Contribution ID: **26**   Type: **not specified**

Long-term Recurrent Convolutional Networks for Visual Recognition and Description

Models based on deep convolutional networks have dominated recent image interpretation tasks; we investigate whether models which are also recurrent, or "temporally deep", are effective for tasks involving sequences, visual and otherwise. We develop a novel recurrent convolutional architecture suitable for large-scale visual learning which is end-to-end trainable, and demonstrate the value of these models on benchmark video recognition tasks, image description and retrieval problems, and video narration challenges. In contrast to current models which assume a fixed spatio-temporal receptive field or simple temporal averaging for sequential processing, recurrent convolutional models are "doubly deep" in that they can be compositional in spatial and temporal "layers".

Author: Ms KATE, Saenko (Texas University)

Presenter: Ms KATE, Saenko (Texas University)