







Learning to Compare with Few Data for Personalised Human Activity Recognition

Nirmalie Wiratunga¹, Anjana Wijekoon¹, and Kay Cooper²

¹ School of Computing, Robert Gordon University,
Aberdeen AB10 7GJ, Scotland, UK
{n.wiratunga,a.wijekoon}@rgu.ac.uk

² School of Health Sciences, Robert Gordon University, Aberdeen AB10 7GJ,
Scotland, UK
k.cooper@rgu.ac.uk

Abstract. Recent advances in meta-learning provides interesting opportunities for CBR research, in similarity learning, case comparison and personalised recommendations. Rather than learning a single model for a specific task, meta-learners adopt a generalist view of learning-to-learn, such that models are rapidly transferable to related but different new tasks. Unlike task-specific model training; a meta-learner's training instance, referred to as a meta-instance is a composite of two sets: a support set and a query set of instances. In our work, we introduce learning-to-learn personalised models from few data. We motivate our contribution through an application where personalisation plays an important role, mainly that of human activity recognition for self-management of chronic diseases. We extend the meta-instance creation process where random sampling of support and query sets is carried out on a reduced sample conditioned by a domain-specific attribute; namely the person or user, in order to create meta-instances for personalised HAR. Our meta-learning for personalisation is compared with several state-of-the-art meta-learning strategies: 1) matching network (MN) which learns an embedding for a metric function; 2) relation network (RN) that learns to predict similarity between paired instances; and 3) MAML, a model agnostic machine learning algorithm that optimizes the model parameters for rapid adaptation. Results confirm that personalised meta-learning significantly improves performance over non personalised meta-learners.

1 Introduction

Integrated human activity recognition (HAR) and assistive technologies promise to enable people to live their life well regardless of their chronic conditions. A systematic review of interventions to promote physical activity [10] illustrated

This work was part funded by SelfBACK, a project funded by the European Union's H2020 research and innovation programme under grant agreement No. 689043. More details available at <http://www.selfback.eu>.

that interventions involving behaviour change strategies are more effective for sustaining longer-term physically active lifestyles than time-limited interventions involving structured exercises alone. Advances in telecommunications and Artificial Intelligence (AI) technologies paves the way for personalised virtual health companions to provide adherence monitoring along side behaviour change digital interventions. Innovative, person-centred strategies to monitor and predict physical activity and exercise behaviours, to scan and anticipate environmental barriers to activity, and to provide social and motivation support are required.

In this paper, we focus on one specific aspect of self-management; which is to reason from sensing data to monitor adherence to personalised self-management plans. A plan requires a user to follow physical activities such as walking and specific physiotherapy exercises. Pervasive and ubiquitous AI enabled devices are arguably best placed to continuously monitor a person’s adherence to self-management plans, make real-time predictions about the likelihood of adherence and the impact of that. What is lacking are HAR algorithms that can adapt to differences in person-specific movements (e.g. gait, disabilities, weight, height); and to do so with few data.

The idea of meta-learning is to train exactly as we would expect to deploy the system [7]. What this means is that rather than treating a specific “activity” as a class to be recognised across all persons; we instead learn to recognise the “person-activity” pair as the class; and importantly do so with a limited number of data instances per person. This can be viewed as a few-shot classification scenario [13,15] commonly used in image classification where the aim is to train with one or few data instances. Meta-learning is arguably the state-of-the-art in few-shot classification [5,11], where a wide range of tasks abstracting their learning to a meta-model, such that, it is transferable to any unseen task. Meta-learning algorithms such as MAML [5] and Relation Networks (RN) [14] are grounded in theories of metric learning and parametric optimisation, and capable of learning generalised models. The meta-learning concept of learning-to-compare aligns well with personalisation where modelling a person can be viewed as a single task; whereby a meta-model must help learn a model that is rapidly adaptable or is applicable to a new person at deployment. Here we propose *Personalised Meta-Learning* to create personalised HAR models, with a small amount of data (about one minute worth of calibration data) extracted from a person’s sensing devices. We make the following contributions:

- present personalised meta-learning in the context of matching networks (MN), relation networks (RN) and MAML;
- perform a comparative evaluation with a self-management dataset from the SELFBACK EU project to compare the utility of personalised meta-learning algorithms over conventional learning algorithms with focus on using few data; and
- provide results from an exploratory study on the transferability of meta-modals from physiotherapy experts to non-experts

2 Reasoning with Sensor Data for HAR

Previous work has demonstrated the effectiveness of applying decision support and reasoning systems to the management of a specific chronic disease. For instance Case-based reasoning (CBR), has been successfully used to incorporate evidence-base practices. For instance in managing diabetes types 1 and 2, CBR uses records that provide details about periodical visits with a physician in a case consisting of features that represent a problem (e.g. weight, blood glucose level), its solution (e.g. levels of insulin) and the outcome (e.g. hyper/hypo(glycemia)) observed after applying the solution [8,9]. More recent work [4], explored the self-management of diabetes type 1 to support monitoring of blood glucose levels before, during and after exercises. Interventions recommend carbohydrate intake based on similar cases retrieved for given HAR an exercise types.

In related work on self-management of low-back pain (LBP) [1], the SelfBACK CBR system recommends personalised care plans from similar patients. Management involves a human activity recognition (HAR) component to monitor the patient activity using sensor data that is continuously polled from a wearable device. Here a combination of patient reported monitoring, and HAR from sensor data, are used by the SelfBACK system to manage exercise adherence. Monitoring allows the system to detect periods of low activity behaviour, at which point a notification is generated to nudge the user to be more active - the intervention. An important contribution of this work is the integration of behaviour change techniques such as goal setting to focus the expected level of activity. Thereafter comparison of expected and actual behaviours to analyse goal achievement. Personalisation is important to ensure that care plans are tailored to the needs of the individual. Although there has been recent work on personalised learning using matching networks [17], more work is needed to understand them in the context of other state-of-the-art meta-learners like MAML and RN with few data.

2.1 HAR Using selfBACK Accelerometer Data

Wearables, such as smart watches or phones, are the most common form of physical activity monitoring devices and sources of delivering digital interventions. These are embedded with inertial measurement devices (e.g. accelerometers or gyroscopes) that generate time-series data which can be exploited for human activity recognition of ambulatory activities, activities of daily living, gait analysis and pose recognition [3,12,17]. In the SELFBACK project the HAR dataset has 6 ambulatory and 3 stationary activities¹. Each activity has approximately 3 min of data with a 100 Hz sampling rate recorded with 33 participants using two accelerometers on the wrist (W) and the thigh (T).

¹ <https://github.com/rgu-selfback/Datasets>.

2.2 Multi-modal Exercise Recognition with MEx Data

Exercise recognition requires more sophisticated sensors such as pressure mats and depth cameras to capture complex human movements. The MEx sensor-rich dataset² contains data from 7 exercises, selected by physiotherapists for the self-management of LBP. Data is recorded with 30 participants, performing 7 exercises, each for 60 s (maximum). Of the 30 participants, 7 were qualified in physiotherapy exercises (i.e expert users) whilst the others were general users (i.e. non-expert users). Figure 1 shows the 4 modalities: a depth camera (DC) with a frame rate 15 Hz & frame size of 240×320 ; a pressure mat (PM) using a frame rate 15 Hz & frame size: 32×16 ; and two accelerometers 100 Hz sampling rate, on the wrist (ACW) & the thigh (ACT).

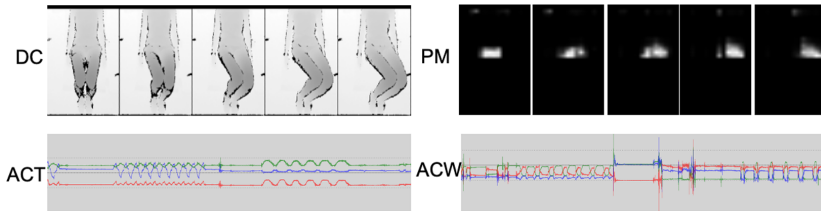


Fig. 1. Multi modal data in the MEx dataset

2.3 Personalisation with Non-iid Data

Analysis of a single person’s pressure mat data, compared to data from the general population shows that there are inherent variations between persons data. For instance in Fig. 2, we have visualised 2-dimensional compressed pressure mat data (using PCA) colour coded by exercise class. The class distribution observed using all of the 30 persons data is very different from that observed with individuals (e.g. Persons 1 and 2 in the figure). We view this as a non-identical and independent (non-iid) distributions problem where personalised meta-learning needs to be able to cope with such distributions at deployment. Accordingly we ensure that meta-modals are trained such that they are exposed to learning from such non-iid samples.

3 Learning to Personalise with Few Data

A meta-learner learns a meta-model, θ , trained over many tasks, where a task is equivalent to a “data instance” or “labelled example” in conventional machine learning. In few-shot classification, meta-learning can be seen as optimisation of a parametric model over many few-shot tasks (i.e. meta-train instances). Personalised meta-learning for HAR learns a meta-model θ from a population, \mathcal{P} ,

² <https://archive.ics.uci.edu/ml/datasets/MEx>.

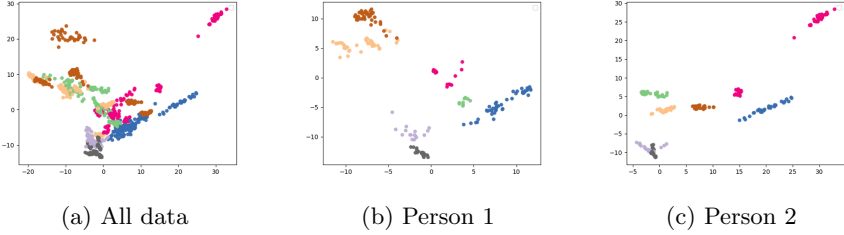


Fig. 2. MEx data distributions visualised with MEx_{PM} data

while treating activity recognition for a person as an independent task. Figure 3 illustrates the task composition for such a setting, where a dataset, \mathcal{D} , is organised over a person population, \mathcal{P} , by creating person-centric tasks, where each “person-task”, \mathcal{P}_i , contains data for a specific person. For example, in Fig. 3, \mathcal{P}_1 has a support set of distinct human activities (or activity classes) formed with data from one person. In this example we have just a single instance (i.e. $K^s = 1$) to represent each class (where $\mathcal{C} = 5$).

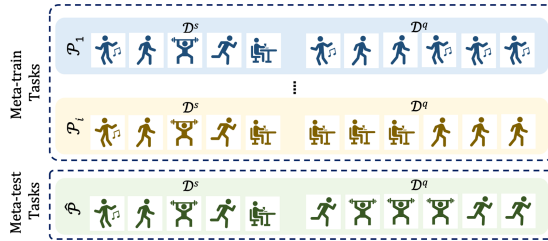


Fig. 3. Composing a meta task for training and testing a meta-learner

Meta-train and meta-test sets, are formed by randomly selection $K^s \times |\mathcal{C}|$ number of labelled data instances from a person, stratified across activity classes, \mathcal{C} , such that there are K^s amount of representatives for each class. We follow a similar approach when selecting a query set, \mathcal{D}^q , for \mathcal{P}_i . Each task contains an equal number of classes but not necessarily the same sub set of classes. Typically the query set, \mathcal{D}^q , has no overlap with the support set, \mathcal{D}^s similar to a train/test split in supervised learning; and unlike the support set, composition of the query set need not be constrained to represent all \mathcal{C} . Once the meta-model is trained using the meta-train tasks, it is tested using the meta-test tasks. An instance of a meta-test task, $\hat{\mathcal{P}}$, has a similar composition to a meta-train task instance, in that it also has a support set, $\hat{\mathcal{D}}^s$, and a query set, $\hat{\mathcal{D}}^q$. Unlike traditional classifier testing; with meta-testing, we use the support set in conjunction with the trained meta-model to classify instances in the query sets. How the meta-test support and query sets are used change depending on the aim of the meta-learning.

3.1 Learning to Match

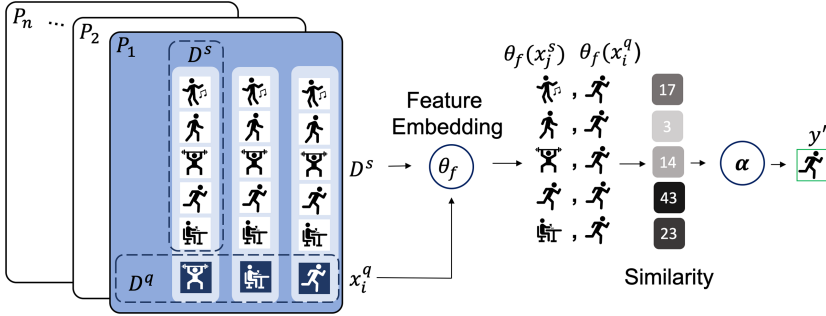


Fig. 4. Training a Matching Network

Matching Networks (MN) [15] can be viewed as an end-to-end neural implementation of the otherwise static kNN algorithm. It aims to learn a feature space by iteratively matching a query instance to a support set, which contains both positive and negative matches to the query instance. It is essentially “training to match” over representative instances from multiple classes in an iteration; which is what sets it apart from other metric learners such as Siamese [2] and triplet networks [6]. Further by training to match (instead of focusing solely on classification alone) makes it possible to add examples from new or unseen classes with no re-training of the model for transfer to related domains [17].

Figure 4 illustrates the Personalised Matching Network, MN^p , where each support set instance, x_i^s in D^s , and a query instance, x^q in D^q , are created for the person-specific task (i.e. using instances from \mathcal{P}_i). All instances in a task are transformed using the feature embedding function, θ_f (a neural network model), into feature vectors. Thereafter the process of matching is applied to every pair formed by each instance, x^q in D^q , with every instance in D^s . In the figure we can see that all pair-wise combinations are formed once D^s is duplicated thrice for a D^q with three query instances.

Similarity between a query instance and each of its support set instance pairs are calculated with an appropriate similarity metric (e.g. Cosine Similarity). Finally an attention mechanism, att , in the form of similarity weighted majority vote estimates the class, y^q (see Eqs. 1 and 2).

$$att(\theta_f(x^q), \theta_f(x_i^s)) = \frac{e^{sim(\theta_f(x^q), \theta_f(x_i^s))}}{\sum^{|S|} e^{sim(\theta_f(x^q), \theta_f(x_i^s))}} \quad (1)$$

$$y^q = \arg \max_i \sum^{|S|} att(\theta_f(x^q), \theta_f(x_i^s)) \times y_i^s \quad (2)$$

During training, the network iteratively updates weights of θ_f to maximise the similarity between the query instance and support set instance pairs from the

same activity class. This is enforced by the loss function, categorical cross-entropy, which quantifies the difference between the estimated, y^q and actual class, y , distributions. One-hot encoding is used to represent classes enabling the attention kernel multiplication with the similarity value (in the range $[0 \dots 100]$).

The concept of “learning to match” is achieved with attention where pairwise similarity computations are used to influence the network’s back propagation and consequent weight updates. This means that the embedding function that is learnt is optimised for matching which is a proxy to class prediction. At deployment, given a meta-test instance, \mathcal{P} , MN predicts the label for a query instance x^q with respect to its support set \mathcal{D}^s . In other words, the network learns to retrieve the best match from the support set elements, thereafter using them with similarity weighted majority voting to predict the class.

3.2 Learning to Relate

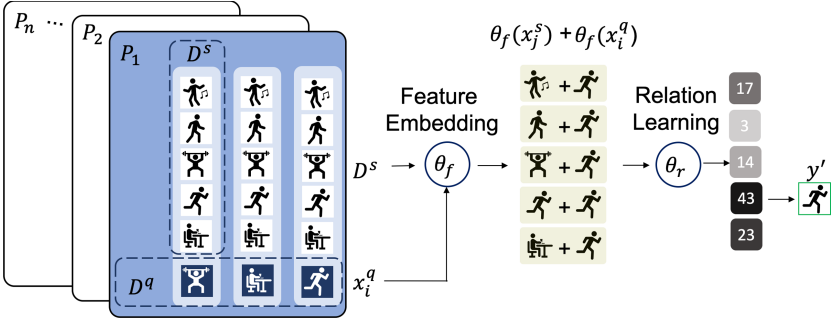


Fig. 5. Training a Relation Network

Personalised Relation Networks RN^p learns to relate by comparing query and support instance pairs using the person-tasks (\mathcal{P}_i) design discussed in Sect. 3. A RN^p has two parametric modules, one for feature representation learning, θ_f (like with MN^p) and a further one for relationship learning, θ_r (Fig. 5). Instead of capturing the relationship with a similarity metric (e.g. cosine) in the feature space, it is predicted as a score, $r_i^{q,s}$, by θ_r which is a Convolutional Neural Net (CNN) based on, $|\mathcal{C}|$ pair-wise relations.

$$r_i^{q,s} = \theta_r(\mathcal{C}(\theta_f(x^q), \theta_f(x_i^s))))), i = 1, 2, \dots, |\mathcal{C}| \quad (3)$$

$$y^q = \arg \max_i r_i^{q,s} \quad (4)$$

Unlike MN^p , the similarity-weighted attention layer is replaced with a parameterised relation learning model, θ_r , which takes as input query and support instance (concatenated) pairs to learn similarity such that matched pairs have

similarity 1 and the mismatched pair have similarity 0. Here learning similarity scores are viewed as a regression problem with mean-squared-error forming the loss function that optimises both θ_f and θ_r . At deployment, as with MN^p , given a test query instance x^q the RN^p predicts the class label, with respect to its support set \mathcal{D}^s .

3.3 Learning to Adapt

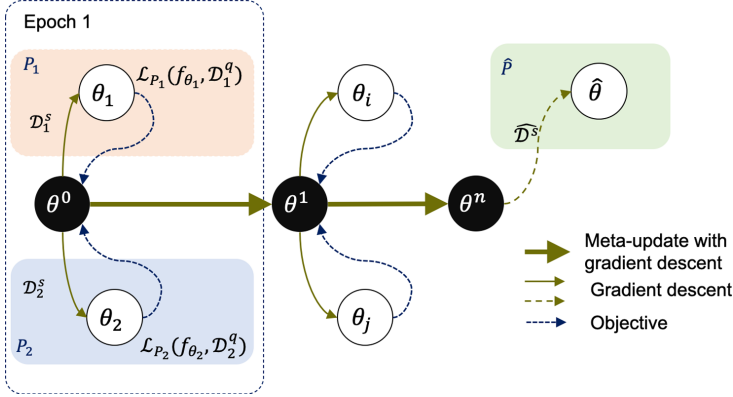


Fig. 6. Training a Personalised Model-Agnostic Meta-Learner, $MAML^p$

Unlike MN^p and RN^p , with Personalised MAML ($MAML^p$) there is less focus on similarity and instance pairing, instead the aim is to learn a generic model prototype (i.e. a meta-model), θ , such that it can be rapidly adapted to any new person encountered at test time \hat{P} . Task design for $MAML^p$ is as described in Sect. 3. Adaptation optimised learning is illustrated in Fig. 6. At the start of each iteration (epoch), a set of person tasks are sampled, P_i to optimise their person-specific model using a generic model θ^j as the model initialisation. Thereafter each person-specific model, θ_i is locally trained by optimising over the \mathcal{D}_i^s using one or few steps of gradient descents. The loss computed using \mathcal{D}^q by each person-task is passed on to the meta-learner; which in turn aggregates these losses and optimises, θ^j , using its own gradient descent step forming the meta-update for the epoch. This process is repeated n epochs, to learn a generic model prototype θ that can be rapidly adapted to a new \hat{P} .

At deployment, \hat{P} , not seen during training, uses its support set, $\hat{\mathcal{D}}^s$ for local training of the parametric model $\hat{\theta}$, initialised by the meta-model θ . Thereafter, the adapted, $\hat{\theta}$ is used to classify instances in \hat{P} 's query set, \mathcal{D}^q . Personalised MAML is model-agnostic, which is advantageous for HAR applications with heterogeneous sensor modalities or modality combinations.

4 Evaluations

The aim of the evaluations is to compare performance of the 3 personalised meta-learners discussed in Sect. 3 with several established benchmark algorithms.

DL: Best performing deep learners from benchmarks published in [16].

Meta-learners (MN, RN & MAML): The 3 meta-learners used in our comparison include; Matching Networks (MN) [15]; Relation Networks [14] (RN); and MAML (using the first-order implementation) [5].

Personalised meta-learners (MN^p , RN^p & $MAML^p$): Personalised versions of meta-learners discussed in Sect. 3.

We follow the person-aware evaluation methodology Leave-One-Person-Out (LOPO) in our experiments; where data from one person is left out to create meta-test instances and the rest used to create meta-train instances. Accuracy on meta-test is presented and any significance reported is at 95% confidence level using the Wilcoxon signed ranked test. Sensor data streams are converted into instances by applying a sliding window of size 5s, and an overlap of 3 and 2.5 for data sources MEX and SELFBACK creating 30 and 88 data instance per person-activity on average (K). We select $K^s = 5$ and $K^q = K - K^s$ to create meta-train and test instances.

MN and MN^p are trained for 20 epochs, and $MAML$, $MAML^p$, RN and RN^p for 100 epochs; all using early stopping. $MAML$ and $MAML^p$, use 5 and 10 as the number of gradients steps when training and testing respectively. MN , MN^p , $MAML$ and $MAML^p$ use a single dense layer with 1200 units as θ_f and θ . The θ_f in RN and RN^p consists of a single layer CNN (64 kernels and kernel size 3×3); θ_r is a single layer CNN (64 kernels and kernel size 3×3), followed by 2 dense layers (120 units and 1 unit); here, the last dense layer has an output of size 1 for the regression task (Sect. 3.2).

4.1 HAR Comparative Study

Results appear in Table 1 on 6 datasets derived from SELFBACK and MEX. As expected personalised meta-learning models significantly outperform conventional DL and (non-personalised) meta-learning models on all datasets. The two visual datasets; MEX_{DC} and MEX_{PM} , recorded the best performance with $MAML^p$. Both accelerometer datasets from MEX and one dataset from SELFBACK achieved best performance with RN^p . Notably, both $MAML^p$ and RN^p fail to outperform the personalised few-shot learning algorithm MN^p on the SB_W dataset which consists of sensing data obtained from the wrist having the greatest degree of freedom and therefore most prone to “noisy” movements. Interestingly, MN^p , has comparable performance against $MAML^p$ on the MEX_{ACT} dataset and outperform RN^p model on the MEX_{DC} dataset. When comparing conventional meta-learners (i.e. RN , $MAML$) and Personalised Few-Shot Learner, MN^p , we see that, MN^p models achieve comparable performances or significantly outperform at least one conventional meta-learner with all four

Table 1. Comparative Study: mean accuracy results, LOPO, 5-shot

Algorithm	ME _{ACT}	ME _{ACW}	ME _{DC}	ME _{PM}	SB _T	SB _W
Best DL	0.9015	0.6335	0.8720	0.7408	0.7880	0.6997
<i>MN</i>	0.9073	0.4620	0.5065	0.6187	0.8392	0.7669
<i>RN</i>	0.9327	0.7279	0.8189	0.8145	0.9334	0.8276
<i>MAML</i>	0.8673	0.6525	0.9629	0.9283	0.8398	0.7532
<i>MN^P</i>	0.9155	0.6663	0.9342	0.8205	0.9124	0.8653
<i>RN^P</i>	0.9436	0.7719	0.9205	0.8520	0.9487	0.8528
<i>MAML^P</i>	0.9106	0.6834	0.9795	0.9408	0.8625	0.8075

experiments; which further confirm the importance of personalisation. Overall, we find that optimisation based meta-learning algorithm (i.e. *MAML^P*) performs well on visual sensing modalities; whilst comparison based meta-learners (i.e. *MN^P* and *RN^P*) perform well on time-series data.

4.2 Discussion

In a real-world situation the data for meta-model training is likely to be provided by physiotherapy experts performing exercises. Thereafter learnt models can be applied to non-physio users. We can simulate this situation with the ME_x dataset where the 7 physio experts can be used to train a meta-modal and observe how these transfer to the rest (23 persons).

Figure 7 plots meta-test accuracy for incrementally increasing values of meta-train epochs for *RN^P* for all 23 persons (in grey) against the average results plot. We can observe the elbow point at about 40-50 epochs. Local learning with no meta-learning as expected is very low (results x-axis = 0). The general trend is that most persons show improvements in transferability with increasing epochs. Even those that struggle to improve accuracy early on seem to benefit from using the meta-modal with increasing epochs. *MAML^P* has benefited from its local training and presents a gradual increase (~5%) with increasing meta-training. For comparison results of *MAML^P* before local training (no adaptation) is included and highlights the importance of personalised model adaptation.

Figure 8 shows the impact of meta-learning on model adaptation with *MAML^P*. The rising and falling cyclic pattern can be explained by observing that local models, $\hat{\theta}_i$, are initialised at each epoch with the meta-model, and have a low meta-test accuracy starting point. As $\hat{\theta}_i$ s are refined through local training, we observe the rising trend in meta-test accuracy. We observed 3 distinct adaptation patterns among the 23 persons (bolded line graphs in Fig. 8): meta-model being easily adapted with 1 or 2 gradient steps (blue); meta-model successfully adapted over several gradient steps (orange); and failure to adapt (green). Overall it is evident that gains from personalised model transfer in most cases gradually improve with increasing meta-training.

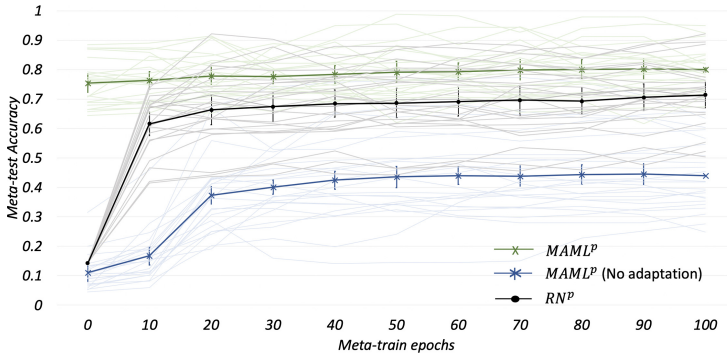


Fig. 7. Transferability of meta-modal from physiotherapy experts to non-experts

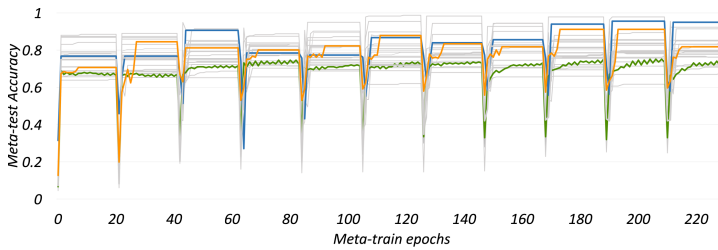


Fig. 8. Adaptation of $MAML^P$ over meta-model training (Color figure online)

5 Conclusion

Personalised meta-learning, supports model transfer to new situations in applications where there is few data. With HAR models can be transferred with a few instances of calibration data obtained from the end-user at deployment. $MAML^P$ uses calibration data to adapt through re-training; whilst MN^P and RN^P uses calibration data directly for matching (without re-training). Our results on MEX and SELFBACK datasets with personalised meta-learning show significant performance improvements over conventional and non-personalised meta-learning algorithms. Importantly we find, while RN^P outperform $MAML^P$, $MAML^P$ performs significantly faster due to the absence of paired matching. We hope that the parameterised learning-to-compare methods discussed here will help inspire new ideas relevant for CBR research.

References

1. Bach, K., Szczepanski, T., Aamodt, A., Gundersen, O.E., Mork, P.J.: Case representation and similarity assessment in the SELFBACK decision support system. In: Goel, A., Díaz-Agudo, M.B., Roth-Berghofer, T. (eds.) ICCBR 2016. LNCS (LNAI), vol. 9969, pp. 32–46. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47096-2_3

2. Bromley, J., Guyon, I., LeCun, Y.: Signature verification using a 'siamese' time delay neural network. *Int. J. Pattern Recognit. Artif. Intell.* **7**(4), 669–688 (1993)
3. Chavarrriaga, R., et al.: The opportunity challenge: a benchmark database for on-body sensor-based har. *Pattern Recogn. Lett.* **34**(15), 2033–2042 (2013)
4. Chen, Y.Y., et al.: Designing a personalised case-based recommender system for mobile self-management of diabetes during exercise. In: *Workshop Proceedings of the 2nd Knowledge Discovery from Health Workshop at the International Joint Conference on AI* (2017)
5. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, pp. 1126–1135. *JMLR. org* (2017)
6. Hoffer, E., Ailon, N.: Deep metric learning using triplet network. In: Feragen, A., Pelillo, M., Loog, M. (eds.) *Similarity-Based Pattern Recognition*, pp. 84–92. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24261-3_7
7. Hospedales, T., Antoniou, A., Micaelli, P., Storkey, A.: Meta-learning in neural networks: a survey. *arXiv preprint arXiv:2004.05439* (2020)
8. Marling, C., Wiley, M., Bunesco, R., Shubrook, J., Schwartz, F.: Emerging applications for intelligent diabetes management. *AI Mag.* **33**(2), 67–67 (2012)
9. Montani, S., Bellazzi, R., Portinale, L., d'Annunzio, G., Fiocchi, S., Stefanelli, M.: Diabetic patients management exploiting case-based reasoning techniques. *Comput. Methods Programs Biomed.* **62**(3), 205–218 (2000)
10. Morris, J.H., MacGillivray, S., Mcfarlane, S.: Interventions to promote long-term participation in physical activity after stroke: a systematic review of the literature. *Arch. Phys. Med. Rehabil.* **95**(5), 956–967 (2014)
11. Nichol, A., Achiam, J., Schulman, J.: On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999* (2018)
12. Reiss, A., Stricker, D.: Introducing a new benchmarked dataset for activity monitoring. In: *2012 16th International Symposium on Wearable Computers (ISWC)*, pp. 108–109. IEEE (2012)
13. Snell, J., Swersky, K., Zemel, R.: Prototypical networks for few-shot learning. In: *Advances in Neural Information Processing Systems*, pp. 4077–4087 (2017)
14. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M.: Learning to compare: relation network for few-shot learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1199–1208 (2018)
15. Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al.: Matching networks for one shot learning. In: *Advances in Neural Information Processing Systems*, pp. 3630–3638 (2016)
16. Wijekoon, A., Wiratunga, N., Cooper, K., Bach, K.: Learning to recognise exercises in the self-management of low back pain. In: *Florida Artificial Intelligence Research Society Conference* (2020)
17. Wijekoon, A., Wiratunga, N., Sani, S., Cooper, K.: A knowledge-light approach to personalised and open-ended human activity recognition. *Knowl.-Based Syst.* **192**, 105651 (2020)