AI-empowered m-health systems for Personalised Services and Digital Phenotyping

Franca Delmastro[1,*], Flavio Di Martino[1] and Mattia G. Campana[1]

[1] Institute for Informatics and Telematics (IIT), CNR, Pisa, 56100, Italy

Abstract

Health services personalization and early risk prediction represent the main research challenges in m-health systems, which can be achieved through the use of AI algorithms and tools applied to physiological and behavioural data collected by wearables and IoT devices in real-world settings. In this paper we present a summary of the results we obtained in our research activities in this area and future works, with particular attention to AI-empowered m-health systems as support for personalised rehabilitation services and malnutrition risk assessment, mobile sensing data analysis for disease detection and the identification of new health and behavioural markers that can support remote patient monitoring and the clinical practice.

Keywords

m-health, AI, wearables, sensor data

1. Introduction

The collection and analysis of mobile sensing data derived from personal and wearable devices open the way to the identification of predictive relationships between human habits, behaviour and health [1]. In addition, the same data, collected in-the-wild and for a medium or long period in a semi-continuous way, contribute to the generation of a Digital Phenotype [2] of the individual that can be used to predict risks and early identify digital markers of important diseases. This type of data includes daily routines, sleep patterns, physical mobility, nutrition, cognitive functioning, speech production, social interactions, and many others, which can be selectively collected based on the individual health profile. m-health applications, integrated with IoT and AAL systems, represent the main infrastructure of the monitoring systems, but today they must be enriched with AI algorithms to automatically detect specific health conditions, risky and adverse situations, and to implement personalized interventions. In addition, if the personal mobile device of the user is able to preliminary process this data directly on-board, it can avoid the transmission of a huge amount of sensitive data to the cloud, preserving the user's privacy and improving the system's trustworthiness.

In our research activity we focus on different aspects in this field: from the definition of customised clinical studies aimed at automatically defining personalised rehabilitation therapies based on physiological data analysis, to the collection and analysis of mobile sensing data for the fast screening of some diseases, and the integration of heterogeneous sensing data (both physiological and behavioural) aimed at identifying new digital markers for individual health and well-being conditions. These systems can be customised both for healthy ageing people, aimed at maintaining a good autonomy and quality of life, and for patients, in order to provide patient-centred and integrated care pathways.

The multidisciplinary characterisation of this research provides multiple impacts: (i) a technological impact with the definition of new AI-empowered decision support systems aimed to provide personalised feedback and therapies; (ii) a medical impact, through the identification of new digital physiological and behavioural markers that can support the clinical diagnosis and the remote patient monitoring; (iii) a social impact by supporting people with daily and unobtrusive monitoring instruments and personalised feedback, reducing the impact on the national health systems and allowing caregivers to predict health trajectories over time.

In addition to the predictive performance analysis of these systems, it is also essential to investigate how their results can be interpretable and explainable in order to improve their acceptance and validation in the clinical practice and the users' trust. To this aim we started also investigating the latest Explainable AI (XAI) methodologies with particular attention to those used for time series and tabular data [3].

In this works we present the main results of our research activity divided in 3 main areas:

- AI-empowered m-health systems for personalised rehabilitation and nutrition support in older adults

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy

*Corresponding author.
franca.delmastro@iit.cnr.it (F. Delmastro); flavio.dimartino@iit.cnr.it (F. D. Martino); mattia.campana@iit.cnr.it (M. G. Campana)
0000-0001-9233-2712 (F. Delmastro); 0000-0001-9778-0493 (F. D. Martino); 0000-0003-3732-4618 (M. G. Campana)

- Smartphone-embedded sensing for the early detection of specific diseases
- Definition of behavioural and physiological markers from wearables and environmental sensors aimed at identifying the health status and early predicting risk conditions for active and healthy ageing.

2. Stress Detection and Malnutrition Risk Assessment

Chronic and acute stress conditions and risk of malnutrition represent two important factors in frail older adults. We decided to focus part of the research activity on these two specific health domains since they represent daily conditions that can characterise both healthy and frail older adults and they can be also integrated with other behavioural parameters for the definition of a general health and well-being index.

As far as the automatic stress detection is concerned, we focused on the analysis of Heart Rate (HR), Heart Rate Variability (HRV) and Electrodermal Activity (EDA) as reference physiological stress sensing data that can be easily collected from wearable sensors, according to [4]. Recently, other studies have introduced also EEG, Blink Reflex through surface EMG, and eye tracking data analysis [5], but the related instruments are invasive and they are not easily accepted, even in clinical environments.

The identification of acute stress events are important in older adults, considering that they are commonly affected by chronic stress conditions in the management of their health status. Cognitive and motor rehabilitation activities, proposed to prevent physical and cognitive decline, can thus represent an additional stressful condition that could mitigate the positive effects of the therapy. Therefore, the analysis of physiological signals during the rehabilitation therapy execution can allow the definition of personalised protocols defined according to the automatic identification of the induced stress level.

We developed a m-health system aimed at collecting and analysing those data during specific rehabilitation protocols that has been validated and evaluated through a pilot study we conducted with a group of frail MCI older adults living in a long-term care (LTC) facility [6]. Each participant has been monitored through two wearable sensors (i.e., Zephyr BioHarness chest strap for HR and HRV and SHIMMER EDA sensor) during specific training sessions based on standard clinical cognitive training alternated by a light physical exercise by using a cycle-ergometer. The proposed m-health system has been designed to highlight the short-time improvements on the cognitive performances generated by the proposed physical exercise, and the related stress response. The first step has been to perform binary stress detection using several machine learning (ML) techniques exploiting the study protocol as implicit ground truth as stressful and non-stressful events' labels. No self-report questionnaires or clinical evaluations have been used to track the user perceived stress during the study. Then, in order to improve the stress detection resolution in terms of number of detected stress levels, we evaluated few prominent ML and deep learning models for solving time series regression tasks [7]. We performed this analysis also on two available multimodal physiological dataset for stress and affect detection: WESAD [8] and SWELL [9]. The considered models have been trained using stress scores derived from different clinical questionnaires aimed at detecting the perceived stress level severity over the time during multiple conditions. Models have been evaluated on each subject separately by using Leave-One-Subject-Out (LOSO) cross validation scheme to test their generalisation capabilities. The obtained results demonstrated that the selected predictive models as well as the used stress ground truth may provide an accurate and detailed individual stress tracking in most cases. In addition, those models could be integrated into a DSS module of the m-health solution for online stress monitoring and they could be further evaluated on new datasets collected from rehabilitation sessions of different target users, including neuro-degenerative patients (e.g., Parkinson).

Malnutrition is a serious and prevalent health problem in the older population, and especially in hospitalised or institutionalised subjects and an accurate and early risk detection is essential for prevention. Also in this case AI-empowered m-health systems may lead to important improvements in terms of a more automatic, objective, and continuous monitoring and assessment. We addressed the challenges in this health domain by exploiting a simple yet efficient m-health application, called DoEatWell (DEW) [10], designed to collect information about nutritional preferences and intake integrated with body composition data collected by a smart bioimpedance scale (i.e., body weight, body mass index, basal metabolic rate, bone mass, body fat, water, and muscle percentage). DEW has been originally designed to be used by LTC care givers, but it can be customised also for independent living scenarios. It has been deployed in a LTC facility in Italy from March 2018, and we collected and analysed data in multiple trial periods (approximately 2 years) from a total sample of 42 subjects. Feature engineering and extraction has been performed in collaboration with a medical specialist in order to define suitable, multimodal input predictors. We first focused on estimating the daily intake of the major macro-nutrients, namely cereals, animal proteins, vegetables, and fruit, which also represent the main components of the "*Healthy Eating Plate*" [11] for the Mediterranean diet. For what concerns body composition assessment, we initially computed only the fat mass index, which is highly correlated with the malnutrition risk. Finally, we also included some behavioural parameters characterizing the completeness and variability of the main daily meals. This data analysis is in line with the information generally requested by the reference clinical screening tools, yet providing a more objective and quantitative assessment. In order to validate our approach as a supervised learning task, we relied on the availability of a periodical clinical screening made by a healthcare professional on a monthly basis through the Mini Nutritional Assessment Short-Form (MNA-SF) tool [12].

We investigated the performances of 6 benchmark ML algorithms, namely Logistic Regression (LR) with Least Absolute Shrinkage and Selection Operator (LASSO) regularisation, Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Classification And Regression Tree (CART), Random Forest (RF), and AdaBoost (AB), as well as different data imbalance management techniques ranging from imbalanced training, dataset oversampling (SMOTE), to cost-sensitive learning. We also included ad-hoc algorithms to directly classify from imbalanced data, such as Random Undersampling Boosting (RUSBoost) and Balanced Random Forest (BRF). We implemented a repeated, stratified random hold-out partition (70% training, 30% test), then we performed hyperparameter tuning through 10-fold CV along with Bayesian optimization algorithm at each iteration. The resulting best model configuration has been evaluated on the held back test set, considering a comprehensive set of evaluation metrics to correctly account for data imbalance.

Obtained results show that tree ensemble models (i.e., RF, AB, and RUSBoost) provide high accuracy and recall in detecting individual nutritional status by combining nutritional intake, dietary habits, and body composition data, with median values of 94% and 92%, respectively. Results also indicate that cost-sensitive learning is the most effective method to deal with data imbalance in our case study, as it also pushes the other models close to the best performers. Instead, not considering body composition data, which is available only in a population subgroup, the classification performances worsen, even if the sample increases.

We also extended the previous analysis by integrating additional heterogeneous information including demographic, anamnestic, and clinical data (e.g., age, chronic disease, therapies) stored in the DEW user health profile. The obtained classification results slightly improved in case body composition is considered, maintaining RF and gradient boosting models as the best ones, with accuracy from 95% to 96% and $F1$ from 93% to 94%. In addition, the new models combining only clinical and nutritional data highlighted significant performance gains, with accuracy gain ranging from $+8.4\%$ to $+17.3\%$, whereas F1 increases from $+16.7\%$ to $+22.5\%$, thus reaching competitive results also on a larger population.

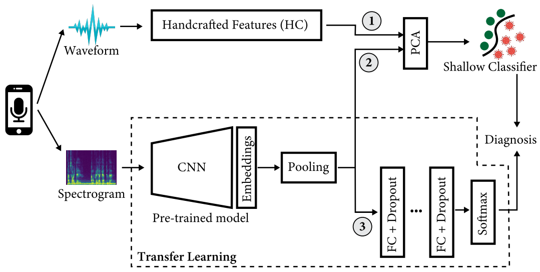

Figure 1: The experimented COVID-19 detection approaches: (1) extraction of handcrafted acoustic features from the audio waveform, which are then classified by a shallow ML model; (2) usage of a pre-trained deep learning model as features extractor, in series with a shallow model to classify the deep audio embeddings; and (3) fine-tuning of the pre-trained deep model for both features extraction and classification.

Finally, we also investigated the utility of XAI techniques to validate the proposed solution, both in terms of objective assessment of the agreement between explanations generated by different methods for each model separately, and a preliminary clinical validation to verify that the input-target relationships learned for the most relevant predictors are in line with the current evidence-based assessment.

Specifically we analysed the following XAI techniques: SHAP, LIME, Anchors, and feature permutations. The model-specific explanation consistency assessment highlighted that each model privileges very similar input subsets to drive predictions, with 90% of pairwise comparisons among rankings showing a degree of overlap ≥ 3 for the top-5 features. In terms of clinical validation, we demonstrated that the global reasoning of the best performing models adheres to the well-established domain knowledge and guidelines, without showing any severe AI bias. As a result, it can be considered *"human-like"* to a large extent, thus enhancing model clinical credibility.

3. Smartphone-embedded sensing for preliminary disease detection

Personal mobile devices, like smartphones and smartwatches, represent pervasive instruments for the collection of user-generated data and signals (e.g., sound, voices, images) that can be analysed to early detect specific diseases. In the last few years, during COVID-19 pandemic, new m-health solutions have been proposed to collect and analyse audio signals generated by smartphone microphones, focusing mainly on human respiratory functions like breathing, speech, and coughing [13]. Schuller et al. [14] have been the pioneers in the investigation of how the automatic analysis of speech and audio data can contribute to fight the pandemic crisis, presenting the potential of Computer Audition techniques (CA, i.e., computer-based speech and sound analysis) [15]. Subsequently, researchers investigated the effective applicability of those techniques in real scenarios even though the collection of objective data from large populations represented a big challenge. In fact, initial studies focused on small patients' cohorts trying to automatically distinguish between COVID-19 cough and cough sounds related to other pathologies [16], while others developed mobile and web apps to directly collect crowdsourced datasets from the population [17]. These projects allowed the sharing of important COVID-19 datasets, which opened the way to other researchers to validate new AI tools to improve the accuracy of the proposed systems. Specifically, we identified 3 datasets: COSWARA [18] and Virufy [19] that are publicly available, and the Cambridge dataset [20] which we can access through a data transfer agreement between CNR and Cambridge University for research purposes.

Classification methods proposed in the literature can be distinguished in three categories. Those relying on hand-crafted acoustic features, like basic frequency-based and temporal features [21], but also sets of features especially designed for voice and paralinguistic applications (e.g., COMPARE [22])This approach is generally outperformed in classification by deep learning models [23]. Therefore, researchers proposed a second approach converting the audio files into a visual representation (e.g., time-frequency spectrogram or Mel-spectrogram) that can be used as input to a Convolutional Neural Network (CNN) model for both features extraction and classification, thus relying directly on DL models [16]. Due to the scarcity of public COVID-19 respiratory sound data, the DL models' training has been performed on small-size datasets, typically composed of a few hundred samples, thus risking overfitting and providing unreliable results. Therefore, we decided to investigate a hybrid approach, based on the integration of hand-crafted features and DL models, focusing on the performance of the recently proposed *Look, Listen and Learn* (L^3-Net) [24] embedding model and YAMNET [25], comparing the obtained results with those obtained by previous works focused on VGGish [26] as deep features extraction model [27] We compared the performances of the aforementioned deep audio embedding models on the same datasets and with the same tasks, performing a series of subject-independent experiments, and we demonstrated that L^3-Net overcomes the other models in terms of standard metrics with different parameters' configurations.

All those solutions take advantage of the Transfer Learning concept to deal with the shortage of COVID-19 audio data. Therefore, to complete our analysis, we also investigated the benefits and drawbacks of feature extraction with respect to fine-tuning applied only to the final fully-connected layers of the neural networks (see Figure 1). Experimental results showed that the fine-tuned models perform considerably worse than their use as feature extractors to input a shallow classifier since their performances mainly depend on the similarity of the pretraining and target tasks [28]. Since the considered respiratory sounds considerably differ from the original model's training data, which actually include heterogeneous sounds extracted from YouTube videos, they do not allow an effective fine-tuning of the analyzed models. A possible solution for this issue could be the fine-tuning not only of the final classification layers but also part of the convolutional components dedicated to the feature extraction. However, this requires a considerable amount of data that is not currently available in public COVID-19 audio datasets [29]. Moreover, since we would like to investigate the performances of the proposed models as components of a m-health system, we provided also a preliminary evaluation of the models' complexity considering the best trade-off between the classification performances and the model's size.

Inspired by recent applications of Deep Learning in the dermatology field [30], we investigated also the use of CNNs to propose a novel m-health system aimed at detecting mpox (formerly known as Monkeypox) from skin lesion pictures captured by smartphone cameras. Adopting Transfer Learning to fine-tune different pre-trained CNNs, we identified MobileNetV3 [31] as the best performing model for our use-case scenario. We evaluated it on a combination of available skin lesions' datasets preprocessed to obtain homogeneous picture as those provided by a smartphone camera. Experiments results present an average accuracy of 93%, 87% sensitivity, and 78% specificity in the binary classification setting, and 88% accuracy, 88% sensitivity, and 96% specificity in distinguishing mpox from other diseases that produce similar skin lesions, including, for example, acne and chickenpox. Finally, we applied Grad-CAM as an XAI algorithm to validate the model's predictions and the *quantization* technique to reduce its memory footprint by 4x with negligible degradation in terms of accuracy. In this way the proposed solution results to be easily deployable on commercial mobile devices, performing the whole data processing on the user's device and thus supporting the fast detection of mpox.

4. Digital Phenotyping for Active and Healthy Ageing

Previous activities mainly focus on single health domains, but the next step towards a comprehensive analysis and classification of individual health and well-being status passes from the integration of heterogeneous data related to multiple health domains. This allows the identification of digital biomarkers that can support clinical evaluations and increase the potentialities of remote and m-health systems. In this area we focus on healthy older adults population, which can exploit the proposed solution as an early predictor of risky conditions and prevention of ageing diseases. To this aim, we proposed m-health systems aimed at collecting and analysing daily monitoring data related to physical activity, sleep patterns, nutrition, and social interactions. Most of this data can be collected by the use of commercial wearables, like smartwatches, and dedicated smart devices installed in the home environment, with a limited request of user's interaction. However, several research challenges are still open in this area, starting from the scarcity of data labeling in-the-wild up to the definition of significant features from each domain and their appropriate fusion for efficient and affordable risks prediction. In addition, personalised interventions (mainly based on behavioural suggestions) can be implemented after a risk prediction in order to observe potential behavioural changes. This activity needs a huge quantity of heterogeneous data from a consistent group of people to provide significant accuracy results. To this aim, in the framework of a recently funded PNRR project called Tuscany Health Ecosystem, we are planning to set up a large scale pilot to validate and evaluate the proposed solutions.

5. Acknowledgements

This research activity has been partially funded by the European Commission under H2020-INFRAIA-2019-1SoBigData-PlusPlus project, by the European Union - Next Generation EU, in the context of The National Recovery and Resilience Plan, Investment 1.5 Ecosystems of Innovation, Project Tuscany Health Ecosystem (THE), CUP: B83C22003920001.

References

[1] D. Cook, M. Schmitter-Edgecombe, Fusing ambient and mobile sensor features into a behaviorome for predicting clinical health scores., IEEE Access (2021).

[2] J.-P. Onnela, S. L. Rauch, Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health, Neuropsychopharmacology 41 (2016) 1691–1696.

[3] F. Di Martino, F. Delmastro, Explainable ai for clinical and remote health applications: a survey on tabular and time series data, Artificial Intelligence Review (2022) 1–55.

[4] N. Sharma, T. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: A survey, Computer methods and programs in biomedicine 108 (2012) 1287–1301.

[5] D. Mannarelli, C. Pauletti, P. Mancini, A. Fioretti, A. Greco, M. De Vincentiis, F. Fattapposta, Selective attentional impairment in chronic tinnitus: Evidence from an event-related potentials study, Clinical Neurophysiology 128 (2017) 411–417.

[6] F. Delmastro, F. D. Martino, C. Dolciotti, Cognitive training and stress detection in mci frail older people through wearable sensors and machine learning, IEEE Access 8 (2020) 65573–65590. doi:10.1109/ACCESS.2020.2985301.

[7] F. D. Martino, F. Delmastro, High-resolution physiological stress prediction models based on ensemble learning and recurrent neural networks, in: 2020 IEEE Symposium on Computers and Communications (ISCC), 2020, pp. 1–6. doi:10.1109/ISCC50000.2020.9219716.

[8] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, K. Van Laerhoven, Introducing wesad, a multimodal dataset for wearable stress and affect detection, in: Proceedings of the 20th ACM international conference on multimodal interaction, 2018, pp. 400–408.

[9] S. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, W. Kraaij, The swell knowledge work dataset for stress and user modeling research, in: Proceedings of the 16th international conference on multimodal interaction, 2014, pp. 291–298.

[10] F. Delmastro, C. Dolciotti, F. Palumbo, M. Magrini, F. Di Martino, D. La Rosa, U. Barcaro, Long-term care: how to improve the quality of life with mobile and e-health services, in: 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), IEEE, 2018, pp. 12–19.

[11] W. Willett, P. J. Skerrett, Eat, drink, and be healthy: the Harvard Medical School guide to healthy eating, Simon and Schuster, 2017.

[12] A. Valentini, M. Federici, M. A. Cianfarani, U. Tarantino, A. Bertoli, Frailty and nutritional status in older people: the mini nutritional assessment as a screening tool for the identification of frail subjects, Clinical interventions in aging 13 (2018) 1237.

[13] T. F. Quatieri, T. Talkar, J. S. Palmer, A framework for biomarkers of covid-19 based on coordination of speech-production subsystems, IEEE Open Journal of Engineering in Medicine and Biology 1 (2020) 203–206. doi:10.1109/OJEMB.2020.2998051.

[14] B. W. Schuller, D. M. Schuller, K. Qian, J. Liu, H. Zheng, X. Li, Covid-19 and computer audition: An overview on what speech &

sound analysis could contribute in the sars-cov-2 corona crisis, Frontiers in Digital Health 3 (2021). URL: https://www.frontiersin.org/article/10.3389/fdgth.2021.564906. doi:10.3389/fdgth.2021.564906.

[15] K. Qian, X. Li, H. Li, S. Li, W. Li, Z. Ning, S. Yu, L. Hou, G. Tang, J. Lu, F. Li, S. Duan, C. Du, Y. Cheng, Y. Wang, L. Gan, Y. Yamamoto, B. W. Schuller, Computer audition for healthcare: Opportunities and challenges, Frontiers in Digital Health 2 (2020). URL: https://www.frontiersin.org/articles/10.3389/fdgth.2020.00005. doi:10.3389/fdgth.2020.00005.

[16] A. Imran, I. Posokhova, H. N. Qureshi, U. Masood, M. S. Riaz, K. Ali, C. N. John, M. I. Hussain, M. Nabeel, Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app, Informatics in Medicine Unlocked 20 (2020) 100378. URL: https://www.sciencedirect.com/science/article/pii/S2352914820303026. doi:https://doi.org/10.1016/j.imu.2020.100378.

[17] J. Han, C. Brown, J. Chauhan, A. Grammenos, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, C. Mascolo, Exploring automatic covid-19 diagnosis via voice and symptoms from crowdsourced data, in: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 8328–8332. doi:10.1109/ICASSP39728.2021.9414576.

[18] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, N. R., P. K. Ghosh, S. Ganapathy, Coswara — a database of breathing, cough, and voice sounds for covid-19 diagnosis, Interspeech 2020 (2020). URL: http://dx.doi.org/10.21437/Interspeech.2020-2768. doi:10.21437/interspeech.2020-2768.

[19] G. Chaudhari, X. Jiang, A. Fakhry, A. Han, J. Xiao, S. Shen, A. Khanzada, Virufy: Global applicability of crowdsourced and clinical datasets for ai detection of covid-19 from cough, arXiv preprint arXiv:2011.13320 (2020).

[20] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, C. Mascolo, Exploring automatic diagnosis of covid-19 from crowdsourced respiratory sound data, Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2020). URL: http://dx.doi.org/10.1145/3394486.3412865. doi:10.1145/3394486.3412865.

[21] T. Drugman, J. Urbain, T. Dutoit, Assessment of audio features for automatic cough detection, in: 2011 19th European Signal Processing Conference, 2011, pp. 1289–1293.

[22] B. W. Schuller, A. Batliner, C. Bergler, F. B. Pokorny, J. Krajewski, M. Cychosz, R. Vollmann, S.-D. Roelen, S. Schnieder, E. Bergelson, A. Cristia, A. Seidl, A. S. Warlaumont, L. Yankowitz, E. Nöth, S. Amiriparian, S. Hantke, M. Schmitt, The INTERSPEECH 2019 Computational Paralinguistics Challenge: Styrian Dialects, Continuous Sleepiness, Baby Sounds & Orca Activity, in: Proc. Interspeech 2019, 2019, pp. 2378–2382. doi:10.21437/Interspeech.2019-1122.

[23] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S.-Y. Chang, T. Sainath, Deep learning for audio signal processing, IEEE Journal of Selected Topics in Signal Processing 13 (2019) 206–219. doi:10.1109/JSTSP.2019.2908700.

[24] R. Arandjelovic, A. Zisserman, Look, listen and learn, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.

[25] D. Ellis, Yamnet: A pretrained audio event classifier, https://github.com/tensorflow/models/tree/master/research/audioset/yamnet, 2019.

[26] S. Hershey, S. Chaudhuri, D. P. W. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold, M. Slaney, R. J. Weiss, K. Wilson, Cnn architectures for large-scale audio classification, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017, pp. 131–135. doi:10.1109/ICASSP.2017.7952132.

[27] M. G. Campana, A. Rovati, F. Delmastro, E. Pagani, L3-net deep audio embeddings to improve covid-19 detection from smartphone data, in: 2022 IEEE International Conference on Smart Computing (SMARTCOMP), 2022, pp. 100–107. doi:10.1109/SMARTCOMP55677.2022.00029.

[28] M. G. Campana, F. Delmastro, E. Pagani, Transfer learning for the efficient detection of covid-19 from smartphone audio data, Pervasive and Mobile Computing 89 (2023) 101754. URL: https://www.sciencedirect.com/science/article/pii/S1574119223000123. doi:https://doi.org/10.1016/j.pmcj.2023.101754.

[29] G. Vrbančič, V. Podgorelec, Transfer learning with adaptive fine-tuning, IEEE Access 8 (2020) 196197–196211. doi:10.1109/ACCESS.2020.3034343.

[30] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, K. Lakshmanna, Skin lesion classification of dermoscopic images using machine learning and convolutional neural network, Scientific Reports 12 (2022) 18134. URL: https://doi.org/10.1038/s41598-022-22644-9. doi:10.1038/s41598-022-22644-9.

[31] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, H. Adam, Searching for mobilenetv3, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019.