# Deep Learning for Human Activity Recognition in Mobile Computing

**Thomas Plötz,** Georgia Tech
**Yu Guan,** Newcastle University, UK

*By leveraging advances in deep learning, challenging pattern recognition problems have been solved in computer vision, speech recognition, natural language processing, and more. Mobile computing has also adopted these powerful modeling approaches, delivering astonishing success in the field's core application domains, including the ongoing transformation of human activity recognition technology through machine learning.*

Deep learning techniques have revolutionized machine learning and its applications in ways no other technologies ever have.[1] This transformation is based on three main breakthroughs, which all co-occurred in the mid-2000s. First, there was the development of very effective modeling and especially pre-training techniques, which became very popular and enabled researchers to bootstrap substantially more complex model architectures and gave rise to *trainable* hierarchical data representations. Then there was the development of massive parallel computing architectures, such as (combinations of many) graphical processing units (GPUs); these have become mainstream and are now widely available, and thus substantially lower the barrier for training complex models by virtually eliminating computational constraints and limitations. And, finally, the increasing availability of massive databases of annotated sample data for many application domains eliminated another barrier to learning complex, deeply hierarchical models. These breakthroughs gave rise to an explosion in deep learning research, with impressive progress in application domains that are considered hard and extremely challenging including, for example, computer vision, speech recognition, and natural language processing.

 0018-9162/18/$33.00 © 2018 IEEE

Mobile computing aims to integrate computing into everyday settings, thus making it available to everyone, always. Mobile computing's key component in general, and for context awareness in particular, is the automated assessment of what a user is doing; this is referred to as human activity recognition (HAR) and is achieved through analyzing sensor data as it is routinely collected by the mobile platform. The constraints and requirements for HAR through mobile computing are significant: substantial resource limitations are imposed by the mobile platform; often near-real-time inference is needed; challenging sensor data need to be processed that include noise, ambiguity, missing data, and so on; typically, only very limited amounts of labeled data are available that can be exploited for training HAR models; and ground truth labeling is often of mixed quality or even ambiguous, which complicates training as well as validation. In the past, a wealth of methods have been developed that utilize and extend conventional machine learning methods to tackle the problem of HAR using mobile platforms under these hard constraints (for example, see the tutorial by Andreas Bulling and his colleagues[2]). Over the years, a huge variety of systems have been developed that quite impressively implement robust HAR for real-world application scenarios (such as in health, sports, or general human–computer interaction).

However, the substantial challenges of HAR in mobile computing somewhat limit the progress that could be made using those conventional machine-learning techniques. With the increasing success and popularity of deep learning methods, demands are now growing to also leverage these techniques in mobile computing scenarios. Starting with initial explorations of how deep learning techniques can help overcome the problem of finding appropriate feature representations for inertial measurement unit data,[3] to developing end-to-end HAR systems using Deep Convolutional Neural Networks,[4] or actual sequential models,[5,6] to recent explorations of transfer learning,[7] ensembles of deep long short-term memory (LSTM) networks,[8] and resource-constrained optimizations,[9] deep learning for HAR in mobile computing scenarios has come a long way. We are now starting to see the full potential of these methods in what is considered a harsh and very challenging domain.

Although impressive progress has been made and deep learning models are now competitive, if not superior to, conventional machine learning models employed for HAR in mobile computing, many problems need to be addressed to enable further progress toward truly robust and reliable systems that are able to automatically analyze human activities. In this article, we outline a roadmap to the next frontiers in deep learning for HAR. Based on a deeper exploration of the problem space and of solutions that have been proposed so far, we provide insights into the existing obstacles to further improvements and discuss possible next steps for addressing such challenges.

## HAR THROUGH MOBILE COMPUTING

HAR corresponds to the problem of automatically recognizing when a person is engaging in certain activities, so it effectively answers the questions of *what* is a person doing and *when*. Although there are extensions to this general problem domain, such as assessing *how well* a person is doing certain activities, as is common in skill assessment domains,[10] the focus of this article is on the former, more traditional interpretation of the activity recognition problem. The basis for HAR systems are observations of activities that are captured using sensors, often focusing on recording movement data, for example, through utilizing body-worn inertial measurement units (including accelerometers, gyroscopes) as they are standard in smartphones or smartwatches.

All sensing modalities record temporal data, thus sequential data streams and automated recognition methods need to solve a dual problem: localize contiguous portions within the data stream that could be relevant to the activity recognition problem at hand (*segmentation*), and *classification* of the extracted segments by automatically assigning class labels (typically from a fixed, finite lexicon). This dual problem is often a "chicken-or-egg" problem because typically information about the activity is required to determine when that activity took place, yet classification requires prior localization within the sensor data stream.

Most problematic for this dual assessment task is that the classification stage will not be able to recover any segments that are—falsely—left out of the initial segmentation step. This issue is the reason why many researchers circumvent the dual problem and instead use a heuristic sliding-window–based processing pipeline. A small analysis window is shifted along the continuous data stream extracting contiguous portions of sensor readings. The windows of extracted data are then analyzed in isolation and great results have been achieved for (quasi-)periodic/repetitive activities, such as walking or climbing stairs, if the length of the analysis window is configured appropriately (that is, by using prior domain knowledge). Conventional machine learning approaches then preprocess the sensor data covered by the analysis window,[11] extract features,[12] and employ probabilistic classification back-ends to assign activity labels to each analysis window.[2] Many HAR systems implement variants of this sliding-window–based analysis pipeline.

## THE DEEP LEARNING REVOLUTION IN HAR

Deep learning techniques come with the promise of overcoming many of the typical problems more conventional machine learning techniques have with challenging pattern recognition tasks. First and foremost, the potential to eliminate the need for manually specifying appropriate feature representations through automatically learning (hierarchical) data representations, integrated into an overarching classification model is most appealing to many. Further potential lies in the sheer modeling power of deep neural networks that allow for learning extremely complex decision functions, which is of major importance when solving challenging analytical problems.

These promises are also very appealing to HAR researchers in the mobile computing domain, and as such the community has adopted deep learning techniques—with great success, as these methods now outperform conventional machine learning techniques in many challenging tasks. This adoption and extension of deep learning methods within the HAR community did not come overnight though, due to the substantial, specific challenges outlined above. As a result, the adoption of deep learning techniques within the field followed a rather organic trajectory, as we describe below.

### Feature learning for HAR

Deep learning methods were first introduced into the field of HAR using mobile computing with the hope of finding more discriminative and especially generalizable feature representations.[3] Relatively little systematic research had addressed the problem of feature design, with almost all previous work using heuristically selected general measures. These features were either calculated in the time domain, on symbolic representations of the sensor data, or were spectra based. The main caveat of these features was that they had to be optimized manually for every single application domain, rendering system design a rather tedious and often error-prone process that could easily lead to suboptimal activity-recognition systems. The observation that led to the design of the first deep learning–based feature extraction method was that "[t]he most straightforward approach to feature design is to investigate the nature of the data to be analyzed and to develop a representation that explicitly captures its core characteristics."[3] Unlike in other domains, such as computer vision or automatic speech recognition, for mobile computing HAR problems, no all-encompassing model exists that could afford the expert-driven design of a universal feature representation. However, deep learning methods had been identified to have the potential to overcome that shortcoming by automatically discovering universal feature representations for such sensor data.

Deep learning–based feature learning utilized auto-encoder networks that aim to learn lower-dimensional representations of input data thereby minimizing the errors when used for reconstructing the original data. Feed-forward neural networks consisting of input and output layers along with an odd number of hidden layers were used, where every layer was fully connected to adjacent layers employing a nonlinear activation function. The innermost layer of the network has a lower dimensionality—an intentional bottleneck—which forces the network to learn a compact, reconstruction error–minimizing representation of the input data, which was then used as universal feature representation for the domain. The models were trained greedily in a bottom-up procedure, treating each pair of adjacent layers in the encoder as a Restricted Boltzmann Machine (RBM), which is a fully connected, bipartite, two-layer graphical model. Sets of stochastic binary hidden units were trained to effectively act as low-level feature detectors. One RBM was trained for each pair of subsequent layers by treating the activation probabilities of the feature detectors in one RBM as input for the next. Once the stack of RBMs was trained, the generative model was unrolled to obtain the final, fully initialized auto-encoder network for subsequent feature learning.

This first attempt to leverage the deep learning methods' potential for HAR scenarios in mobile computing resulted in generalizable, rich feature representations for movement sensor data. These have been used very successfully in a number of application scenarios. More importantly, this development kick-started widespread adoption and further development of sophisticated deep learning–based modeling approaches for end-to-end recognition systems.

## Convolutional neural networks for time-series analysis

Initial work on deep learning focused on extracting rich and generalizable feature representations of the sensor data. In the early 2010s, deep learning methods started to gain substantial traction in many research and application domains. Indeed, researchers in the mobile computing domain started working on end-to-end recognizers that combined the promising representation-learning aspects of deep learning methods with their superb recognition capabilities, which are facilitated by the deep, hierarchical analysis structure.

Deep convolutional neural networks (CNNs) have been studied for decades and impressive image processing results had been achieved in which single images (namely two-dimensional input data), are analyzed.[13] The key to CNNs' success lies in the employment of convolutional filter hierarchies that consecutively extract feature representations of increasing complexity from raw sensor data. Activity recognition researchers who analyze time-series, namely sequential data streams, utilized CNNs by employing the well-known sliding-window procedure to extract data blocks that are two-dimensional in principle (number of samples per window times number of sensor channels) and thus allow the use of CNNs in the same way as (static) image data as explained above. This trick was a major breakthrough that essentially paved the way for end-to-end recognizers for time-series analysis that goes beyond mere feature learning to incorporate the powerful recognition backend. Numerous examples of CNN-based activity recognition systems exist today, all of which achieve excellent recognition performance on challenging tasks. [4,5]

## Sequential modeling for HAR

CNNs are also now widely used for sequential data analysis as outlined above. However, CNNs can only be used for this kind of data by using the "trick" of employing a sliding-window preprocessing step that carves out consecutive sample data, thereby "pretending" that the time-series data is actually static data and thus analyzable using CNNs. Of course, the temporal order is preserved within the analysis window, and thus CNNs actually can implicitly deal quite effectively with sequential data. However, just as with the caveat of sliding-window procedures in conventional machine learning (as discussed above), great care has to be taken when parameterizing the approach. Inappropriate window lengths inevitably lead to suboptimal recognition results as the filters of the CNN will capture irrelevant or only partial information that shall be analyzed.

Strictly speaking, the analysis of sequential data shall employ sequential models. This is true in general, and in particular for deep learning methods due to their substantially increased complexity (compared with more conventional machine learning techniques) and potential for error. As such, HAR researchers now employ more sequential deep learning models. Prominent examples include generic recurrent neural networks, and more importantly LSTM models. LSTM models are particularly attractive, as their specialized internal structure implements a memory that includes a forget function to very effectively and selectively focus on those sensory data stream parts that are actually—according to the training process—relevant for the recognition process. As such, these models not only integrate representation learning and classification, but also effective segmentation, which is of utmost importance for HAR. The capabilities of the various modeling variants have been effectively analyzed so that specific suggestions can be given on when to use a particular modeling variant as well as what (hyper-)parameters should be subject to optimization efforts.[6] Currently, the most effective and thus successful modeling variants are combinations of CNNs and LSTM models that integrate the great feature learning capabilities of CNNs with the sequential modeling capabilities of LSTM.[5] Figure 1 illustrates the most relevant model architectures that are currently employed in the HAR community within the mobile computing domain.

## Tackling small training sets

Mobile computing's biggest challenge in using deep learning techniques for HAR scenarios is likely the absence of large scale, *labeled* training datasets. This is in stark contrast to other domains in which deep learning methods have been used very successfully, such as computer vision with its vast databases of annotated images that led to the creation of very deep models. Although it is straightforward to record virtually unlimited amounts of unlabeled data, the majority of

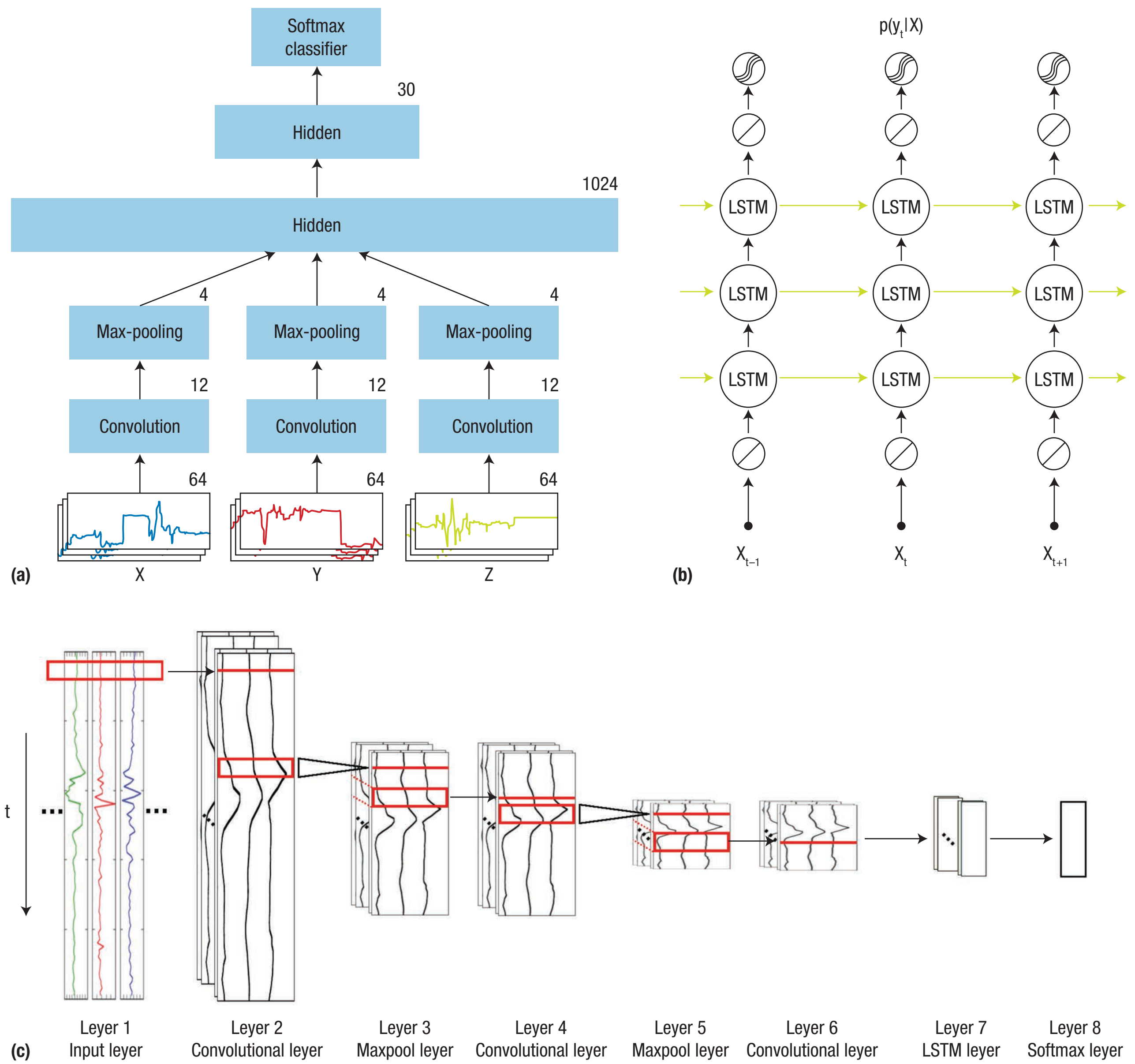


**FIGURE 1.** Deep learning model architectures for human-activity recognition (HAR) in mobile computing. (a) Convolutional neural networks (CNNs) for frame-based HAR (used with permission from M. Zeng et al., "Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors," *Proc. Int. Conf. Mobile Computing, Applications and Services* (MobiCASE 14), 2014, pp. 1–18); (b) Deep long short-term memory (LSTM) networks for HAR (used with permission from N.Y. Hammerla, S. Halloran, and T. Ploetz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables," *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI 16), 2016); (c) Combinations of CNNs and LSTM networks (used with permission from F.J.O. Morales and D. Roggen, "Deep Convolutional Feature Transfer across Mobile Activity Recognition Domains, Sensor Modalities and Locations," *Proc. Int'l Symp. Wearable Computers*, (ISWC 16), 2016).

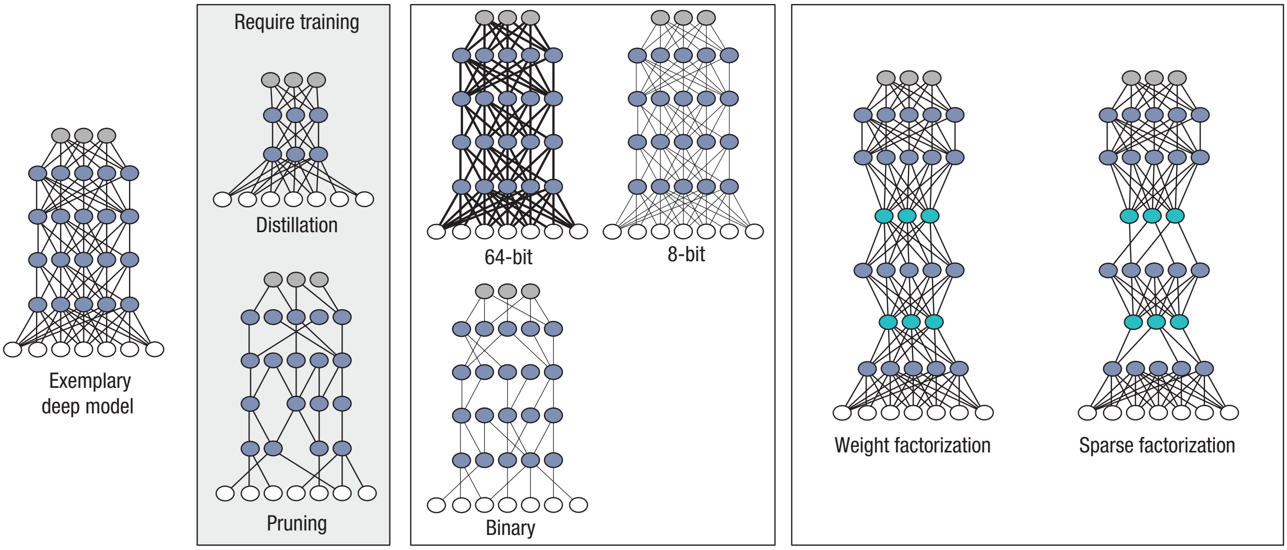


**FIGURE 2.** Dealing with resource constraints in deep learning–based HAR on mobile platforms through model surgery, compression, and reduction (used with permission from N.D. Lane et al., "Squeezing Deep Learning into Mobile and Embedded Devices," *IEEE Pervasive Computing*, vol. 16, no. 3, 2017, pp. 82–88).

contemporary deep learning techniques require annotated training data, as they follow the supervised learning paradigms. The absence of large-scale labeled datasets is a serious restriction, and if not addressed carefully can very quickly lead to overfitting, as the complex deep neural networks, with their millions of parameters, will simply memorize the small training sets.

Unlike in other domains, in HAR scenarios, it is not straightforward to simply concentrate efforts on collecting and annotating larger datasets. The reason for this is that mobile users often cannot be easily observed for ground truth annotation, and asking users to provide their own labeling can only scale to some extent.[14] Instead, researchers have focused on developing techniques that make more efficient use of existing datasets. There are two main directions for these techniques: first, the use of transfer learning techniques that lets models that were trained for a particular scenario benefit from other essentially similar domains;[7] and second, employing classifier ensembles, which are collectives of differently trained models that, when combined, lead to increased modeling capabilities due to capturing more data variability during training.[8] Transfer learning has so far only been explored at the feature extraction level, and only moderate improvements of classification accuracies could be achieved. However, the general proof has been given that feature representations can be exchanged between domains in principle and that moderate performance improvements can be achieved. Ensemble learning approaches respond to the assumption that the small labeled datasets, which are available for supervised training of deep learning models, contain redundancies and potentially erroneous data that have a negative impact on the training procedure and thus lead to suboptimal models. By randomly removing portions of the sensor data, and by combining variants of the original deep learning models trained on the same training data (for example, by combining different loss functions), it was shown that ensembles of LSTMs make more efficient use of small, potentially problematic training datasets and thus lead to less overfitting and better recognition performance overall.

## Dealing with resource constraints

Mobile computing pushes at least the inference stage of complex pattern analysis tasks onto substantially resource-constrained platforms, such as smartphones, smartwatches, or even smart textiles. These environments are characterized by limitations with regards to onboard memory, computational power, and—most crucially—dependence on battery runtime for "always on" operation. Such harsh computation environments are in stark contrast to the typically cloud-based infrastructures in which deep learning methods are trained, which have access to virtually unlimited resources both in terms of memory and (massively parallelized) computation power.

Mobile computing researchers and practitioners long ago moved on from exploring whether proof-of-concept of deep learning methods were beneficial for HAR analysis tasks. Substantial efforts have since been devoted to making deep learning models usable for (near) real-time inference on mobile platforms.[9] The majority of these efforts target automated optimization such that complex models fit onto platforms with little onboard memory and avoid unnecessary computations. Such optimizations enable near-real-time inference (model evaluation), in mobile computing scenarios. Figure 2 gives an overview of common model optimization techniques.

Reducing deep neural networks' memory footprint (with their typically millions of parameters), largely corresponds to model surgery, which eliminates parts of the model that are deemed unnecessary or at least contribute only minimally to the overall recognition process. The optimization criterion here is to reduce the number of layers, nodes, and weights without hurting the overall recognition performance too much. A number of techniques have been described in the literature that, for example, aim at "distillation" (compression) or elimination (pruning) of parts of a model.[9] Furthermore, researchers have successfully experimented with reducing the precision of the internal representation of model parameters (for example, from 16 bits to 8 bits, which dramatically reduces the memory footprint), while—interestingly—not having an overly negative effect on recognition capabilities, if done correctly. Along with smaller models go substantial reductions in runtimes when evaluating the pruned models at inference time, and it has been shown that very sophisticated deep neural networks can now be used for inference on mobile platforms, such as Qualcomm's SnapDragon, a common suite of SoC solutions that are widely used in modern smartphones and wearables such as smartwatches.

## CHALLENGES AND NEW FRONTIERS

Deep learning methods for HAR in mobile computing have come a long way, now representing the most promising category of analysis methods for a number of challenging analytical tasks. The field as a whole is very dynamic and creative in advancing the state-of-the-art so that improvements can be realized in recognition capabilities, and also so these powerful models can be used in next-generation mobile computing platforms. To do this, we need to tackle a number of substantial challenges and develop new methods. We have outlined a roadmap for addressing HAR in mobile computing's most pressing issues, and will describe them below.

### Exploiting unlabeled data

Most contemporary deep learning methods and applications are based on supervised training procedures; namely, those models that are derived from sample data and thereby associated with ground truth annotation. In domains such as computer vision or automatic speech recognition—where there are plenty of very large, annotated datasets available—the true potential of the complex deep neural networks can be witnessed through impressive recognition capabilities.

In the field of mobile computing, it is straightforward to record sample data, as these devices are "always on" and continuously recording sensor data. In contrast, it is, however, very challenging if not impossible to collect similarly large amounts of ground truth annotations. The main reason for this problem is practical in nature as it is typically infeasible to either (continuously) observe a user of mobile computing for the sake of labeling their activities, or to ask them repeatedly to provide their own ground truth annotations.

A promising way forward to overcome this sparse annotation problem is to essentially push the boundaries of unsupervised or semi-supervised learning techniques. The origins of HAR deep learning in mobile computing are in feature learning from unlabeled data where RBM-based auto-encoder networks generated generalizable data representations.[3] With the availability of very large datasets (such as the 2,000-person-year accelerometer dataset collected by the UK Biobank Consortium[15]) and novel modeling techniques (such as LSTM auto-encoders), there is huge potential to make significant progress by more effectively exploiting unlabeled data.

Semi-supervised learning combines labeled and unlabeled training data to derive more capable recognition systems. It has already been shown that transfer learning could be one feasible way to derive models that somewhat generalize across tasks or even domains.[7] With the availability of larger datasets, it should become possible to substantially push the (currently) modest improvements achieved by using transfer learning methods. Another push could come from more substantial investment in the concept of ensemble-based recognition, wherein collectives of differently trained classifiers collaboratively solve complex recognition tasks. Promising results have emerged from the combining LSTM networks, giving us evidence for the feasibility of such an approach.[8]

## Personalization

One of the most challenging—and desirable—aspects of HAR is personalization. Personalization empowers technology to implement user-specific models that are not reliant on vast amounts of user-specific, annotated sample data and do not impose unnecessary burdens on the user themselves when bootstrapping a HAR system. Key to personalization is the adaptation of generic models, thereby enabling us to focus on, for example, targeted transformations of feature representations or overall model architecture optimization to better capture the idiosyncrasies of a particular user.

Complex deep learning models primarily aim at generalization, thus they incorporate enormous numbers of parameters and sophisticated model architectures. Future personalization research should focus on analyzing unlabeled, user-specific data, for example, by analyzing what the user-independent model actually covers. Such a comparative analysis, combined with a deeper understanding of the modeling process itself (see below) will enable personalization schemes that have the potential to lead to improved recognition performance for individual users without increasing the burden for the user unnecessarily.

## Model optimization

Model optimization has been a key item on the wider research agenda for deep learning–based HAR in mobile computing. As outlined before, the rationale is to reduce the complex models' memory footprint and to accelerate processing during inference time—all of which must be done without hurting the actual recognition capabilities. Impressive results have already been achieved, but further efforts are required to make even the most complex (and most capable) deep model architectures accessible to mobile platforms.

Some researchers have already started working on dedicated compilers that translate models into highly optimized, possibly hardware-specific representations. This is in line with efforts to develop dedicated deep learning hardware, such as, Google's TensorFlow Processing Units. For mobile computing, the efforts should, however, aim at the inverse direction; that is, instead of building optimized *hardware*, researchers should focus on optimizing *the models* for existing (and future) hardware platforms. Such model optimization efforts will also support the aforementioned personalization efforts.

## Intelligibility

More and more functions in our everyday life are reliant on automated analysis, such that they are no longer visible to people who use them. This is important, particularly for mobile computing–based automation in which reliability is of mission critical importance. Even though deep learning methods lead to impressive recognition systems, there is growing animosity—especially in the practitioners' community—with regard to the fact that these complex models essentially resemble "black boxes" that do not explain how decisions are made. This aspect of non-intelligibility is a generic one that not only affects deep learning but also all machine learning–based methods. Yet, as deep learning models become increasingly complex, the issue becomes more and more pertinent as users become more and more uncomfortable with uninterpretable, mission-critical automated decision-making processes.

Substantial research efforts must be devoted to developing intelligible modeling techniques. Ideally, it should become clear what exactly a model is capturing and how it derives its decisions. In addition to responding to users' demands, intelligible deep learning for HAR also bears potential to actually improve the models' recognition capabilities as well as to further optimize the models for integration into resource-constrained mobile platforms.

## Effective segmentation of time-series data

Essentially, activity recognition corresponds to a time-series analysis problem. As such, two fundamental problems need to be tackled: segmentation and classification. So far, most analytical approaches, including deep learning–based ones, more or less circumvent the segmentation problem by employing variants of a sliding-window preprocessing step as explained above. In more recent work, researchers have moved away from fixed-length sliding-window approaches through the use of, for example, random-length frames, and even by positioning them randomly on the stream of input data for more effective mini-batch-learning procedures in ensemble approaches.[8] Alternatives of sample-by-sample processing, as suggested for example by Nils Hammerla and his colleagues,[6] might not always be an option, as such methods require more complex model architectures (such as LSTM models, which are not always feasible for some resource-constrained application domains), or are simply too challenging to train in some scenarios.

Alternatives could revert to the traditional two-stage approach comprising an explicit segmentation step followed by classification. More recent modeling techniques, including attention models, are attractive because they can automatically focus classification efforts on relevant portions of a sensor data stream, which effectively resembles segmentation.

Just as with other machine learning–based pattern-recognition domains, deep learning has revolutionized the field of HAR in mobile computing. Impressive recognition capabilities based on challenging analysis tasks are now a reality, thanks to years of research and fundamental breakthroughs that have effectively enabled us to leverage the fantastic representation learning and classification capabilities of deep neural network architectures. The state-of-the-art in the field now consists of combinations of (multiple) convolutional layers for effective feature learning and (multiple) layers containing sequential modeling nodes such as LSTM cells. Moreover, researchers are very effectively tackling the fundamental problem of severely resource-constrained operation, as this is common in the mobile computing domain. Through effective model surgery and model compression, deep learning–based recognition is now possible to run on mobile devices with both limited memory and computational power, effectively enabling near-real-time inference—required for many novel mobile computing interaction schemes.

Mobile computing's next generation of deep learning methods for HAR must tackle a number of substantial challenges that, if unaddressed, limit progress in improving recognition capabilities and hinder wider adoption of these techniques in the field. The most pressing challenges are related to dealing with noisy, often uncertain sensor data; the absence of large-scale annotated training datasets; the need for personalization of recognition systems without relying on unrealistic amounts of specialized training data or lengthy adaptation procedures; increasing model optimization to enable integration of complex models on severely resource-constrained platforms; the need for open-ended recognition schemes; and the need for intelligible and interpretable model inference.

With these exciting developments and the many important challenges related to deep learning's use in HAR for mobile computing, we look forward to further advances from the community in order to tackle the next frontiers. Deep learning is here to stay; and thanks to further advancements in modeling, data exploitation, and model optimization, an entire new generation of methods and applications are possible. We look forward to seeing where these future advances will take the field as a whole.

## ABOUT THE AUTHORS

**THOMAS PLÖTZ** is an associate professor in the School of Interactive Computing at the College of Computing at Georgia Tech. His research interests include computational behavior analysis with primary applications in the health domain, and developing machine learning techniques for sensor and general time-series data analysis. Plötz received his doctorate degree in computer science from Bielefeld University. Contact him at thomas.ploetz@gatech.edu.

**YU GUAN** is a lecturer in data science at the School of Computing Science, Newcastle University. His research interests include machine learning and its various applications such as behavior analysis, wearables, ubiquitous computing, computer vision, and biometrics. Guan received his PhD in computer science from the University of Warwick. Contact him at yu.guan@newcastle.ac.uk.

## REFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
2. A. Bulling, U. Blanke, and B. Schiele, "A Tutorial on Human Activity Recognition Using Body-worn Inertial Sensors," *ACM Computing Surveys*, vol. 46, no. 3, Jan. 2014, pp. xx-xx.
3. T. Ploetz, N. Hammerla, and P. Olivier, "Feature Learning for Activity Recognition in Ubiquitous Computing," *Proc. 22nd Int'l Joint Conf. Artificial Intelligence* (IJCAI 11), 2011, pp. 1729–1734.
4. M. Zeng et al., "Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors," *Proc. Int. Conf. Mobile Computing, Applications and Services* (MobiCASE 14), 2014, pp. 1–18.
5. [F. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors*, vol. 16, no. 1, 2016, p. 115.
6. N.Y. Hammerla, S. Halloran, and T. Ploetz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables," *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI 16), 2016, xx.
7. F.J.O. Morales and D. Roggen, "Deep Convolutional Feature Transfer across Mobile Activity Recognition Domains, Sensor Modalities and Locations," *Proc. Int'l Symp. Wearable Computers,* (ISWC 16), 2016.
8. Y. Guan and T. Ploetz, "Ensembles of Deep LSTM Learners for Activity Recognition Using Wearables," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* (IMWUT 17), 2017.
9. N.D. Lane et al., "Squeezing Deep Learning into Mobile and Embedded Devices," *IEEE Pervasive Computing*, vol. 16, no. 3, 2017, pp. 82–88.
10. A. Khan et al., "Beyond Activity Recognition," *Proc. Int'l Joint Conf. Pervasive Ubiquitous Computers* (UbiComp 15), 2015, pp. 1155–1166.
11. D. Figo et al., "Preprocessing Techniques for Context Recognition from Accelerometer Data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, 2010, pp. 645–662.
12. N. Hammerla et al., "On Preserving Statistical Characteristics of Accelerometry Data Using Their Empirical Cumulative Distribution," *Proc. Int. Symp. Wearable Computers* (ISWC 13), 2013.
13. Y. LeCun, S. Chopra, and R. Hadsell, "A Tutorial on Energy-Based Learning," *Predicting Structured Data*, 2006.
14. T. Miu, P. Missier, and T. Ploetz, "Bootstrapping Personalised Human Activity Recognition Models Using Online Active Learning," *Proc. IUCC*, 2015.
15. A. Doherty et al. , "Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study," *PLoS One*, vol. 12, no. 2, 2017, pp. e0169649–14.