

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/320065040>

High Speed Video-based health monitoring using 3D Deep Learning

Conference Paper · August 2017

CITATIONS

0

READS

48

4 authors:



Sambuddha Ghosal

Iowa State University

9 PUBLICATIONS 13 CITATIONS

SEE PROFILE



Adedotun Akintayo

Iowa State University

13 PUBLICATIONS 22 CITATIONS

SEE PROFILE



Paige Boor

Iowa State University

1 PUBLICATION 0 CITATIONS

SEE PROFILE



Soumik Sarkar

Iowa State University

98 PUBLICATIONS 503 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Video/Frame based health monitoring using Deep Learning [View project](#)



Automated image analytics for phenotyping root and hypocotyl growth of Arabidopsis plants [View project](#)

All content following this page was uploaded by [Sambuddha Ghosal](#) on 27 September 2017.

The user has requested enhancement of the downloaded file.

High Speed Video-based health monitoring using 3D deep learning

Sambuddha Ghosal
Department of Mechanical
Engineering
Iowa State University
Ames, USA
sghosal@iastate.edu

Adedotun Akintayo
Department of Mechanical
Engineering
Iowa State University
Ames, USA
akintayo@iastate.edu

Paige Boor
Department of Mechanical
Engineering
Iowa State University
Ames, USA
pkboor@iastate.edu

Soumik Sarkar
Department of Mechanical
Engineering
Iowa State University
Ames, USA
soumiks@iastate.edu

Abstract— Recently, deep learning models have been shown to outperform other state of the art techniques in handling and analyzing large dimensional data (both spatial and temporal), by learning the hierarchical features to perform various tasks such as, classification and bulk structure detection given a large corpus of 2-dimensional (2D) data or images. As an extension, embedding of 3-dimensional (3D) spatiotemporal data (where the data have spatial features evolving over time) has been performed using a 3D convolutional neural network (3D CNN) framework. In this paper, we apply for the first time such a 3D CNN architecture for early detection of combustion flame instability using hi-speed flame videos. We demonstrate the performance of our proposed framework on an experimental data set of hi-speed (i.e., sampling at 3 KHz) video consisting of flame images collected from a laboratory scale swirl-stabilized combustor. Such early detection of combustion instability can eventually enable health monitoring and enhance fuel efficiency of an engine or power generation system that uses a combustor. The main contribution of the paper is development of a 3D CNN model that is pre-trained by the encoder part of a 3D convolutional autoencoder that outperforms the recent results reported using 2D deep learning frameworks (that are unable to leverage the temporal correlations among the consecutive frames) on the same data set. For training the 3D network, multiple 2D frames are stacked into short temporal segments and the performance of the model is evaluated in detecting regions of stability, instability and most importantly regions of evolving flame intermittencies that is crucial for an early warning system. In general, the model development and analysis presented in this paper opens up the door for leveraging 3D CNN models for prognostics and health monitoring of human-engineered systems using large volume spatiotemporal data.

Keywords—health monitoring, combustion instability, high-speed video, 3D Deep Learning, spatiotemporal learning

I. CONTRIBUTION

We propose a deep learning-based tool that has enormous potential for preventing catastrophic failure in engines by early detection of combustion instability. Control modules can leverage such information to enable recovery of the combustion process at least to a gracefully degraded condition. Such diagnostics tools are of particular interest because of the current focus on reduction of fossil fuel use in jet engines and power plants. However, the equivalent mixtures obtained by reducing the fuel-to-air ratio in the combustors have been found to result in combustion instabilities [1]. Among the common effects of instabilities that hurt the health and reliability of the engines are the blow-out and excessive flame heating that cracks the wall of the chamber. Therefore, the safety of the customers and operators of such systems become compromised.

As sensors such as pressure or chemiluminescence may not be able to detect sufficiently early, recent research studies have leveraged high-speed flame video and sparse DMD has been applied to detect stability and instability in the flame images by essentially extracting the primary modes of the evolving flame [2]. Also, a neural-symbolic framework was explored for extracting salient features from multi-modal sensors data that uses labeled data of stable and unstable frames. The framework fuses the features from the images with the information provided by the pressure sensor [3]. An end-to-end (from image to image) type 2-dimensional convolutional selective autoencoder (2D CAE) was designed to jointly learn the features that encode the information in the stable frames as well as the coherent structure information present in the unstable frames [4]. The 2D CAE network was successful at revealing the intermittent instabilities, however, due to the inability of capturing temporal characteristics, the performance of such models was sub-optimal especially in detecting intermittent instability indicators before the flame becomes completely unstable. To capture such dynamics of the frames, a 3D convolutional neural network (3D CNN) can be trained. The training inputs of the network are composed of short time stacks of the 2D image frames that explicitly model

the dynamics of neighboring dependent frames. An example of the time evolution of the stable and unstable frames is shown in Figure 1. The frames shown to be stacked to form the 3D ‘super voxels’ that motivate the implementation of our architecture. There were 63,000 frames available for training the images, out of which 35,000 frames were taken from the stable region and 28,000 frames were collected for the unstable region. A detailed breakdown of the frames are available [3, 4].

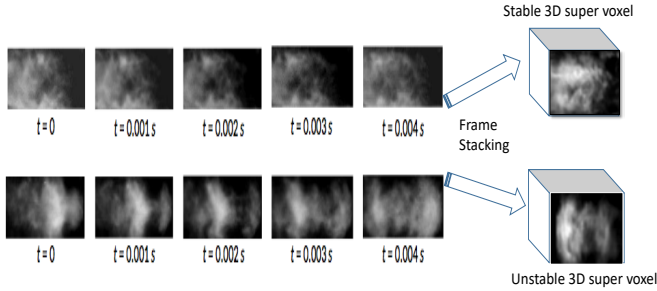


Figure 1: Examples of stable and unstable flame frames in a typical temporal evolution and their respective 3D super voxels.

The network in Figure 2 is made up of a series of convolution and maxpooling layers whose feature maps perform the detection of image and feature edges, shapes and objects. Convolutions are initiated by choosing appropriate kernel sizes and numbers that are shifted and convolved on the input images, following which some activations such as rectified linear units [5] are applied for learning the joint kernels in local neighborhoods that are useful for describing the explanatory features at that layer. Maxpooling layers are included to learn the scale invariance of the input images. The fully connected layers embed the high dimensions in a low dimension space, activates ensure that the orientations do not necessarily hurt the network performance and produces the classification of the network at the coding layer with an activation type called softmax [5].

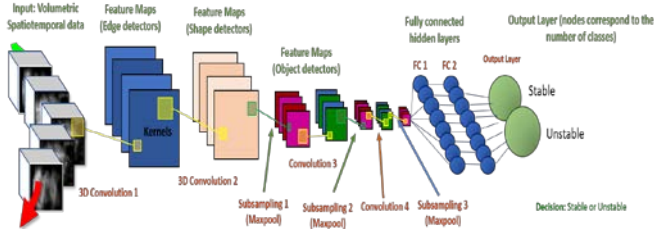


Figure 2: 3D CNN architecture consisting of example image snapshots, hierarchical layers and model parameters that were trained to learn the suitable features from frames in the stable and unstable frames.

Network training is completed by back-propagating the error in a reverse direction to the activation. In that process, the weights are modified by using some standard algorithm such as the information-theoretic cross entropy function [5] which examines the level of filter achieved at each run of the algorithm.

After various inferences using different network structure, we found a short time segment of 32 frames suitable for better

detection of the intermittent unstable frames before the onset of complete instability.

II. RESULT

Inference is done using unseen test videos from various experiments with different air-to-fuel ratio protocols such as: Protocols 500_{40to28} , 500_{40to30} and $500to600_{40}$ where the digits are the air flow rate (in litres/minute) and the subscripts denote the fuel flow rate (also in litres/min). The protocols model flame transition from stable to unstable regions in compliance with increasing the relative air-to-fuel ratio since the air flow has been reduced in the experiments. The machine learning tool described by the network in Figure 2 is examined to determine the dynamics that take place in between the stable and unstable frames with the goal of determining the instabilities present stable frames as well as the onset of instability. The results in Figure 3 show the instability levels against the 3D frame numbers (#).

From Figure 3, we note that the 3D CNN framework is able to efficiently detect onset of instability (termed as ‘‘intermittency’’) in the flame flow marked by the peaks in the plots (encircled in red). These regions could not be detected earlier using the 2D based methods reported in literature. Thus, 3D CNN proves to be more effective in detecting these crucial points thus effectively analyzing the flame flow in a spatiotemporal manner. Apart from detecting the intermittencies efficiently, the framework can also predict the regions of stability and total instability effectively.

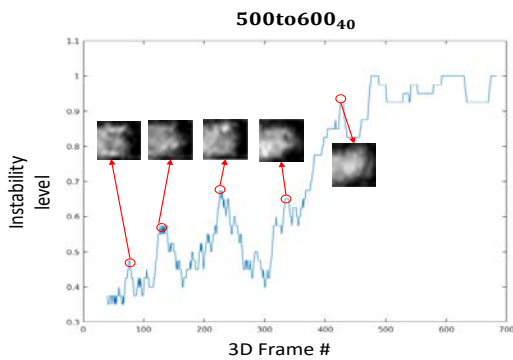
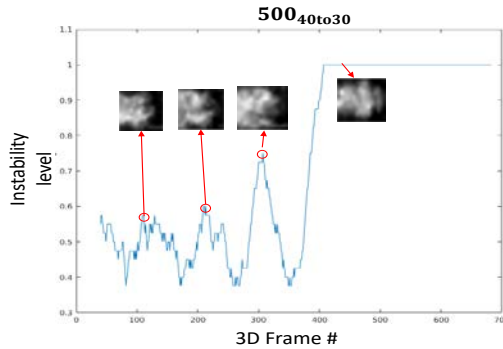
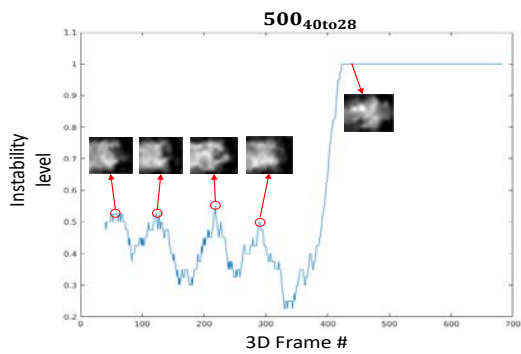


Figure 3 3D CNN Results for the 3 considered protocols with example frames from each of the unstable intermittencies in the insets.

ACKNOWLEDGEMENTS

This work has been supported in part by the U.S. Air Force Office of Scientific Research under the YIP grant on “Neural-Symbolic approach to Real-time Decision-making in Complex Aerospace Systems”. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

REFERENCES

- [1] S. Sarkar, K. G. Lore, S. Sarkar, V. Ramanan, S. R. Chakravarthy and S. Phoha, Early detection of combustion instability from hi-speed images via deep learning and symbolic time series analysis. In Annual conference of The Prognostics and Health Management, pp.1-9, 2015. (references)
- [2] S. Ghosal, V. Ramanan, S. Sarkar, S. Chakravarthy, S. Sarkar, “Detection and Analysis of Combustion Instability From Hi-Speed Flame Images Using Dynamic Mode Decomposition,” P2016.Proceedings of the ASME DSCC, 2016.
- [3] S.Sarkar, K. G. Lore, S. Sarkar, Early Detection of Combustion Instability by Neural Symbolic Analysis on Hi-speed Video, Neural Information Processing Systems (NIPS), pp.1-9, 2015.
- [4] A. Akintayo, K. G. Lore, S. Sarkar and S. Sarkar, Prognostics of Combustion Instabilities from Hi-speed Flame Video using A deep Convolutional Selective Autoencoder, International Journal of Prognostics and Health Management, vol. 7, no 023 (Special Issue Big Data and Analytics), pp.1-14, 2016.
- [5] A. Krizhevsky, I. Sutskever and G. E. Hinton, Imagenet classification with deep convolutional neural networks, Neural Information Processing Systems, pp.1-9,2012