My name is Bhat Dittakavi and I am a second-year Ph.D student from the Artificial Intelligence department at IIT-Hyderabad. I have more than two decades of technology entrepreneurship experience. Using technologies such as analytics for solving business problems has been my passion. My research areas of interest include building computer vision-based technologies for fitness and wellness.

Building AI technologies that address real-world problems in fitness training and wellness is a problem worth solving. Worldwide, more than 2 billion adults are overweight. The solution is fitness and healthcare that have global appeal. Fitness market for Yoga and Pilates alone is expected to reach $215 billion by 2025. Previous efforts in this area have been largely limited to analysing sensory data from hand-held or hand-worn devices. Works were done in exercise analysis using deep learning very rudimentary. AI works done for action recognition in sports are limited to very few such as basketball and cricket that have lots of demand and hence some funding. Currently, no AI research leverages vision and language for real-time training for personal fitness purposes. Our ultimate long-term objective is to build something like Alexa with eyes.

My thesis work is aimed at exercise pose analysis for real-time fitness training. There are two major challenges here. Somebody parts blocked by some other body parts make it hard for the AI system to recognise the type and correctness of the pose. This occlusion problem is hard as the AI system has no knowledge about the parts that it can't see. We are running various experiments using different deep learning methods whose collective wisdom might help solve this problem partially. The second challenge is about the response time of the system. As the person in front of the mobile camera performs the exercises in sequence, our system has to quickly identify the correctness of the poses being done and also localize and correct the body parts or region where the exercisers are getting it wrong. This requires our deep learning models to be very light-weight, quick in analysis and interpretation. EfficientNet and MobileNet are some of the state-of-the-art deep learning networks meant for mobile applications. We are evaluating how these methods can be adopted for exercise pose analysis.

Finally, human pose estimation, i.e. identification of x and y coordinates of 18 skeletal joints, is the natural starting point to recognize actions involved with humans. State-of-the-art human pose estimation methods are limited to action recognition datasets that are built for sports. Most of these methods don't address the occlusion well. On top of it, adopting these methods for exercises such as Yoga and Pilates is challenging as we don't have any datasets available. We are parallel working on building a yoga pose dataset for Surya Namaskar (Sun Salutations) workout that would be of immense help to us.

In summary, we are building a yoga dataset for running our own CNN based experiments and working on coming up with novel methods for occlusion handling and pose correction in exercise. Our future work will include building an AI-based visual dialogue system using CNNs and RNNs that enable it to have a dialogue with the exerciser in real-time in exerciser's native language.