrongly-labelled data that can inhibit the model learning. Training larger and more advanced CNN architectures, gave little to no improvement, perhaps due the noise in the data. We, therefore, believe that CNN detectors can perform even better with better data generation. More data and higher data diversity can also increase the performance, hopefully in the areas of clutter ridges where the CNN detectors show a good performance increase, but still have room for improvement. More ways of augmenting the data might also increase the diversity of the data, by adding more combinations of target and noise. The test and training data sets were collected at separate locations with different conditions, this indicates that the model is able to generalize.

VI. Conclusion

In this paper, we attempted a new machine learning approach to target detection in range-Doppler images. The goal was to decrease the P_{FA} for small targets in high clutter environments. This would, in turn, increase the P_D for the same P_{FA} and in effect increase the detection distance of the system. The CNN detectors were trained solely on measured data, labelled by using position and velocity estimates from a UAV. The first approach was the CNN detector, which proved to have similar performance to the CA-CFAR in the high SNR scenarios while outperforming the CA-CFAR when the SNR is low or the target is in a non-homogeneous noise environment. At most, the CNN improves the P_D close to 0.4 compared to the CA-CFAR. The second approach was the temporal CNN detector, which uses 3 consecutive range-Doppler images as channels to include temporal information. This gives an approximate 0.1-0.3 increase P_D compared to the CNN detector. The current test results indicate that the temporal CNN can increase the detection distance of close to 30%, while retaining the same P_{FA} as the CA-CFAR. Further work will address the position estimates of the UAV to improve the accuracy of the training and test data. More data will also be collected in environments of higher diversity.

Acknowledgement

Some of the measurments presented in this paper were conducted as a joint effort of FFI, TNO and Thales Nederland B.V. The measurements were organized within the framework of the NATO SET-245 research group.

References

[1] M. A. Richards, W. A. Holm, and J. Scheer, *Principles of Modern Radar: Basic Principles*. Electromagnetics and Radar, Institution of Engineering and Technology, 2010.

[2] P. Ghandi and S. Kassam, "Analysis of {CFAR} processors in non-homogeneous background," *IEEE Trans. Aerospace and Electronic Systems*, vol. 14, no. 4, 1998.

[3] Q. Qi and W. Hu, "One efficient target detection based on neural network under homogeneous and non-homogeneous background," *International Conference on Communication Technology Proceedings, ICCT*, vol. 2017-Octob, pp. 1503–1507, 2018.

[4] D. Mata-Moya, N. Del-Rey-Maestre, V. M. Peláez-Sánchez, M. P. Jarabo-Amores, and J. Martín-De-Nicolás, "MLP-CFAR for improving coherent radar detectors robustness in variable scenarios," *Expert Systems with Applications*, vol. 42, no. 11, pp. 4878–4891, 2015.

[5] P. P. Gandhi and V. Ramamurti, "Neural networks for signal detection in non-gaussian noise," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2846–2851, 1997.

[6] S. Wunsch, J. Fink, and F. K. Jondral, "Improved detection by peak shape recognition using artificial neural networks," *2015 IEEE 82nd Vehicular Technology Conference, VTC Fall 2015 - Proceedings*, no. 0, pp. 1–5, 2016.

[7] N. Gálvez, J. l. Pasciaroni, O. Agamennoni, and J. Cousseau, "Radar signal detector implemented with artificial neural networks," 06 2004.

[8] F. Amoozegar and M. K. Sundareshan, "Robust neural network scheme for constant false alarm rate processing for target detection in clutter environment," *Proceedings of the American Control Conference*, vol. 2, pp. 1727–1728, 1994.

[9] Kück, "Constant False Alarm Rate Detection of Radar Signals with Artificial Neural Networks," 1996.

[10] K. Cheikh and F. Soltani, "Application of neural networks to radar signal detection in K-distributed clutter," *IEE Proceedings: Radar, Sonar and Navigation*, vol. 153, no. 5, pp. 460–466, 2006.

[11] J. Akhtar and K. E. Olsen, "Go-cfar trained neural network target detectors," in *2019 IEEE Radar Conference (RadarConf)*, pp. 1–5, April 2019.

[12] M. V. Carretero, R. I. A. Harmanny, and R. P. Trommel, "Smart-CFAR , a machine learning approach to floating level detection in radar," *EuRAD 2019*, pp. 161–164, 2019.

[13] Y. Zhang, Z. Qiu, T. Yao, D. Liu, and T. Mei, "Fully Convolutional Adaptation Networks for Semantic Segmentation," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 6810–6818, 2018.

[14] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-Normalizing Neural Networks," tech. rep.

[15] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and F. F. Li, "Large-scale video classification with convolutional neural networks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1725–1732, 2014.