Name: Bharghav Srikhakollu

CS549 – ST/Scalable Querying & ML – Spring 2022

Topic:

Learning to Infer Graphics Programs from Hand-Drawn Images

https://arxiv.org/pdf/1707.09627.pdf

What is the problem discussed in the paper?

The paper showed how they learn to change over basic hand drawings into graphics programs composed in a subset of LATEX. The process combines procedures from deep learning and program blend. They learn a convolutional neural network that proposes conceivable drawing primitives that clarify a picture. These drawing primitives are a detail (spec) of what the illustrations program ought to draw. These programs have builds like variable ties, iterative loops, or straightforward sorts of conditionals. With an illustrations program in hand, they can correct blunders made by the profound arrange, degree likeness between drawings by use of comparative high-level geometric structures and extrapolate drawings.

Why is it important?

In this approach, they used STN – Spatial Transformer Network in place of CNN because

1. CNNs lack the ability to be spatially invariant in a computationally and parameter efficient manner
2. Max pooling layers in CNN satisfy this property where the receptive fields are fixed and local.
3. STN is a dynamic mechanism that can actively spatially transform an image or feature map.

The approach and research are useful in correcting errors made by the neural network using prior probability of programs, modelling similarity between drawings, extrapolating figures.

What are the main ideas of the proposed solution for the problem?

The approach typically mentioned in 2 stages. First, is to read perceptual input into neural network and prepare a 'trace set' which typically represents about the line, rectangles, circles etc. In Second stage, using programming synthesis this trace set is being converted into a programmable code. STN are incorporated into CNN to benefit spatial attention. The network may predict the results incorrectly, so they used SMC – Sequential Monte Carlo Sampling which uses pixel wise distance as a surrogate for likelihood function. In the results they identified better simulation with SMC + NN in comparison to LSTM, SMC, NN. In the approach, actual hand drawings are not used directly instead noise is introduced into rendering of training target images. In program synthesis stage, cost function is calculated. At the end of the experiment, they calculated the compression factor to identify the trace set and the actual program.

Reference citation: Learning to Infer Graphics Programs from Hand Drawn Images - YouTube