

Deep Learning in Medical Imaging: fMRI Big Data Analysis via Convolutional Neural Networks

Extended Abstract

Amirhessam Tahmassebi*
Department of Scientific Computing,
Florida State University
Tallahassee, Florida
atahmassebi@fsu.edu

Amir H. Gandomi
School of Business, Stevens Institute
of Technology
Hoboken, New Jersey
a.h.gandomi@stevens.edu

Ian McCann
Department of Scientific Computing,
Florida State University
Tallahassee, Florida
iam10@my.fsu.edu

Mieke H.J. Schulte
Addiction, Development, and
Psychopathology (ADAPT) Lab,
Department of Psychology, University
of Amsterdam
Amsterdam, The Netherlands
miekehjschulte@gmail.com

Anna E. Goudriaan
Amsterdam Institute for Addiction
Research, Department of Psychiatry,
Academic Medical Center, University
of Amsterdam
Amsterdam, The Netherlands
a.e.goudriaan@amc.uva.nl

Anke Meyer-Baese
Department of Scientific Computing,
Florida State University
Tallahassee, Florida
ameyerbaese@fsu.edu

ABSTRACT

This paper aims at implementing novel biomarkers extracted from functional magnetic resonance imaging (fMRI) images taken at resting-state using convolutional neural networks (CNN) to predict relapse in heavy smoker subjects. In this regard, two classes of subjects were studied. The first class contains 19 subjects that took the drug N-acetylcysteine (NAC), and the second class contains 20 subjects that took a placebo. The subjects underwent a double-blind smoking cessation treatment. The resting-state fMRI of the subjects' brains were recorded through 200 snapshots before and after the treatment. The relapse data was assessed after 6 months past the treatment. The data was pre-processed and an undercomplete autoencoder along with various similarity metrics was developed to extract salient features that could differentiate the pre and post treatment images. Finally, the extracted feature matrix was fed into robust classification algorithms to classify the subjects in terms of relapse and non-relapse. The XGBoost algorithm with 0.86 precision and an AUC of 0.92 outperformed the other classification methods in prediction of relapse in subjects.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning approaches; Neural networks; Supervised learning by classification; Feature selection; Cross-validation;**

* Corresponding Author

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

PEARC '18, July 22–26, 2018, Pittsburgh, PA, USA

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6446-1/18/07.

<https://doi.org/10.1145/3219104.3229250>

KEYWORDS

Deep Learning, Convolutional Neural Network, Autoencoder, fMRI, Big Data

ACM Reference Format:

Amirhessam Tahmassebi, Amir H. Gandomi, Ian McCann, Mieke H.J. Schulte, Anna E. Goudriaan, and Anke Meyer-Baese. 2018. Deep Learning in Medical Imaging: fMRI Big Data Analysis via Convolutional Neural Networks: Extended Abstract. In *PEARC '18: Practice and Experience in Advanced Research Computing, July 22–26, 2018, Pittsburgh, PA, USA*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3219104.3229250>

1 INTRODUCTION

Tobacco use remains the single largest preventable cause of death and disease in the United States. Based on the Centers for Disease Control (CDC)¹ report, in 2015, 36.5 million adults in the United States smoked cigarettes and more than 16 million Americans live currently with a smoking-related disease. In addition to this, every year more than 480,000 Americans die for smoking cigarettes.

The scientific goal of this study is to develop a cessation treatment using a compound that will reduce a patient's dependence on nicotine [3]. NAC as one of the potentially effective compounds was used in this study. NAC is a derivative of the amino acid cysteine pro-drug which is approved as a mucolytic agent and an acetaminophen antidote. NAC restores the basal level of glutamate in the accum-bens which may reduce the drug seeking behavior. This study aims at showing that NAC affects brain functions related to addiction [6][8]. Therefore, it would be possible to predict relapse in subjects i.e. after 6 months of treatment. In this study, CNNs were used to implement an autoencoder to extract features related to the high activity areas of brain where more oxygen-rich blood is flowing and fMRI is able to map these areas. Then, several similarity metrics were used to compare pre-treatment (baseline) and post-treatment reduced images. Finally, the extracted feature matrix was fed into various machine learning algorithms for classification.

¹<https://www.cdc.gov>

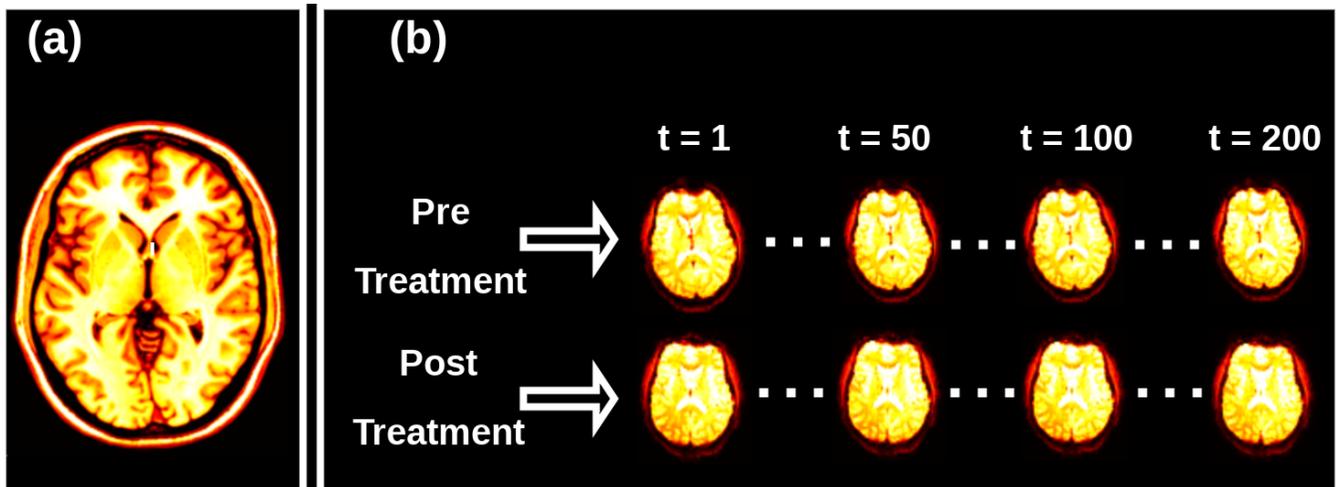


Figure 1: (a) Axial MRI slice of anatomical scan of brain of a subject with size of $240 \times 240 \times 220$. (b) Axial MRI slices of pre-treatment and post-treatment functional scans with size of $80 \times 80 \times 37$ through 200 snapshots.

2 DATA

The data acquisition was done at the Spinoza Center of the University of Amsterdam which is equipped with a 3.0 T Intera MRI scanner (Philips Health care, Best, The Netherlands) with a 32-channel SENSE head coil using 39 smokers: 19 of them received the drug NAC, and the rest received a placebo for two weeks. Anatomical and functional MRI scans of the subjects' brain at resting-state were taken at baseline (before treatment), and after two weeks of NAC treatment and the relapse data were assessed at six months past the NAC treatment [8]. A threshold of 10 cigarettes was chosen as the cut-off and according to the relapse data, 26 subjects relapsed and 13 subjects completely stopped smoking. Two hundreds 3D functional MRI scans of each subject's brain of size $80 \times 80 \times 37$, and one 3D anatomical scan of size $240 \times 240 \times 220$ were given in 4D spatio-temporal NIFTI (Neuroimaging Informatics Technology Initiative) format. Figure 1 shows an axial MRI slice of (a) anatomical and (b) functional scans of brain of a subject.

The artifacts due to the long process of the scans, possible movements of the subjects, and physiological noise were pre-processed via the standard pipeline using the Statistical Parametric Mapping (SPM12) software to increase the blood oxygen-level dependent (BOLD) signal to noise ratio (SNR). The pre-processing stage include: (1) motion correction, (2) segmentation, (3) realignment, (4) temporal slice timing, (5) smoothing, (6) normalization, and (7) co-registration [11][13][14][15][16][10].

3 METHODOLOGY

In this paper, an autoencoder consisting of multiple convolutional layers was developed to learn the features of the fMRI images for each subject in order to predict the relapse. An autoencoder is a neural network which is trained to attempt to copy its input to its output using two parts: (1) an encoder function $h = f(x)$, and (2) a decoder function which reproduces a reconstruction $x' = g(h)$ of the input [4][9]. Figure 2 presents the general schematic structure of an autoencoder. However, not in all cases the output of the decoder

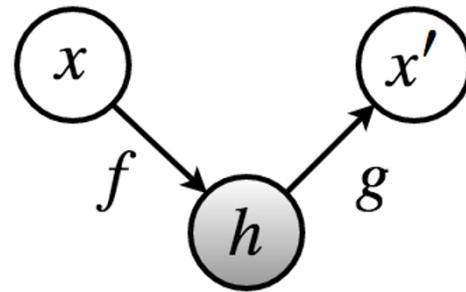


Figure 2: The general schematic structure of an autoencoder, mapping an input x to reconstruction x' via code h . The two essential components are: (1) encoder f which maps the input x to h , and (2) decoder which maps h to x' .

is the point of interest. In this study, it is desired that the trained autoencoder would extract some salient properties from the MRI images that could be used to predict the relapse in subjects. One of the possible ways is to employ undercomplete autoencoders by applying constraints on the input x to have smaller dimension. In this way, salient features can be extracted from the full dimension input ($80 \times 80 \times 37$) with a smaller dimension (i.e. $10 \times 10 \times 8$). In fact, learning an undercomplete representation forces the autoencoder to capture the most salient features of the training data [4].

The developed pipeline in this paper was written in Python employing various libraries including Keras, TensorFlow, Numpy, Nibabel, and Scikit-Learn [1][2][5][7]. As shown in Table 1, the developed autoencoder contains six 2D convolutional layer with the same padding. In fact, the encoder includes the first five convolutional layer using a linear rectifier (ReLU) as the activation function and a Sigmoid function used as the activation function of the last convolutional layer (decoder). A stride size of (3×3) , a

Table 1: The autoencoder layer settings.

Layer Type	Output Shape	# of Parameters
Input Image	$80 \times 80 \times 37$	0
Conv2D	$80 \times 80 \times 16$	5,344
MaxPooling2D	$40 \times 40 \times 16$	0
Conv2D	$40 \times 40 \times 8$	1,160
MaxPooling2D	$20 \times 20 \times 8$	0
Conv2D	$20 \times 20 \times 8$	584
MaxPooling2D	$10 \times 10 \times 8$	0
Conv2D	$10 \times 10 \times 8$	584
UpSampling2D	$20 \times 20 \times 8$	0
Conv2D	$20 \times 20 \times 8$	584
UpSampling2D	$40 \times 40 \times 8$	0
UpSampling2D	$80 \times 80 \times 8$	0
Conv2D	$80 \times 80 \times 37$	2,701
Total Trainable Parameters		10,957

pool size of (2×2) , and a sample size of (2×2) were used in all convolutional, max pooling, and up sampling layers, respectively. Binary cross-entropy was used as the loss function and the Adadelta algorithm which is robust to sparsity was employed to optimize the hyper-parameters [9].

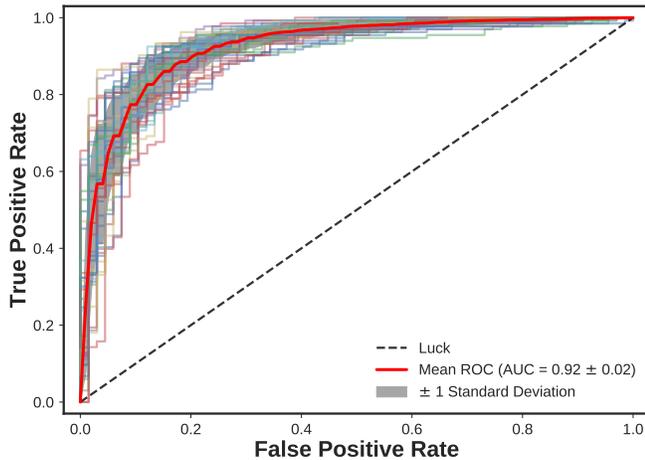


Figure 3: ROC curves for classification using XGBoost employing leave-one-out cross-validation. The lighter curves demonstrate the ROC for each fold, the red curve illustrates the mean value of the lighter curves, and the shaded gray area shows the confidence interval of the classification.

The autoencoder was applied on both pre-treatment and post-treatment scans. The compressed images after the third MaxPooling2D layer with a size of $(10 \times 10 \times 8)$ were fed into eight similarity comparison metrics including (1) correlation coefficient (CC), (2) correlation ratio (CR), (3) L_1 -norm based correlation ratio (CRL1), (4) mutual information (MI), (5) normalized mutual information (NMI), (6) Euclidean distance (ED), (7) Dice coefficient (DC), and (8) Jaccard coefficient (JC). It was desired to extract salient features

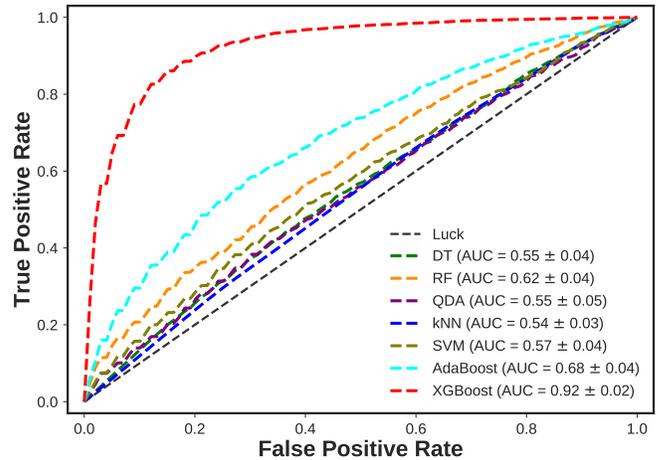


Figure 4: ROC curves for classification using various machine learning algorithms including DT, RF, KNN, SVM, QDA, AdaBoost, and XGBoost employing leave-one-out cross-validation.

via comparing each of the 200 snapshots of the pre-treatment and post-treatment for each subject. This procedure was resulted in a feature matrix of size 7800×8 . Then, the feature matrix was fed into robust classification algorithms along with Bayesian optimization to find the tuned hyper-parameters for each classifier.

As discussed, two NIFTI image files (pre-treatment and post-treatment) were given for each subject (total 78 image files). Each NIFTI image which contains 200 snapshots requires $\sim 100MB$ on disk. However, the NIFTI format contains multiple compression layers and reading the NIFTI file of the each subject into NumPy array turned into $\sim 1.3GB$ which was led into a big data challenge ($\sim 100GB$). The training process of the autoencoder on only one subject using a normal equipment (Intel Core i7 2.2 GHz \times 8 processor & 8 GB 1867 MHz DDR3 memory) took around 8 hours. Therefore, the developed pipeline was slightly changed to apply multiple computation stages in parallel. To overcome over-fitting, leave-one-out cross-validation was employed which also requires better equipment. Therefore, the developed pipeline ran on a high performance computing (HPC) machine using 5 nodes at the Research Computing Center (RCC) at the Florida State University. The wall-clock time was improved dramatically and the training and testing process including visualization stages were done in less than three days. It should be noted that the autoencoder model had to be trained 15600-times ($39 \times 2 \times 200$) which cost the major computational run-time of the project and it was almost impossible to be done using any normal computing equipment.

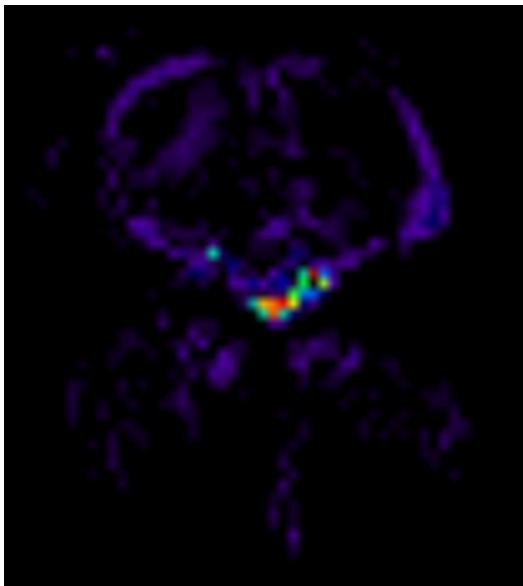
4 RESULTS

In this paper, various machine learning algorithms were used for the classification of the subjects into relapse and non-relapse classes employing leave-one-out cross-validation to overcome over-fitting. The salient features were extracted from fMRI scans using convolutional layers developed as an autoencoder. The feature matrix

Table 2: Classification metrics for several machine learning algorithms employing leave-one-out cross-validation.

Classifier	F1 Score	Precision	Recall	AUC
DT	0.68 ± 0.02	0.68 ± 0.02	0.67 ± 0.03	0.55 ± 0.04
RF	0.79 ± 0.01	0.68 ± 0.01	0.93 ± 0.02	0.62 ± 0.04
KNN	0.72 ± 0.02	0.69 ± 0.01	0.75 ± 0.03	0.54 ± 0.03
QDA	0.78 ± 0.01	0.67 ± 0.01	0.94 ± 0.02	0.55 ± 0.05
SVM	0.80 ± 0.0	0.66 ± 0.0	100.0 ± 0.0	0.57 ± 0.04
XGBoost	0.90 ± 0.01	0.86 ± 0.01	0.95 ± 0.02	0.92 ± 0.02
AdaBoost	0.80 ± 0.01	0.69 ± 0.01	0.94 ± 0.01	0.68 ± 0.04

was used as the input of the classifiers. The results of seven classification algorithms including (1) decision tree (DT), (2) random forest (RF), (3) quadratic discriminant analysis (QDA), (4) k-th (k=3) nearest neighbors (kNN), (5) support vector machine (SVM) with a radial basis function (RBF) kernel, (6) adaptive boosting (AdaBoost), and (7) extreme gradient boosting (XGBoost) were presented in Table 2. Figure 3 shows the receiver operating characteristic (ROC) curves for leave-one-out cross-validation using XGBoost. As seen in Table 2, XGBoost outperformed the other machine learning algorithms and showed reasonable results. As shown in Figure 3, all the presented ROC curves of the 39-folds are within the shaded gray confidence interval. In addition to this, the mean ROC curves of the employed classifiers were illustrated in Figure 4. Each curve is also presented along with an area under curve (AUC) metric. As the ROC curve gets closer to top left corner, the AUC value will be higher and the model would show better accuracy. In contrast, as the ROC curve gets closer to the dashed black line (Luck), it indicates that the predictions are more stochastic and cannot be generalized.

**Figure 5: The mapped extracted features by the developed autoencoder from a subject from the non-relapse class.**

Finally, the extracted features from a subject from the non-relapse class were mapped on the subject's brain template. The highest intensity that is indicated in red was seen close to the mesolimbic system which is in agreement with the previously published results [11][12][13][14].

REFERENCES

- [1] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. 2014. Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics* 8 (2014), 14.
- [2] François Chollet et al. 2015. Keras. (2015).
- [3] Michael C Fiore, Thomas E Novotny, John P Pierce, Gary A Giovino, Evridiki J Hatzianandreu, Polly A Newcomb, Tanya S Surawicz, and Ronald M Davis. 1990. Methods used to quit smoking in the United States. *Jama* 263, 20 (1990), 2760–5.
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.
- [5] Krzysztof Gorgolewski, Christopher D Burns, Cindee Madison, Dav Clark, Yaroslav O Halchenko, Michael L Waskom, and Satrajit S Ghosh. 2011. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front Neuroinform* 5 (08 2011). <https://doi.org/10.3389/fninf.2011.00013>
- [6] Steven D LaRowe, Pascale Mardikian, Robert Malcolm, Hugh Myrick, Peter Kalivas, Krista McFarland, Michael Saladin, Aimee McRae, and Kathleen Brady. 2006. Safety and tolerability of N-acetylcysteine in cocaine-dependent individuals. *The American journal on addictions* 15, 1 (2006), 105–110.
- [7] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.
- [8] MHJ Schulte, AE Goudriaan, AM Kaag, DP Kooi, W van den Brink, RW Wiers, and L Schmaal. 2017. The effect of N-acetylcysteine on brain glutamate and gamma-aminobutyric acid concentrations and on smoking cessation: A randomized, double-blind, placebo-controlled trial. *Journal of Psychopharmacology* 31, 10 (2017), 1377–1379. <https://doi.org/10.1177/0269881117730660>
- [9] Amirhessam Tahmassebi. 2018. iDeepLe: Deep Learning in a Flash. In *Disruptive Technologies in Information Sciences*, Vol. 10652. International Society for Optics and Photonics, 106520S. <https://doi.org/10.1117/12.2304418>
- [10] Amirhessam Tahmassebi, Ali Moradi Amani, Katja Pinker-Domenig, and Anke Meyer-Baese. 2018. Determining disease evolution driver nodes in dementia networks. In *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Vol. 10578. International Society for Optics and Photonics, 1057829. <http://doi.org/10.1117/12.2293865>
- [11] Amirhessam Tahmassebi, Amir H Gandomi, Ian McCann, Mieke HJ Schulte, Lianne Schmaal, Anna E Goudriaan, and Anke Meyer-Baese. 2017. An evolutionary approach for fMRI big data classification. In *Evolutionary Computation (CEC), 2017 IEEE Congress on. IEEE*, 1029–1036. <http://doi.org/10.1109/CEC.2017.7969421>
- [12] Amirhessam Tahmassebi, Amir H Gandomi, Ian McCann, Mieke HJ Schulte, Lianne Schmaal, Anna E Goudriaan, and Anke Meyer-Baese. 2017. fMRI Smoking Cessation Classification Using Genetic Programming. In *Workshop on Data Science meets Optimization*. http://ds-o.org/images/Workshop_papers/Gandomi.pdf
- [13] Amirhessam Tahmassebi, Amir H. Gandomi, and Anke Meyer-Baese. 2017. High Performance GP-Based Approach for fMRI Big Data Classification. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact (PEARC'17)*. ACM, New York, NY, USA, Article 57, 4 pages. <https://doi.org/10.1145/3093338.3104145>
- [14] Amirhessam Tahmassebi, Amir H Gandomi, Mieke HJ Schulte, Anna E Goudriaan, Simon Y Foo, and Anke Meyer-Baese. 2018. Optimized Naive-Bayes and Decision Tree Approaches for fMRI Smoking Cessation Classification. *Complexity* 2018 (2018). <https://doi.org/10.1155/2018/2740817>
- [15] Amirhessam Tahmassebi, Katja Pinker-Domenig, Georg Wengert, Marc Lobbes, Andreas Stadlbauer, Francisco J Romero, Diego P Morales, Encarnacion Castillo, Antonio Garcia, Guillermo Botella, et al. 2017. Dynamical graph theory networks techniques for the analysis of sparse connectivity networks in dementia. In *Smart Biomedical and Physiological Sensor Technology XIV*, Vol. 10216. International Society for Optics and Photonics, 1021609. <https://doi.org/10.1117/12.2263555>
- [16] Amirhessam Tahmassebi, Katja Pinker-Domenig, Georg Wengert, Marc Lobbes, Andreas Stadlbauer, Norelle C Wildburger, Francisco J Romero, Diego P Morales, Encarnacion Castillo, Antonio Garcia, et al. 2017. The driving regulators of the connectivity protein network of brain malignancies. In *Smart Biomedical and Physiological Sensor Technology XIV*, Vol. 10216. International Society for Optics and Photonics, 1021605. <https://doi.org/10.1117/12.2263554>