M. Nugue, J.-M. Roche, G. Le Besnerais, C. Trottier, R. W. Devillers (ONERA)

J. Pichillou (CNES)

A. Chan-Hon-Tong, A. Boulch, A. Hurmane (ONERA)

E-mail: adrien.chan_hon_tong@onera.fr

DOI: 10.12762/2020.AL15-07

Scaling Up Information Extraction from Scientific Data with Deep Learning

This paper presents two use cases where deep learning is able to help scientists by removing the burden of manual review of large volumes of physical data. Such examples highlight why deep learning could become a transverse tool across many scientific fields.

Introduction

Designing an experimental protocol to validate a scientific hypothesis, and analyzing the resulting data is an intellectual process that cannot be solved by current artificial intelligence. However, in this process at the edge of the human mind, there are often sub-tasks that are repetitive, time-consuming and that do not involve expert skill or knowledge. This paper is aimed at showing that existing artificial intelligence, like deep learning, can solve such tasks.

For example, there are common scientific issues where understanding a physical phenomenon is slowed down by the need to extract information from large volumes of physical data, and in this context the application of deep learning can be very relevant. This article presents such scientific use cases, and the deep learning algorithms that have been able to scale information extraction where humans alone have not.

Specifically, this article focuses on two successful ONERA scenarios linked to the DELTA project¹: solid-propellant combustion and material resistance.

Solid-propellant combustion analysis by shadowgraphy

Solid propellants are widely used for spatial and military applications. For example, they are used for the first stage of the Ariane V launcher. Classically, aluminum particles are included in the composite to increase the efficiency of the solid-propellant combustion, improving the thrust by 10% [4]. However, aluminum particles can trigger various negative effects, such as diphasic losses, film formation on the nozzle or pressure instability (e.g., thermoacoustic instabilities [6]).

Understanding the physical phenomena associated with burning aluminum droplets is a critical issue to design new generations of engines and/or new solid propellant compositions. Typically, numerical simulation results are highly dependent on two critical input parameters: the initial size distribution of the droplets, and their initial velocity distribution when they leave the burning surface [7]. Defining precise representative droplet size distributions is not as easy as it may seem since agglomeration phenomena [5] strongly modify the droplet diameters due to the aluminum powders introduced in the composite material. Similarly, accurate velocity data are not available today under representative burning propellant conditions (for instance, recent data were published [3] but only for 0.1MPa burning conditions). Hence, obtaining access to these two data would be a breakthrough to update physical models and the numerical simulations for solid-propellant combustion.

For the past couple of years, the ONERA solid propellant research team has been using an experimental setup to characterize aluminum droplet combustion under relevant solid propellant conditions [15]. High-speed visualization of droplets is achieved via a focused shadowgraphy diagnostic (see Figure 1). The diagnostic



Figure 1 – Shadowgraphy setup enabling aluminum droplets to be visualized during solid-propellant combustion

https://delta-onera.github.io/



Figure 2 – Images obtained by shadowgraphy

is sensitive to refractive index gradients, leading to a good contrast between the liquid aluminum droplets and the surrounding hot gases (see Figure 2).

The experimental images in Figure 2 show that ejected burning aluminum droplets are clearly visible to human eyes as bright grey balloons over the dark background. Thus, from a theoretical point of view, there is no obstacle for obtaining the size and speed of particles: *it only requires somebody to scan the existing* experimental images and to extract the data.

However, each combustion test generates between 1,000 and 10,000 images, each showing several dozen aluminum droplets. Considering the volume of data, extracting the desired parameters is impossible using human hands/eyes. In other words, today, combustion scientists know how to generate data containing the information but not how to extract the information from this large volume of data. Making progress on an algorithm enabling this information extraction could allow this scientific obstacle to be overcome.

Modeling some material properties with microscopy

The ability to predict the fracture behavior of a material under strain is essential to design a car, building or plane. However, in aeronautics, the use of a large variety of materials (especially composite materials) makes it difficult to derive the fracture behavior solely from solid mechanic physical equations. Thus, fracture mechanics sometimes relies on experimental data to characterize a special kind of material [9].

Such an experiment consists in measuring a material state at different levels of strain, and producing a stress versus strain curve. Measurements can include acoustic measurements, image correlation (when it is possible to obtain an image before deformation), thermal imaging, or microscopy imaging.

A typical example of such a microscopy experiment is provided in Figure 3. Carbon epoxy laminates enhance damages within their plies. These damages are observed at all 3 scales of the material, i.e., the microscale (fiber and matrix), the mesoscale (ply), and the macroscale (laminate). The main damages that can appear during a tension test of a [0/90]s laminate are illustrated in Figure 3. First of all, when damage at the microscale appears, it is in the form of fiber/matrix debonding or micro cracks within the resin (3.1 in the figure). Loading the sample causes the damages at the fiber scale to grow upward and leads to a transverse matrix crack at the ply scale (3.2) parallel to the fiber direction. These cracks cross the entire thickness

AerospaceLab

of the ply. In some composites, micro delamination can be observed at the crack tips. Finally, when the strength limit of the sample is reached, the crack density increases (3.3) and leads to fiber failure due to redistribution of the load (3.4).

Here again, scientists have designed an experimental protocol allowing the behavior of these materials to be modelled under stress. However, this involves the analysis of all resulting data to count and/ or characterize all fractures. Unfortunately, the volume of resulting data is also too large to be processed by human hands/eyes in these experiments. Again, making progress on the development of an algorithm enabling this information to be extracted could allow this scientific obstacle to be overcome.

A few words about deep learning

Deep learning (which is becoming the state-of-the-art in image classification with [10]) consists simply of neural networks (which existed at least since [12]) with a large number of layers. This simple enlargement of neural networks, allowed by the increase in available GPU, eventually leads to very impressive performance gap with previous neural networks attempts. However, deep learning has been successfully applied to many tasks: image classification [10], object detection in images [8], natural image segmentation [1], semantic segmentation in remote sensing images [11], natural language processing like automatic translation [22], sound classification [17], cyber intrusion detection [21], malware classification [25], games (with a reinforcement learning framework) [19], medical diagnosis [2], etc.

Supervised classification: More generally, deep learning is a sub-field of supervised machine learning, which is a sub-field of machine learning, which in turn is a sub-field of artificial intelligence. The simplest form of supervised machine learning is binary supervised classification: it consists in estimating an unknown function *f* of \mathbf{R}^{D} to $\{-1,1\}$ from training data; i.e., a set of samples $(x_1, y_1 = f(x_1)), \dots, (x_N, y_N = f(x_N))$. A learning step in this training data leads to a function \hat{f} (classically $\hat{f}(x) = g(x, \hat{w})$ and learning consists in selecting weights *w* to set up a function family g(x, w)). This function \hat{f} is an approximation



Figure 3 – Typical experiment to characterize a material

f accepting that \hat{f} may be different from f. Usually, the quality of the approximation is evaluated by comparing \hat{f} and f for a disjoint set of samples known as the test data. Of course, given that multiple functions f will lead to the same training data, there is no mathematical possibility of finding the correct function in the discrete case (this statement is confirmed in [23]). However, in practice, given that the function f to be approximated generally has a good regularity property, in practice very interesting results can be achieved.

For example, the function f could use an image as input and provide cat or dog as the output depending on whether the image is an image of a cat or an image of a dog. Using a human being, f can be trivially evaluated on a particular image just by asking the human whether this is an image of a cat or an image of a dog. However, nobody is able to compute this function exactly, even with an arbitrarily powerful computer, since the steps leading to the decision are unknown. However, it is possible to collect training data and to approximate this function quite efficiently, especially with deep learning.

Supervised segmentation: A very natural extension to supervised classification on images is supervised segmentation, where the goal is to classify each pixel of an input image (see Figure 4).

Contrarily to most other classifiers, deep learning based classifiers are much more flexible. Simply changing a few layers (or even just changing the layer structure) allows a classifier to be converted into a segmentation algorithm. In addition, specific networks for



Classification + Localization

Object Detection

Figure 4 - Illustration of different natural extensions of classification

segmentation have been designed, like UNET [18], whose architecture is presented in Figure 5. The first half of the network is considered to be an encoder, i.e., it extracts more and more relevant information from the image, but loses more and more precise spatial information. Then, the second half would decode the decoder; i.e., it restores precise spatial location by combining different levels of encoded information. At the end of the process, the network produces a map with the same spatial size as the input image, in which each pixel is associated with a likelihood per class.

This framework is especially relevant for the two scientific use cases presented in this paper. Indeed, in these two use cases, relevant physical information can be obtained from a segmentation map.

Furthermore, when processing scientific data, deep learning can safely assume that hacked data will not be encountered (indeed, sensibility to adversarial attacks [13, 24, 16, 20, 14] is an important issue for deep learning).

Particle segmentation by deep learning

As presented in Figures 1 and 2, the shadowgraphy diagnostic provides experimental images of the solid-propellant combustion showing aluminum droplets. Thus, a straightforward idea is to apply UNET (see Figure 4) to shadowgraphy images (Figure 2).

First, given that we are considering supervised segmentation, it is unfortunately unavoidable to consider the issue of designing a semantic segmentation dataset. Thus, this step consists in manually annotating some images for semantic segmentation; i.e., providing a class for each pixel.

This is a very repetitive and time-consuming task. The good news is that, contrarily to common thinking, UNET does not require too many images to reach a relevant state depending on the type of data.



Figure 5 – UNET is a representative example of a segmentation network

AerospaceLab



Figure 6 - Illustration of the manual annotation required to apply UNET



Figure 7 - Illustration of a UNET prediction on a testing image

Given that each pixel leads to a decision in semantic segmentation, an image corresponds to 1 million data points. This is a number quite close to the size of the well-known IMAGENET dataset. Off course, pixels are much more correlated than the IMAGENET examples, but, on the other hand, the variability of shadowgraphy images may be lower than the variability of a natural image. Here, the difficulty is mainly due to the low contrast, irregular shape and very complex features of shadowgraphy images. With these ideas in mind, we only annotate 30 images from 3 experimental videos (10 images per video).

Then, we perform a cross-validation: we learn from 2 videos (20 images) and test the model on the last video (10 images). This way, we can evaluate the generalization to new videos. Eventually, the algorithm is able to generalize (quite fairly) to new videos: see Figure 7.

Now, our goal is to estimate the droplet size distribution during combustion. A perfect segmentation of the images enables the size distribution to be evaluated accurately. However, predicted segmentation will always be more or less noisy, and it is not trivial to derive the resulting noise in the droplet size distribution. Also, estimating the distribution on a small subset of the images (e.g., the annotated data) may not be sufficiently accurate. Thus, we cannot use as a metric the distance between the distribution estimated by annotation and that obtained by the algorithm. Hence, we propose here to only quantify the segmentation (and not the resulting size distribution). In our cross-validation evaluation, the multi-class accuracy of UNET (i.e., the number of pixels correctly predicted divided by the number of pixels) is above 90% (training is not convex, so multiple runs do not lead exactly to the same accuracy even if variance is not an issue).

After post-processing, the resulting size distributions with different post processing (estimated on all frames of the video) are presented in Figure 8 alongside the size before burst (which is known to change during burst). Even though none of these curves can be considered as reference distributions, we can see that they seem consistent. Estimated distributions have a greater number of large particles, which is expected due to the aggregation phenomenon during burst.

As a partial conclusion, although the developed process is not complete (particularly with regard to guaranteeing the estimated distribution), it leads to very promising results and already allows input data to be refined for numerical simulations of solid-propellant combustion.



Figure 8 – Illustration of the extracted size distribution

AerospaceLab

Crack segmentation by deep learning

Like in the previous use case, UNET is a straightforward idea for crack segmentation. Here, other algorithms from the computer vision community could be relevant. Indeed, the semantic object of interest is not defined by itself but by opposition to a normal situation. Thus, abnormality detection, or even, edge detection could have been investigated at first glance. However, abnormality detection is currently less mature than supervised detection, and edge detection may be deceiving considering that a non-uniform material leads to strong contrast. Thus, using supervised deep learning to capture the structure of cracks, which are the only edges of interest, is a good shot.

The processing of these images is considered as a binary semantic segmentation problem where the goal is to give to each pixel either a crack tag or a background tag.

These data are in a way harder to obtain than shadowscopy ones (from a deep-learning point of view) because of the unbalance of crack pixels vs. background pixels, and the required spatial accuracy (as cracks are linear area not delimited surfacique object).

Indeed, the number of background pixels is much greater than the number of crack pixels. Thus predicting all pixel as background leads to a very small error i.e. a good local minimum.

However, applying UNET with an *ad hoc* weighting of error proved to be sufficient to learn a relevant model for crack segmentation. The weighting consists in penalizing an error on a crack pixel much more than an error on a background pixel. This way, the network is forced to go outside the local minimum where all pixels are background ones.

The resulting performances of the network are quite interesting as shown by Figures 9 and 10.

Conclusion

In this paper, we focus on two real-life use cases where the size of the experimental data is an obstacle in the understanding of a physical phenomenon. We show how, in these two contexts, deep learning was successfully applied to help scientists to achieve the processing of the experimental data.

However, this success is not that surprising, since the selected experimental data are images, which is one of the primary areas of deep learning development. However, this article still shows that state-ofthe-art deep learning is becoming more and more mature for use in processing scientific experimental data

Acknowledgement

AerospaceLab

We wish to thank Antoine HURMANE and Francois Henri LEROY, for their help with this paper. Alexandre BOULCH is now a researcher at Valeo, but this work was done while he was working at ONERA.



Figure 9 – Application of UNET to detect cracks on a glass/epoxy 2D woven composite. (a) Microscopic observation of the entire sample under tension, (b) Zoom on cracks, (c) Manual assessment of the crack (in yellow), (d) Cracks predicted by our UNET model (in pink).



Figure 10 – Application of UNET to detect cracks on a carbon/epoxy [0/90] laminated composite. (a) Microscopic observation of the entire sample under tension, (b) Zoom on cracks, (c) Manual assessment of a crack (in yellow), (d) Cracks predicted by our UNET model (in green).

References

Aerospace Lab

- [1] V. BADRINARAYANAN, A. HANDA, R. CIPOLLA Segnet: A deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling. arXiv preprint arXiv:1505.07293, 2015.
- [2] O. BERNARD, A. LALANDE, C. ZOTTI, F. CERVENANSKY, X. YANG, P.-A. HENG, I. CETIN, K. LEKADIR, O. CAMARA, M. A. G. BALLESTER, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: is the Problem Solved? IEEE transactions on medical imaging, 37(11):2514-2525, 2018.
- [3] Y. CHEN, D. GUILDENBECHER, K. HOFFMEISTER, M. COOPER, H. STAUFFACHER, M. OLIVER, E. WASHBURN Study of Aluminum Particle Combustion in Solid Propellant Plumes using Digital In-Line Holography and Imaging Pyrometry. Combustion and Flames, 182:225-237, 2017.
- Y. FABIGNON, J. ANTHOINE, D. DAVIDENKO, R. DEVILLERS, J. DUPAYS, D. GUEYFFIER, J. HIJLKEM, N. LUPOGLAZOFF, J.-M. LAMET, L. TESS, A. GUY, C. ERADS Recent Advances in Research on Solid Rocket Propulsion. AerospaceLab Journal, 11:AL11-13, 2016.
- [5] S. GALLIER A stochastic Pocket Model for Aluminium Agglomeration in Solid Propellants. Propellants Explos. Pyrotech, 34:97-105, 2009.
- [6] S. GALLIER, F. GODFROY Aluminum Combustion Driven Instabilities in Solid Rocket Motors. Journal of propulsion and power, 25(2):509-521, 2009.
- [7] A. GENOT, S. GALLIER, T. SCHULLER Thermo-Acoustic Instabilities Driven by Fuel Droplet Lifetime Oscillations. Proceedings of the Combustion Institute, 37(4):5359-5366, 2019.
- [8] R. GIRSHICK Fast R-CNN. Proceedings of the IEEE International Conference on Computer Vision, pages 1440-1448, 2015.
- [9] C. HUCHETTE Sur la complémentarité des approches expérimentales et numériques pour la modélisation des mécanismes d'endommagement des composites stratifiés. Thesis, 2005.
- [10] A. KRIZHEVSKY, I. SUTSKEVER, G. E. HINTON Imagenet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems, pages 1097-1105, 2012.
- [11] A. LAGRANGE, B. LE SAUX, A. BEAUPERE, A. BOULCH, A. CHAN-HON-TONG, S. HERBIN, H. RANDRIANARIVO, M. FERECATU Benchmarking Classification of Earth-Observation Data: from Learning Explicit Features to Convolutional Networks. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pages 4173-4176. IEEE, 2015.
- [12] Y. LECUN, B. E. BOSER, J. S. DENKER, D. HENDERSON, R. E. HOWARD, W. E. HUBBARD, L. D. JACKE Hand-Written Digit Recognition with a Back-Propagation Network. Advances in Neural Information Processing Systems, pages 396-404, 1990.
- [13] S. M. M. DEZFOOLI, A. FAWZI, O. FAWZI, P. FROSSARD Universal Adversarial Perturbations. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.
- [14] A. NGUYEN, J. YOSINSKI, J. CLUNE Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 427-436, 2015.
- [15] M. NUGUE, R. W. DEVILLERS, G. LE BESNERAIS, N. CESCO Particle Detection and Size Evaluation in Solid Propellant Flames via Experimental Analysis to Improve Two-Phase Flow Simulation in Rocket Motors. Space Propulsion 2016 Conference, Rome, page 3124722, 2016.
- [16] N. PAPERNOT, P. MCDANIEL, S. JHA, M. FREDRIKSON, Z. B. CELIK, A. SWAMI The Limitations of Deep Learning in Adversarial Settings. Security and Privacy (EuroS&P), 2016 IEEE European Symposium on, pages 372-387. IEEE, 2016.
- [17] K. J. PICZAK Environmental Sound Classification with Convolutional Neural Networks. 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), pages 1-6. IEEE, 2015.
- [18] O. RONNEBERGER, P. FISCHER, T. BROX *U-Net: Convolutional Networks for Biomedical Image Segmentation.* International Conference on Medical image computing and computer-assisted intervention, pages 234-241. Springer, 2015.
- [19] D. SILVER, J. SCHRITTWIESER, K. SIMONYAN, I. ANTONOGLOU, A. HUANG, A. GUEZ, T. HUBERT, L. BAKER, M. LAI, A.BOLTON, et al. Mastering the Game of Go Without Human Knowledge. Nature, 550(7676):354, 2017.
- [20] C. SZEGEDY, W. ZAREMBA, I. SUTSKEVER, J. BRUNA, D. ERHAN, I. J. GOODFELLOW, R. FERGUS Intriguing Properties of Neural Networks. technical report arxiv:1312.6199, 2013.
- [21] A. TUOR, S. KAPLAN, B. HUTCHINSON, N. NICHOLS, S. ROBINSON *Deep Learning for Unsupervised Insider Threat Detection in Structured Cybersecurity Data Streams.* Workshops at the Thirty-First AAAI Conference on Artificial Intelligence, 2017.
- [22] A. VASWANI, N. SHAZEER, N. PARMAR, J. USZKOREIT, L. JONES, A. N. GOMEZ, L. KAISER, I. POLOSUKHIN Attention is All you Need. Advances in neural information processing systems, pages 5998-6008, 2017.
- [23] D. H. WOLPERT The Lack of a Priori Distinctions Between Learning Algorithms. Neural computation, 8(7):1341-1390, 1996.
- [24] C. XIE, J. WANG, Z. ZHANG, Y. ZHOU, L. XIE, A. YUILLE Adversarial Examples for Semantic Segmentation and Object Detection. The IEEE International Conference on Computer Vision (ICCV), Oct 2017.
- [25] Z. YUAN, Y. LU, Z. WANG, Y. XUE *Droid-Sec: Deep Learning in Android Malware Detection*. ACM SIGCOMM Computer Communication Review, volume 44, pages 371-372. ACM, 2014.



Matthieu Nugue obtained his Engineering Degree in Fluid Mechanics and Energy from *Polytech Marseille*, and his Master's degree in two-phase flow from Aix-Marseille University in 2015. From 2015 to 2018 he accomplished his PhD, co-funded by the CNES and ONERA, in the field of image processing in

the Rocket Propulsion Modeling Team (MPF) at ONERA. From 2018 to 2019 he worked on computer vision for robotics applications as a member of the innovation team at Audensiel. Since 2020, he has been working with the research and development team at Reminiz on deep learning for video analysis.



Jean-Michel Roche

Research engineer

Head of the "*Manufacturing & non-destructive testing of composite materials*" research team.

PhD thesis in Acoustics (ONERA).

He has been working on non-destructive inspection and structural health monitoring of materials and structures for over 10 years.



Guy Le Besnerais graduated from ENSTA in 1989 and obtained his PhD degree from *Université Paris Sud* in 1993. Since 1994 he works at ONERA – The French Aerospace Lab, currently with the Information Processing and System Department (DTIS) in Palaiseau (91) as Research Director. He obtained the

"Habilitation à diriger les recherches" (HDR) in 2008 and is affiliated to University of Paris-Saclay. His research activities include methodology for solving inverse problems, performance modeling for imaging measurement systems, and embedded vision for robotics applications.



Camille Trottier

Research engineer in the "Manufacturing & non-destructive testing of composite materials" research team.

(2013 - Université du Mans) Graduated in Optics and Nondestructive evaluations.

(2013 - 2016 / EDF) PhD thesis in ultrasound for non-destructive evaluations. (2016 - 2018 / THALES) Engineer specializing in signal processing and radar algorithms.

(2018 - to date/ ONERA) He has conducted research on the inspection, characterization and monitoring of damage in composite and metallic materials, using non-destructive techniques (thermal & ultrasonic).



Robin Devillers has been a research engineer at ONERA since 2013. He is in charge of experimental measurements associated with solid propellant studies, ranging from material characterization to combustion imaging. He worked previously for 4 years at the National Research Center of Canada in Ottawa, on

optical diagnostics for soot emission characterization in gas flares, with measurement campaigns performed in oil refineries in Mexico. He obtained his PhD from *École Centrale Paris* for his work on laser techniques for temperature measurements in engines at IFP. He has a MSc and an engineering degree from *École Centrale Paris*.



Julien Pichillou obtained his Energetics and Propulsion engineering diploma from INSA de Rouen in 2000, and has been a thermal engineer specializing in space launcher applications (thermal insulation development, thermal environment definitions) for 14 years. Since 2015 he has been employed as a

solid propulsion engineer in the CNES launcher directorate, working on P120C solid rocket motor development and on thrust oscillations in SRM.



Adrien Chan-Hon-Tong has been a researcher in machine learning at ONERA since 2014 in the IVA team. He has worked from 2014 to 2018 on the development of an object detection algorithm, working on satellite images and/or UAV videos (for military purposes). Since 2018, he has been focusing on safe-

ty issues raised by machine learning on critical platforms (autonomous driving, lethal autonomous weapon systems): performance evaluation of operational conditions, robustness to adversarial attack and data poisoning.



Alexandre Boulch received his engineering degree from the *École Polytechnique*, his M.Sc in Computer Vision and Machine Learning from ENS Cachan and his PhD from the Eastern Paris University (UPE) with a thesis focusing on geometric and semantic reconstruction for buildings. His research fields are

computational geometry and machine learning for remote sensing data, including Lidar and satellite imagery.



Antoine Hurmane graduated (Engineering Diploma) from the French Institute for Advanced Mechanics (IFMA), and received a Master Degree from Blaise Pascal University (Clermont- Ferrand France) in 2011. He received his PhD in mechanics from the University of Technology of Compiègne in 2015. Since then he has

been working at ONERA – The French Aerospace Lab as a research scientist in the Composite Materials and Structures Department. His research includes the development of damage and failure approaches for composite materials, and experimental characterization of damage and failure of composite materials.

