A deep learning framework to estimate the pose of robotic arms and predict their movements
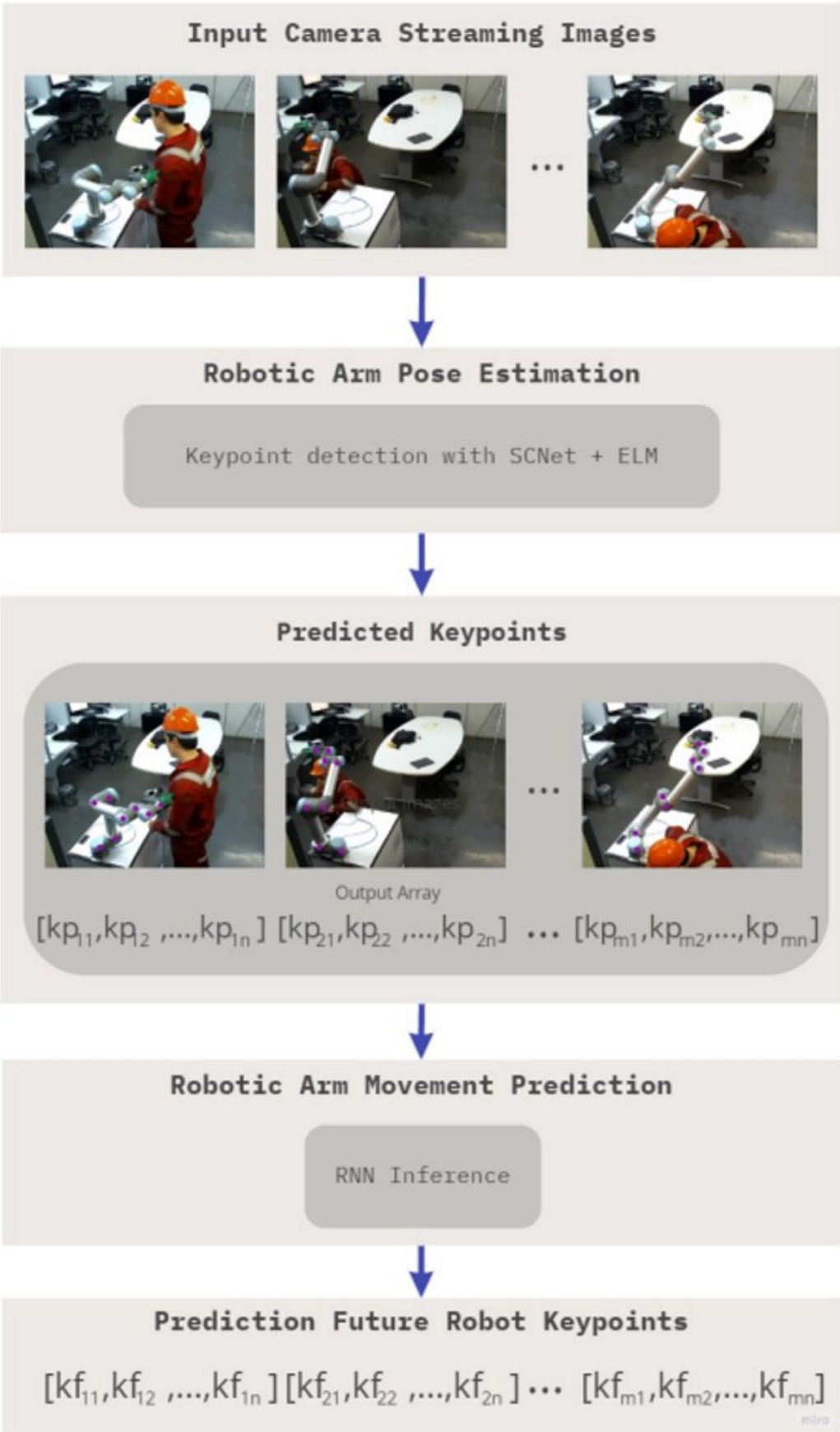
June 22 2022, by Ingrid Fadelli

Input Camera Streaming Images

Robotic Arm Pose Estimation

Keypoint detection with SCNet + ELM

Predicted Keypoints

Output Array

$[kp_{11}, kp_{12}, ..., kp_{1n}]$ $[kp_{21}, kp_{22}, ..., kp_{2n}]$ $\cdots$ $[kp_{m1}, kp_{m2}, ..., kp_{mn}]$

Robotic Arm Movement Prediction

RNN Inference

Prediction Future Robot Keypoints

$[kf_{11}, kf_{12}, ..., kf_{1n}]$ $[kf_{21}, kf_{22}, ..., kf_{2n}]$ $\cdots$ $[kf_{m1}, kf_{m2}, ..., kf_{mn}]$

Pose detection and Pose prediction fluxogram. Credit: Rodrigues et al.

As robots are gradually introduced into various real-world environments, developers and roboticists will need to ensure that they can safely operate around humans. In recent years, they have introduced various approaches for estimating the positions and predicting the movements of robots in real-time.

Researchers at the Universidade Federal de Pernambuco in Brazil have recently created a new deep learning model to estimate the pose of robotic arms and predict their movements. This model, introduced in a paper pre-published on arXiv, is specifically designed to enhance the safety of robots while they are collaborating or interacting with humans.

"Motivated by the need to anticipate accidents during [human-robot interaction](#) (HRI), we explore a framework that improves the safety of people working in close proximity to robots," Djamel H. Sadok, one of the researchers who carried out the study, told TechXplore. "Pose detection is seen as an important component of the overall solution. To this end, we propose a new architecture for Pose Detection based on Self-Calibrated Convolutions (SCConv) and Extreme Learning Machine (ELM)."

Estimating a [robot](#)'s pose is an essential step for predicting its future movements and intentions, and in turn reducing the risk of them colliding with objects in their vicinity. The approach for pose estimation and movement prediction introduced by Sadok and his colleagues has two key components, namely an SCConv and an ELM model.

The SCConvs component improves their model's overall spatial and channel dependencies. The ELM approach, on the other hand, is known to be an efficient approach to classify data.

"We observed that there were no existing studies that combine these two technologies in the context of our application," Sadok explained. "Hence, we decided to see if such combination improves our application. We also improved the framework by applying movement prediction, considering the pose detection, using recurrent neural networks (RNN)."

Scenario sample – Human and Robot. Credit: Rodrigues et al.

First, Sodok and his colleagues compiled a custom dataset containing images of scenes in which a robotic arm is interacting with a nearby human user. To create these images, they specifically used UR-5, a robotic arm created by Universal Robots.

The researchers annotated these images, particularly the robotic arm's frames. This then allowed them to use the new dataset to train SCNet, their framework's SCConv-based component.

"Our goal was to improve the observed error as compared to other known architectures, such as VGG or ResNet," Sadok said. "To extract features, we used SCNet and applied the EML at the network's end. We next used the Long Short-Term Memory (LSTM) algorithm and the Gated Recurrent Unit (GRU) to predict the movement. We consider this a new approach to solving this problem."

Sadok and his colleagues evaluated their framework's performance in a series of initial tests, where they tried to estimate the pose and predict future movements of a UR-5 arm, as it was assisting a human user with maintenance-related tasks. They found that it achieved very promising results, detecting the robotic arm's pose and predicting its future movements with good levels of accuracy.

"We believe that our main contributions are the generation of a framework capable of detecting the pose of a robotic arm and its movements, consequently improving the safety of the arm," Sadok said. "We also expanded the applicability of SSConv and EML and validated their combined capabilities."

In the future, the framework developed by this team of researchers could be used to improve the safety of both existing and newly developed robotic systems. In addition, the SCConv and ELM algorithms they used could be adapted and applied to other tasks, such as human pose estimation, object detection and object classification.

"We now plan to expand our framework to human pose detection and jointly provide a robot and pose estimation," Sadok added. "By combining both data, we can work on the joint prediction of both

movements, preventing even further risks resulting from their interaction as in a factory plant and better classify the level of risk."

More information: Iago Richard Rodrigues et al, A framework for robotic arm pose estimation and movement prediction based on deep and extreme learning models. arXiv:2205.13994v1 [cs.RO], arxiv.org/abs/2205.13994

Citation: A deep learning framework to estimate the pose of robotic arms and predict their movements (2022, June 22) retrieved 19 April 2024 from https://techxplore.com/news/2022-06-deep-framework-pose-robotic-arms.html