

Q&A: How to train AI when you don't have enough data

March 28 2024, by Sarah McQuate



Shown on the left is an image of a baby with a 2D "stick figure" created using a set of detected keypoints. On the right is the stick figure model of the baby's 3D pose. The red stick figure shows the "ground truth" and the blue stick figure is the 3D pose estimated with the researchers' algorithm. Credit: Adapted from *arXiv* (2023). DOI: 10.48550/arxiv.2311.12043

Artificial intelligence excels at sorting through information and detecting patterns or trends. But these machine learning algorithms need to be trained with large amounts of data first.

As researchers explore potential applications for AI, they have found scenarios where AI could be really useful—such as analyzing X-ray image data to look for evidence of rare conditions or [detecting a rare](#)

[fish species](#) caught on a commercial fishing boat—but there's not enough data to accurately train the algorithms.

Jenq-Neng Hwang, University of Washington professor of electrical and computer and engineering, specializes in these issues. For example, Hwang and his team developed a method that teaches AI to monitor how many distinct poses a baby can achieve throughout the day. There are limited training datasets of babies, which meant the researchers had to create a unique pipeline to make their algorithm accurate and useful.

The team recently presented this work at the [IEEE/CVF Winter Conference](#) on Applications of Computer Vision 2024. The research is [available](#) on the *arXiv* preprint server.

UW News spoke with Hwang about the project details and other similarly challenging areas the team is addressing.

Why is it important to develop an algorithm to track baby poses?

We started a collaboration with the UW School of Medicine and the Korean Electronics and Telecommunications Research Institute's AI Lab. The goal of the project was to try to help families with a history of autism know whether their babies were also likely to have autism. Babies before 9 months don't really have language skills yet, so it's difficult to see if they're autistic or not.

Researchers developed one test, called the [Alberta Infant Motor Scale](#), which categorizes various poses babies can do: If a baby can do this, they get two points; and if they can do that, they get three points; and so on. Then you add up all the points and if the baby is above some threshold, they likely don't have autism.

But to do this test, you need a doctor to observe all the different poses. It becomes a very tedious process because sometimes after three or four hours, we still haven't seen a baby do a specific pose. Maybe the baby could do it, but at that moment they didn't want to. One solution could be to use AI. Parents often have a baby monitor at home. The baby monitor could use AI to continuously and consistently track the various poses a baby does in a day.

Why is AI a good fit for this task?

My background is studying traditional image processing and computer vision. We were trying to teach computers to be able to figure out human poses from photos or videos, but the trouble is that there are so many variations. For example, even the same person wearing different outfits is a challenging task for traditional image processing to correctly identify that person's elbow on each photo.

But AI makes it so much easier. These models can learn. For example, you could train a machine learning model with a variety of motion captured sequences showing all different kinds of people. These sequences could be annotated with the corresponding 3D poses. Then this model could learn to output a 3D model of a person's pose on a sequence it has never seen before.

But in this case, there aren't a lot of motion captured sequences of babies that also have 3D pose annotations that you could use to train your machine learning model. What did you do instead?

We don't have a lot of 3D pose annotations of baby videos to train the machine learning model for privacy reasons. It's also difficult to create a dataset where a baby is performing all the possible potential poses that

we would need. Our datasets are too small, meaning that a model trained with them would not estimate reliable poses.

But we do have a lot of annotated 3D motion sequences of people in general. So, we developed this pipeline.

First we used the large amount of 3D motion sequences of regular people to train a generic 3D pose [generative AI](#) model, which is similar to the model used in ChatGPT and other GPT-4 types of large language models.

We then finetuned our generic model with our very limited dataset of annotated baby motion sequences. The generic model can then adapt to the small dataset and produce high quality results.

Are there other tasks like this: Good for AI, but there's not a lot of data to train an algorithm?

There are many types of scenarios where we don't have enough information to train the model. One example is a [rare disease](#) that is diagnosed by X-rays. The disease is so rare that we don't have enough X-ray images from patients with the disease to train a model. But we do have a lot of X-rays from healthy patients. So, we can use generative AI again to generate the corresponding synthetic X-ray image without disease, which can then be compared with the diseased image to identify disease regions for further diagnosis.

Autonomous driving is another example. There are so many real events you cannot create. For example, say you are in the middle of driving and a few leaves blow in front of the car. If you use [autonomous driving](#), the car might think something is wrong and slam on the brakes, because the car has never seen this scenario before. This could result in an accident.

We call these "long-tail" events, which means that they are unlikely to happen. But in daily life we always see random things like this. Until we figure out how to train autonomous driving systems to handle these types of events, autonomous driving cannot be useful. Our team is working on this problem by combining data from a regular camera with radar information. The camera and radar persistently check each other's decisions, which can help a machine learning algorithm make sense of what's happening.

Additional co-authors on the baby poses paper are Zhuoran Zhou, a UW research assistant in the electrical and computer engineering department; Zhongyu Jiang and Cheng-Yen Yang, UW doctoral students in the electrical and computer engineering department; Wenhao Chai, a UW master's student studying electrical and computer engineering; and Lei Li, a doctoral fellow at the University of Copenhagen.

More information: Zhuoran Zhou et al, Efficient Domain Adaptation via Generative Prior for 3D Infant Pose Estimation, *arXiv* (2023). [DOI: 10.48550/arxiv.2311.12043](https://doi.org/10.48550/arxiv.2311.12043)

Provided by University of Washington

Citation: Q&A: How to train AI when you don't have enough data (2024, March 28) retrieved 19 May 2024 from <https://techxplore.com/news/2024-03-qa-ai-dont.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.