

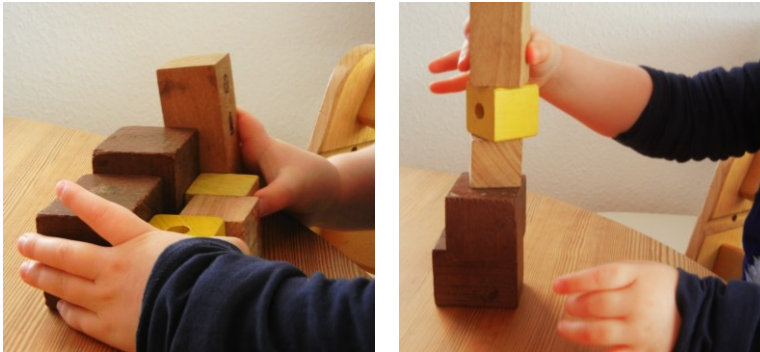
Computer Vision Assessment of Children's Fine Motor Skills in Block Stacking

M. J. Tomasik (University of Zurich) K. K. Nakka (Swiss Federal Institute of Technology) and M. Salzmann (Swiss Federal Institute of Technology)

Problem

Assessment of children's fine motor skills (FMS; Bredekamp & Copple, 1997; Lillard, 2005) using visual data from a block stacking task

- important indicator of school readiness (Johnson, Gallagher, Cook & Wong, 1995)
- strongest predictor for special education referral controlling for other skills and sociodemographic background factors (Roth, McCaul & Barnes, 1993)
- highly predictive for learning gains in primary schools (McPhillips & Jordan-Black, 1993; Murrell, 2010; Son & Meisels, 2006)



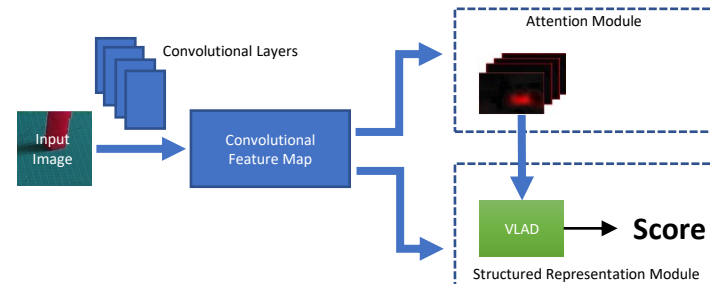
FMS typically examined with standardized procedures such as Bayley Scales of Infant and Toddler Development (Bayley, 2008)

- procedure is time-consuming
- require training for implementation and evaluation

Solution

Modification of a deep artificial neural network (Nakka & Salzmann, 2018) to regress a fine motor score given a still image of a block stack built by child as input

- usual approaches of image recognition suffer from the drawback that they aggregate information from the entire image, regardless how relevant this information is to the recognition task
- the approach by Nakka and Salzmann (2018) first produces class-specific attention maps which are then incorporated into the Vectors of Locally Aggregated Descriptors (VLAD) module
- Our network jointly identifies the informative image regions and learns a structured representation



Sample of $N = 56$ children aged 24 to 64 months conducted block stacking test from the Bayley Scales of Infant and Toddler Development (Bayley, 2008)

- block stacking task (5-10 minutes)
- cognitive testing (15 minutes, not included here)

Preliminary Results

Prediction of children's FMS from one single image with reasonable accuracy, Pearson's $r = .56$; $CI_{95} = [.31-.73]$

Base of the stack more relevant for prediction than the actual height (although height was the parameter directly representing children's FMS)

More research needed when analysing video information

