MOX-Report No. 85/2021

Recurrence-specific supervised graph clustering for subtyping Hodgkin Lymphoma radiomic phenotypes

Cavinato, L., Gozzi, N., Sollini, M., Carlo-Stella, C., Chiti, A.,

& Ieva, F.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-1179-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/EMBC46164.2021.9629625

Recurrence-specific supervised graph clustering for subtyping Hodgkin Lymphoma radiomic phenotypes

L. Cavinato[1], N. Gozzi[2], M. Sollini[3], C. Carlo-Stella[3], A. Chiti[3], F. Ieva [1]
[1]Department of Mathematics, Politecnico di Milano, Milan, Italy
[2]Nuclear Medicine, Humanitas Research Hospital, Pieve Emanuele, Italy
[3]Department of Biomedical Sciences, Humanitas University, Rozzano, Italy
[1]{lara.cavinato, francesca.ieva}@polimi.it
[2]{noemi.gozzi}@humanitas.it
[3]{martina.sollini, carmelo.carlostella, arturo.chiti}@hunimed.eu

***Abstract*— The prediction at baseline of patients at high risk for therapy failure or recurrence would significantly impact on Hodgkin Lymphoma patients treatment, informing clinical practice. Current literature is extensively searching insights in radiomics, a promising framework for high-throughput imaging feature extraction, to derive biomarkers and quantitative prognostic factors from images. However, existing studies are limited by intrinsic radiomic limitations, high dimensionality among others. We propose an exhaustive patient representation and a recurrence-specific multi-view supervised clustering algorithm for estimating patient-to-patient similarity graph and learning recurrence probability. We stratified patients in two risk classes and characterize each group in terms of clinical variables.**

I. Introduction

Hodgkin Lymphoma (HL) is an haematological malignancy of the lymphatic system, part of the immune system. Although the majority of patients are cured, up to 15% of patients with early stage and up to 30% of patients with advanced stage of HL will be primary refractory or experience recurrence, with percentages increasing with therapy progression [1]. For this reason, the identification of cases at high risk for first-line therapy failure or recurrence would significantly impact on HL patients treatment, resorting to alternative therapy planning. Today, prognostic stratification guiding the therapeutic strategy in HL relies mainly on staging and clinical risk factors, such as the presence of B symptoms [2] and bulky disease [3]. Unfortunately, current staging systems and prognostic factors provide limited information about the lymphoma biology and fail in identifying refractory HL patients at baseline [4].

Recently, quantitative image mining, namely radiomics, has emerged. It has been shown to be promising in predicting patients' outcome in several cancers, including HL, shedding lights on detecting refractory/relapsing patients and long-term responders [5].

Radiomics consists of the employment of high-throughput methods, extracting numerical features from regions of interest, e.g. cancer lesions, thus transforming medical images into easy-to-handle matrix data [6]. Radiomic features, also known as texture features, can be divided into conventional (i.e. Standard Uptake Value) and higher-order variables. For instance, such variables include histogram-derived variables, shape-derived variables, Gray Level Co-occurrence (GLCM) matrix-derived variables, Gray Level Run Length (GLRLM) matrix-derived variables, Neighbour Gray Level Difference (NGLDM) matrix-derived variables and Gray Level Zone Length (GLZLM) matrix-derived variables. According to this taxonomy, every lesion is described with different views. Together with genomics, radiomics undoubtedly represents the new frontier of cancer research, paving the way for personalized patients' treatment. However, several methodological aspects have not been clarified yet, preventing its translational application in clinical practice. These include its close-source nature, sensitivity to acquisition settings and reconstruction parameters, lack of interpretability and methodological biases [7]. In addition, preliminary findings have shown spatial intra-tumor heterogeneity to be fundamental in understanding tumor severity and evolution, impacting on pre-treatment prognosis [8]. A lack of consensus about quantitative definition of heterogeneity has led to the need of evaluating each lesion severeness and, with it, to the issue of exhaustively representing patients' disease [9].

On the other hand, predicting the refractory/relapsing class is often limited to the number of minority class observations, i.e. relapsing patients, such that both fixed-effects and mixed-effects classification models and even classification models optimized for imbalanced data eventually fail. Patients stratification in an unsupervised setting could provide a different perspective for deducing baseline insights about the patients' disease progression, in order to implement effective and personalized treatment strategies. Although not yet explored in radiomics, such questions are usually addressed with unsupervised clustering approaches from the integration of quantitative and qualitative data of patients [10] [11]. Nevertheless, the patients subtyping approach can be far improved when considering the clinical relapsing response variables as a regularization factor for clusters' building, with the aim of discovering clinically more relevant cancer profiles.

The majority of cancer subtype identification models have been proposed in genomic literature, where high-dimentional, multi-omic datasets need to be properly processed for genetic patients' profiling. Such methods range from the simple combination of single-omic distance matri-

ces (LRAcluster [12]), to the fusion of single-omic similarity networks accounting for data complementarity (Similarity Network Fusion [13]), to Cancer Integration via Multi-kernel Learning [14]. More recent works have focused on both multi-omic views and dimensionality reduction, such as Joint and Individual Variation Explained [15], Monte Carlo consensus clustering [16], Multi-Omics Factor Analysis [17]. and Deep learning based integration approach [18]. Moreover, the latest trend is to overcome pure unsupervised approaches and perform cancer subtyping in a supervised way, as to guarantee a good clustering performance for patient stratification while accounting for survival time and clinically relevant information (Survival Supervised Graph Clustering, S2GC in short [19]).

The contribution of this work is two-fold. First, we propose a multi-lesion patient's representation accounting for clinical and personal information, as well as patient-wise radiomic description across his/her multi lesion information, in terms of mean and dispersion values. We then exploit the multi-view approach of the S2GC [19] to propose a recurrence-specific supervised clustering built upon such patients' representation, for indentifying clinically-relevant HL phenotypical subtypes. Moreover, thanks to the multi-view approach, we are able to interpret groups evaluating the separate contribution of all radiomic feature groups, i.e. the above-mentioned views, as well as their consistency/agreement throughout observations. This would also enable the radiomic framework to scale up to many more extracted features.

II. Data collection

One hundred and thirty-six patients diagnosed with HL were retrospectively included in the study. Personal and clinical disease information was recorded, including staging, sex, age, presence of B-symptoms, types of administered therapy (e.g. chemotherapy and radiotherapy), number of nodal lesions, number of extranodal lesions, distribution of lesions' volume (mean, standard deviation) as well as total tumor volume. Lesions' radiomic information were retrieve from semi-automatic segmentation of $[^{18}F]$FDG-PET/CT by experienced radiologists. LIFEx (www.lifexsoft.org, [20]) was employed for both $[^{18}F]$FDG-avid lesions' segmentation and feature extraction.

The study, performed in accordance with the Declaration of Helsinki, was approved by the ethics committee of IRCCS Humanitas Research Hospital. The signature of a specific informed consent and legal requirements of clinical trials were waived in view of the observational retrospective study design.

III. Methods

A. Patient representation

As previously stated, although radiomic lesion description allows to quantitatively exploit imaging information in matrix data format, there are several obstacles that prevent radiomics from taking root in clinical practice and research.

In multi-lesions cancers, like HL, in addition to the intrinsic problems of radiomics, an exhaustive patient representation needs to be built. In fact, each lesion is associated with a radiomic description vector that cannot be treated as independent from the others as it belongs to the same patient. It follows that the radiomic vectors need to be synthesized in one-row representation prior to be juxtaposed to the vector of patients' clinical covariates.

We dealt with these two issues by treating lesions' radiomic features as different samples of the same patient. Patient-wise radiomic profile was obtained by averaging all radiomic features of lesions belonging to the same subject with equal lesions' weights, such that each lesion contributes equally to its patient representation. Of course, different weighting strategies could be implemented and are left for future experiments.

Intra-patient variability has been described by multivariate point-cloud dispersion of patients' lesions and such index has been used to represent patients' intra-tumor heterogeneity. Dispersion has been computed for all, only nodal and only extranodal lesions. Radiomic, volume, dispersion and personal variables have been thus concatenated to build the overall patient representation, divided into *omic views*, i.e. clinical, histogram, shape, SUV, GLCM, GLRLM, NGLDM and GLZLM groups.

B. Recurrence-specific survival clustering model

In order to find clinically-relevant clusters in HL patients, two steps have been performed: a recurrence-specific graph embedding of patients has been estimated to compute the patient-wise similarity matrix S and a spectral clustering algorithm has been applied to S for class categorization.

The recurrence-specific graph embedding has been built upon the optimization of the following objective function [19]:

$$\begin{aligned}
&\min_{w;S} \sum_{k=1}^{m} \left(-\sum_{i=1}^{n} \delta_i \left(X_i^k w^k - log \sum_{j \in R_i} exp(X_j^k w^k) \right) \right) \\
&+ \lambda \sum_{k \neq j} \|X^k w^k - X^j w^j\|_2^2 + \eta \sum_{k=1}^{m} \|w^k\|_1 \\
&+ \min_{S} \gamma \sum_{i=1}^{n} \sum_{j=1}^{n} (\|X_i - X_j\|^2 + \|X_i w - X_j w\|^2) S_{i,j} + \mu S_{i,j}^2 \\
&s.t. \sum_{j}^{n} S_{I,j} = 1, S_i \succeq 0; i = 1, 2, \ldots, n.
\end{aligned} \tag{1}$$

The first term entails the minimization of the recurrence prediction according to the cumulative Cox proportional Hazard method over all views m, where X_i is the p-dimensional patient's feature vector and w the coefficient variables that need to be estimated; the second term is the co-regularization term to perform the pair-wise agreement integration between the prediction of different views, with control parameter λ; the third term is the $l1$ sparsity regularization term for managing high data dimensionality, with control parameter η; the final term performs the recurrence-based graph learning

S accounting for the distances between observations, i.e. patients, in terms of radiomic views and predicted recurrence time; n is the total number of patients and each entry of S ($S_{i,j}$) indicates the similarity between X_i and X_j. The larger $S_{i,j}$, the smaller the distance between radiomic covariates of patients and their recurrence predictions. Indeed, the graph has patients on each node and each edge represents the similarity between the two patients it connects.

On one hand, we find the best w such that a model for the cancer recurrence time is learnt for each patient. On the other hand, similarity graph S is estimated on both clinical-radiomic and recurrence time patients' information. These two tasks are thus jointly performed, in a mutual empowerment fashion. As the model simultaneously computes two tasks, an alternating optimization algorithm has been used to optimize the corresponding problem, iteratively solving one task while keeping the other fixed [19].

Once recurrence-specific embedding graph is obtained as described above, spectral clustering has been performed to single out k clinically-relevant HL subtypes, i.e. clusters [21].

C. Fitting recurrence curves

The spectral clustering is performed on the patient-to-patient similarity graph and the best number of clusters k is searched. A recurrence curves is thus fitted on each recurrence-specific cluster, i.e. risk class, according to Kaplan-Meier estimator. Each KM recurrence curve is defined as the probability of recurring in a given length of time while considering time in many small interval [22].

The log-rank test (Mantel-Cox test) is used to evaluate the statistical difference between the groups in terms of recurrence time and p-values, while Hazard Ratio (HR) and median recurrence times were annotated for clusters characterization. Confidence Interval (CI) is also provided. Finally, median recurrence time is defined as the time at which recurrence curve intercepts the 50% probability line.

IV. Results

As k was set to 2 according to maximization of similarity criterion, two clusters were obtained, dividing the patients' population in two recurrence-specific risk classes.

In Figure 1, the KM recurrence probability curves for cluster 1 (in red) and cluster 2 (in blue) are displayed.

The logrank test on the two curve tested a significant difference between the two groups' recurrence rate (p=1.94e-06). Moreover, Hazard Ratio (HR) has been computed and resulted to be 0,163 with 95% Confidence Interval [0.0496, 0,537]. Finally, median recurrence time was 659 for group 2, infinity for group 1. Since the HR is the ratio of the risk of recurrence in group 1 with respect to the risk of recurrence in group 2, in HL application it results that group 2 is at high risk for recurrence while membership to group 1 seems to prevent from recurrence in Hodgkin Lymphoma. This is further confirmed by median recurrence time, which is little less than two years for group 2, while group 1 seems not to experience recurrence.

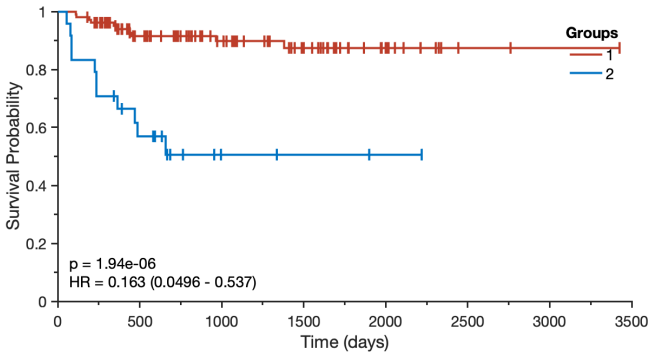

Fig. 1. Recurrence probability curves for cluster 1 (in red) and cluster 2 (in blue).

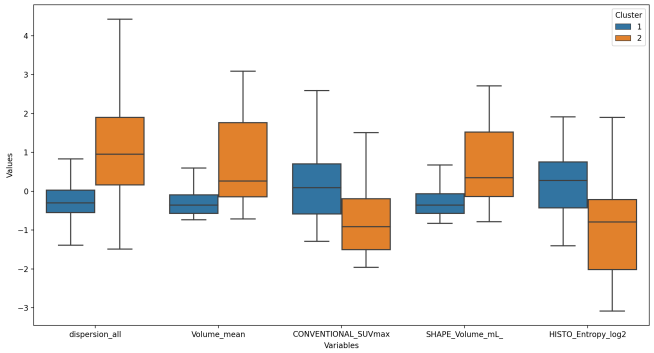

Fig. 2. Recurrence-specific cluster characterization for cluster 1 (blu) and cluster 2 (orange): cluster 1 is at high risk of recurrence and cluster 2 at lower risk.

V. Discussion

The two groups were assessed and a high risk of recurrence was noticed in cluster 2 while a lower risk was found in cluster 1. While characterizing the two clusters, a large number of radiomic features emerged as significant in discriminating the groups, as expected by construction.

Figure 2 shows the distributions of variables representing some of the different radiomic views in the two groups, namely patient-wise dispersion over all lesions (nodal and extranodal), mean volume of lesions, SUV_max values, shape values and energy of the first order histogram. Group-wise boxplots describe the differences of variables in patients.

These findings provide risk factors for HL recurrence and largely agree with those found in medical literature, proving the reliability of the model. Indeed, patients belonging to risk class 1, i.e. having lower risk of recurrence, exhibit statistically relevant lower lesions' dispersion, lower lesions' volume, more spherical lesions, and higher entropy within lesions with respect to patients belonging to risk class 2, who instead express higher dispersion and volume, lower entropy and more irregular lesions' shapes. As dispersion is intended as a proxy of intra-tumor heterogeneity, we notice how higher heterogeneous diseases are associated to poorer clinical outcomes, while lesion-specific entropy values are linked to more positive responses, as shown in other medical applications [23]. The employment of patient-wise lesions'

dispersion as intra-tumor heterogeneity proxy paves the way towards the assessment of tumor evolution in a personalized medicine fashion, beyond standard prognosis recommendations. Moreover, the bigger the whole tumor and the bigger the lesions, the higher the recurrence probability within short time. This again underlines the match of our model's findings with current guidelines in clinical research and practice, where patients with bigger lesions at baseline are prescribed to undergo radiotherapy along with chemotherapy, as they are flagged with bad prognosis. Finally, sphericity is known to correlate with clinical prognostic factors in other types of cancer [24], as our model highlighted in HL.

Beside the most intuitive variables described above, a multitude of higher order indexes have emerged as significant, making provision for radiomic features clinical interpretation in terms of cancer severity and therapy response. In fact, as current clinical literature struggles in finding a biological meaning of higher order radiomic variables, recurrence-specific cluster characterization could provide a robust and reliable tool for feature explanation.

VI. Conclusions

In this work we proposed a disease patients representation encompassing clinical and imaging information: clinical information included personal data as well as qualitative cancer characterization; lesions' radiomic information was summarized in terms of average radiomic profile and intra-patient lesions' dispersion, i.e. dissimilarity. We fed such patients' representation into a recurrence-specific multi-view supervised clustering algorithm for estimating patient-to-patient similarity graph and learning recurrence probability, leading to the stratification of patients in two risk classes. Patients subtypes were characterized according to their risk, and dispersion, volume, shape, SUV and entropy perfectly profiled non-recurrent and recurrent cancer phenotypes.

Thanks to its multi-view and supervised nature, such method opens the way for high dimensional radiomic data processing, while keeping strong interpretability qualities.

Acknowledgements

We acknowledge all personnel in Nuclear Medicine for the assistance during the PET/CT scans. We thank M. Biroli and F. Ricci for data collection.

References

[1] A. S. LaCasce, "Treating hodgkin lymphoma in the new millennium: Relapsed and refractory disease," *Hematological oncology*, vol. 37, pp. 87–91, 2019.

[2] R. Sharma, D. Cunningham, P. Smith, G. Robertson, O. Dent, and S. J. Clarke, "Inflammatory (b) symptoms are independent predictors of myelosuppression from chemotherapy in non-hodgkin lymphoma (nhl) patients–analysis of data from a british national lymphoma investigation phase iii trial comparing chop to pmitcebo," *BMC cancer*, vol. 9, no. 1, pp. 1–10, 2009.

[3] S. M. Ansell, "Hodgkin lymphoma: 2018 update on diagnosis, risk-stratification, and management," *American journal of hematology*, vol. 93, no. 5, pp. 704–715, 2018.

[4] A. Mottok and C. Steidl, "Biology of classical hodgkin lymphoma: implications for prognosis and novel therapies," *Blood*, vol. 131, no. 15, pp. 1654–1665, 2018.

[5] S. A. Milgrom, H. Elhalawani, J. Lee, Q. Wang, A. S. Mohamed, B. S. Dabaja, C. C. Pinnix, J. R. Gunther, L. Court, A. Rao *et al.*, "A pet radiomics model to predict refractory mediastinal hodgkin lymphoma," *Scientific reports*, vol. 9, no. 1, pp. 1–8, 2019.

[6] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.

[7] S. Rizzo, F. Botta, S. Raimondi, D. Origgi, C. Fanciullo, A. G. Morganti, and M. Bellomi, "Radiomics: the facts and the challenges of image analysis," *European radiology experimental*, vol. 2, no. 1, pp. 1–8, 2018.

[8] E. Sala, E. Mema, Y. Himoto, H. Veeraraghavan, J. Brenton, A. Snyder, B. Weigelt, and H. Vargas, "Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging," *Clinical radiology*, vol. 72, no. 1, pp. 3–10, 2017.

[9] M. Sollini, M. Kirienko, L. Cavinato, F. Ricci, M. Biroli, F. Ieva, L. Calderoni, E. Tabacchi, C. Nanni, P. L. Zinzani *et al.*, "Methodological framework for radiomics applications in hodgkin's lymphoma," *European Journal of Hybrid Imaging*, vol. 4, pp. 1–17, 2020.

[10] H. Muhammad, C. S. Sigel, G. Campanella, T. Boerner, L. M. Pak, S. Büttner, J. N. IJzermans, B. G. Koerkamp, M. Doukas, W. R. Jarnagin *et al.*, "Unsupervised subtyping of cholangiocarcinoma using a deep clustering convolutional autoencoder," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 604–612.

[11] S. Roy, A. P. Singh, and D. Gupta, "Unsupervised subtyping and methylation landscape of pancreatic ductal adenocarcinoma," *Heliyon*, vol. 7, no. 1, p. e06000, 2021.

[12] D. Wu, D. Wang, M. Q. Zhang, and J. Gu, "Fast dimension reduction and integrative clustering of multi-omics data using low-rank approximation: application to cancer molecular classification," *BMC genomics*, vol. 16, no. 1, p. 1022, 2015.

[13] B. Wang, A. M. Mezlini, F. Demir, M. Fiume, Z. Tu, M. Brudno, B. Haibe-Kains, and A. Goldenberg, "Similarity network fusion for aggregating data types on a genomic scale," *Nature methods*, vol. 11, no. 3, p. 333, 2014.

[14] R. Chen, "Supervised-unsupervised cancer subtyping based on multi-task learning," Ph.D. dissertation, State University of New York at Buffalo, 2020.

[15] E. F. Lock, K. A. Hoadley, J. S. Marron, and A. B. Nobel, "Joint and individual variation explained (jive) for integrated analysis of multiple data types," *The annals of applied statistics*, vol. 7, no. 1, p. 523, 2013.

[16] C. R. John, D. Watson, D. Russ, K. Goldmann, M. Ehrenstein, C. Pitzalis, M. Lewis, and M. Barnes, "M3c: Monte carlo reference-based consensus clustering," *Scientific reports*, vol. 10, no. 1, pp. 1–14, 2020.

[17] R. Argelaguet, B. Velten, D. Arnol, S. Dietrich, T. Zenz, J. C. Marioni, F. Buettner, W. Huber, and O. Stegle, "Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets," *Molecular systems biology*, vol. 14, no. 6, p. e8124, 2018.

[18] K. Chaudhary, O. B. Poirion, L. Lu, and L. X. Garmire, "Deep learning–based multi-omics integration robustly predicts survival in liver cancer," *Clinical Cancer Research*, vol. 24, no. 6, pp. 1248–1259, 2018.

[19] C. Liu, C. Wenming, S. Wu, W. Shen, D. Jiang, Z. Yu, and H. San Wong, "Supervised graph clustering for cancer subtyping based on survival analysis and integration of multi-omic tumor data," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020.

[20] C. Nioche, F. Orlhac, S. Boughdad, S. Reuzé, J. Goya-Outi, C. Robert, C. Pellot-Barakat, M. Soussan, F. Frouin, and I. Buvat, "Lifex: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity," *Cancer research*, vol. 78, no. 16, pp. 4786–4789, 2018.

[21] J. Liu, J. Han, C. Aggarwal, and C. Reddy, "Spectral clustering." 2013.

[22] J. M. Bland and D. G. Altman, "Survival probabilities (the kaplan-meier method)," *Bmj*, vol. 317, no. 7172, pp. 1572–1580, 1998.

[23] V. M. Eskov, A. A. Khadartsev, V. V. Eskov, O. E. Filatova, and D. U. Filatova, "Chaotic approach in biomedicine: individualized medical treatment," *Journal of Biomedical Science and Engineering*, vol. 2013, 2013.

[24] A. Davey, M. van Herk, C. Faivre-Finn, H. Mistry, and A. McWilliam, "Is tumour sphericity an important prognostic factor in patients with lung cancer?" *Radiotherapy and Oncology*, vol. 143, pp. 73–80, 2020.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

84/2021 Torti, A.; Galvani, M.; Urbano, V.; Arena, M.; Azzone, G.; Secchi, P.; Vantini, S.
Analysing transportation system reliability: the case study of the metro system of Milan

83/2021 Colasuonno, F.; Ferrari F.; Gervasio, P.; Quarteroni, A.
Some evaluations of the fractional p-Laplace operator on radial functions

78/2021 Bucelli, M.; Dede', L.; Quarteroni, A.; Vergara, C.
Partitioned and monolithic algorithms for the numerical solution of cardiac fluid-structure interaction

80/2021 Sollini, M., Bartoli, F., Cavinato, L., Ieva, F., Ragni, A., Marciano, A., Zanca, R., Galli, L., Pai
[18F]FMCH PET/CT biomarkers and similarity analysis to refine the definition of oligometastatic prostate cancer

81/2021 Massi, M.C.; Gasperoni, F.; Ieva, F.; Paganoni, A.
Feature Selection for Imbalanced Data with Deep Sparse Autoencoders Ensemble

79/2021 Ferraccioli, F.; Sangalli, L.M.; Finos, L.
Some first inferential tools for spatial regression with differential regularization

82/2021 Massi, M.C.; Ieva, F.
Learning Signal Representations for EEG Cross-Subject Channel Selection and Trial Classification

75/2021 Cicci, L.; Fresca, S.; Pagani, S.; Manzoni, A.; Quarteroni, A.
Projection-based reduced order models for parameterized nonlinear time-dependent problems arising in cardiac mechanics

73/2021 Marcinno, F.; Zingaro, A.; Fumagalli, I.; Dede', L.; Vergara, C.
A computational study of blood flow dynamics in the pulmonary arteries

77/2021 Guo, M.; Manzoni, A.; Amendt, M.; Conti, P.; Hesthaven, J.S.
Multi-fidelity regression using artificial neural networks: efficient approximation of parameter-dependent output quantities