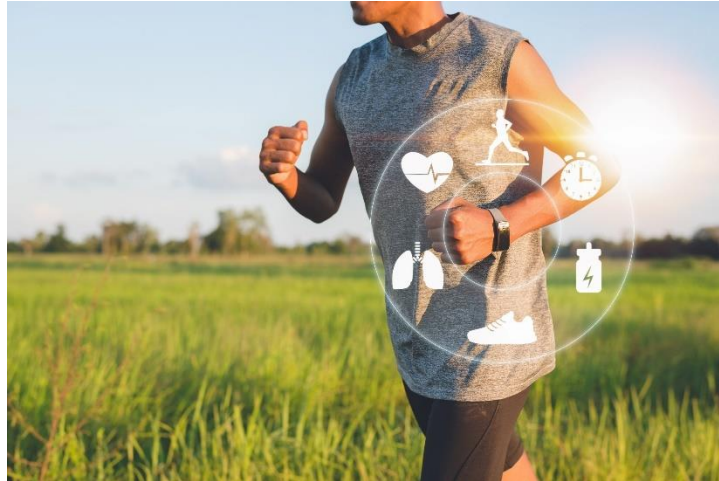


## **To Enhance Physical Activity Tracking Big Data Boosts Wearable Tech**

Researchers used the large-scale accelerometer dataset from the United Kingdom (U.K.) Biobank consisting of unlabeled data for 700,000 person days to build models to monitor [physical activity](#) levels with more accuracy and generalizability.



### **Introduction**

The field of healthcare has seen a rapid increase in the development and use of wearable devices with sensors that can be used for wellness and [fitness tracking](#), remote monitoring of patients, clinical trials that require real-time data, early detection of disease, personalized medicine, and conducting health studies at a large scale. These devices provide summary metrics on movement, sleep quality, step counts, pace, and sedentary time. However, reliable algorithms are needed to obtain information on human activity from the data collected by the sensor.

While fields such as natural language processing and computer vision have advanced significantly due to the availability of surplus data to train these learning models, the dearth of large-scale datasets that can be used to train algorithms has constrained the progress in developing models that reliably and accurately recognize [human activity](#). The lack of sufficient data to train these models has also confounded the findings on deep-learning models, suggesting that deep-learning models do not perform better than conventional methods such as simple statistics.

### **Study**

In the present study, the researchers used the accelerometer dataset from the U.K. Biobank to train the deep-learning models to recognize physical activity accurately. The U.K. Biobank conducted a [large-scale accelerometer](#) study in which close to half a million participants were recruited. Over a hundred thousand of these participants wore an accelerometer on their wrist for one week in their natural environment, as opposed to the lab-based setting. This provided approximately 700,00 person-days of free-living human motion data.

The researchers used a self-supervised learning approach, which has successfully been used for examples such as generative pre-trained transformers or GPT. Recent studies have used numerous self-supervised learning approaches such as masked reconstruction, multi-task self-supervision, bootstrapping, and contrastive learning to examine the analysis of data from [wearable sensors](#). The present study applied the multi-task self-supervision method to the large

U.K. Biobank dataset to show how a pre-trained model can be generalized to a wide range of activity-based datasets with health and clinical significance.

The multi-task self-supervision learning method was first applied to the large-scale accelerometer dataset from the U.K. Biobank to train the deep convolutional neural network. Subsequently, eight benchmark datasets were used to evaluate the performance of the pre-trained [neural network](#) and assess the quality of representation on different populations and activity types.

Labeled datasets were used to evaluate the model's success in transfer learning. Additionally, the study also used a weighted sampling approach to circumvent the problem of uninformative periods of low movement. Real-world data collected from motion sensors have periods of inactivity, and such [static signals](#) do not change during transformation, presenting problems for self-supervised learning tasks. Therefore, to improve the convergence and stability of the training process, the researchers applied a weighted sampling approach where data windows were sampled in proportion, and a standard deviation of those samples was used for the analysis.

### **Findings**

The results showed that when the models trained in this study were tested on eight benchmark datasets, they outperformed the baselines with a median relative improvement of 24.4%. Furthermore, the model could be generalized across a wide range of motion sensor devices, living environments, cohorts, and [external datasets](#).

The pre-training method of multi-task self-supervision was also found to be effective in improving downstream recognition of human activity, even in unlabeled, [small datasets](#). The self-supervised pre-training could also perform better than the supervised method.

The researchers stated that this study demonstrated that the multi-task self-supervision method of learning can be applied to datasets from wearable sensors and built accurate and generalizable activity recognition models using [deep-learning algorithms](#).

The team of researchers also released the pre-trained models to the research community working in digital [health](#) so that high-performing models could be built based on them for use in various other domains involving limited labeled data.

### **Conclusion**

To summarize, the study used a large-scale, unlabeled dataset from the U.K. Biobank consisting of accelerometer data to pre-train deep-learning models through a self-supervised approach. These pre-trained models performed beyond baseline levels in accurately analyzing motion sensor data across datasets varying in cohorts, sensor devices, and [living environments](#). The researchers believe these models can be built upon and used for various scenarios involving limited amounts of labeled data.

### **Source:**

<https://www.news-medical.net/news/20240421/Big-data-boosts-wearable-tech-New-study-enhances-physical-activity-tracking.aspx>