Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Jaroslav Macke		
Název práce	Learning to solve geometric construction problems from images		
Rok odevzdání	2021		
Studijní program	Computer Science	**Studijní obor**	Artificial Intelligence

Autor posudku	Josef Šivic	**Role**	Vedoucí práce
Pracoviště	CIIRC ČVUT		

Posudek byl připraven ve spolupráci s Jiřím Sedlářem (CIIRC ČVUT), který práci spoluškolil.

Text posudku:

Thesis overview. The objective of the thesis is to investigate whether it is possible to automatically solve geometric construction problems when the problem is specified only with image data as input, i.e. with images of the initial and the target configurations. Using image data as input is novel as current automatic solvers of geometric construction problems require specifying analytical models of the input geometric primitives. The problem is interesting as it is a step towards a long-term objective of developing automatic systems that combine machine learning with machine reasoning over unstructured inputs, for example, automatic "virtual AI assistants" that can reason about input text and associated illustrations (e.g. automatic patent lawyer assistant).

The objective of the thesis has been achieved by developing a visual recognition system based on a convolutional neural network that is trained in a supervised manner using automatically generated example constructions. At test time, given an image of the input and the target geometric configuration, the system predicts a sequence of construction steps to reach the target configuration. When trained on examples of the same problem, the resulting system solves all 68 geometric construction problems in the first six level packs of the geometric game Euclidea with a mean accuracy of 92%. The system can also solve new problems unseen at training but with lower accuracy.

Content of the thesis. The thesis is divided into seven chapters. The first chapter states the goal, motivation, challenges, and reviews the related work. The second chapter describes the Euclidea game environment for solving geometric construction problems including the different geometric tools and the procedure for generating different instances of the geometric problems. The third chapter investigates the complexity of the exhaustive search for geometric constructions when the geometric primitives are known. The fourth chapter describes the main contribution of the thesis, which is a supervised learning approach for solving geometric construction problems from the image input. The approach is building on a state-of-the-art Mask R-CNN convolutional neural network architecture for detecting objects in images and extends this architecture for solving geometric construction problems. Chapter 5 extends this approach to solving unseen geometric problems. This is achieved by exploring the space of hypotheses, suggested by the supervised learning approach, using a tree-search algorithm. Chapter 6 reports the experimental results. First, the chapter demonstrates the benefits of several algorithm design choices showing significant improvements of the developed system

over the vanilla state-of-the-art Mask R-CNN object detection baseline. Then the chapter evaluates the proposed supervised approach on the first six level-packs of the Euclidea geometric construction game. Finally, it quantifies the success rate of solving new geometric constructions unseen at training time. The final chapter 7 states the contributions of the thesis and reviews possible directions of future work.

Evaluation. Overall, the student has demonstrated independent work on a challenging new problem that has not been addressed before. He has mastered a state-of-the-art object detection approach, adapted it to solving geometric construction problems, and himself proposed a number of extensions to improve its performance managing to automatically solve more than sixty geometric construction problems of the Euclidea game. While the writing process required substantial feedback from the supervisor/consultants, particularly regarding English grammar and text structure, the final version is comprehensible and provides a solid basis for future research. With some additional effort including a comparison with analytical systems for solving geometric constructions, we would like to submit the work to a machine learning conference, such as NeurIPS or ICLR.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Datum **Podpis**
24.1.2021