



Counting Mathematical Diagrams with Machine Learning

Henrik Kragh Sørensen^(✉) and Mikkel Willum Johansen

Section for History and Philosophy of Science, Department of Science Education,
University of Copenhagen, Øster Voldgade 3, 1350 Copenhagen K, Denmark
{henrik.kragh,mwj}@ind.ku.dk

Abstract. The role and use of diagrams in mathematical research has recently attracted increasing attention within the philosophy of mathematics, leading to a number of in-depth case studies of how diagrams are used in mathematical practice. Though highly interesting, the study of diagrams still largely lack *quantitative* investigations which can provide vital background information regarding variations e.g. in the frequency or type of diagrams used in mathematics publication over time.

A first attempt at providing such quantitative background information has recently been conducted [9], making it clear that the manual labour required to identify and code diagrams constitutes a major limiting factor in large-scale investigations of diagram-use in mathematics.

In order to overcome this limiting factor, we have developed a machine learning tool that is able to identify and count mathematical diagrams in large corpora of mathematics texts. In this paper we report on our experiences with this first attempt to bring machine learning tools to the aid of philosophy of mathematics. We describe how we developed the tool, the choices we made along the way, and how reliable the tool is in identifying mathematical diagrams in corpora outside of its training set. On the basis of these experiences we discuss how machine learning tools can be used to inform philosophical discussions, and we provide some ideas to new and valuable research questions that these novel tools may help answer.

Keywords: Mathematical diagrams · Machine learning · Philosophy of mathematical practice · Digital humanities · Regional convoluted neural networks

1 Introduction

Historians and philosophers of mathematics are developing an ever-increasing interest in the role that diagrams play in mathematical reasoning and research practice [2–6]. This line of research has been highly successful in unearthing the multi-faceted and complex roles which diagrams play in mathematics; and yet the philosophical study of mathematical diagrams still largely lacks quantitative data providing vital background information for the qualitative investigation of

selected cases. Recently, a quantitative approach has led to new insights into the development in the use of diagrams over the twentieth century [9]. Among their findings is an apparent ‘valley’ in the use of diagrams, which seems to coincide with the rise of Bourbaki-style formalistic styles in mathematics during the mid-20th century [8,9].

Despite their obvious interest to the historian and the philosopher, even those quantitative studies are based on a sample from only three journals and only include volumes in five year intervals. Judging from these investigations, the major limiting factor in large-scale quantitative investigations of diagrams seems to be the huge amounts of manual labour required to identify and code diagrams by hand. Thus, to substantiate and expand the quantitative approach, an automated procedure is required to count (and subsequently classify and analyse) diagrams in mathematical texts. To this end, recent developments in machine learning may be able to lend a hand to the historian and the philosopher of mathematical practice.

In this paper, we report on our construction of a machine learning system for automated detection of mathematical diagrams. Without providing the system with any definition of a mathematical diagram, we trained an object detector by feeding it instances of diagrams from a (relatively) small set of mathematical papers. Upon iterated training, our detector was able to predict diagrams outside its training base with a (to us) surprising accuracy and precision.

We open the paper by describing how we trained the system, and we report basic measurements of its accuracy. In the final section of the paper, we discuss how an automatic diagram detector may contribute to our philosophical and historical understanding of mathematics. There, we argue that the existence of such a system opens a variety of new philosophical research questions concerning the role and diversity of mathematical diagrams which it has hitherto not been feasible to pursue.

2 Methods

Any object detector involves a number of crucial choices of which model (and implementation) to use and how to build a good training set for the task at hand. We chose to build our diagram detector on one of the well-known existing models of object detectors based on regional convoluted neural networks, known as *Fast R-CNN* [7], implemented under the *keras*-framework and publicly available [1]. And we chose to build our training set from diagrams found in the volumes of the *Journal für die reine und angewandte Mathematik*, colloquially known as *Crelle’s Journal* after its first editor. The volumes of Crelle’s Journal published from its inception in 1826 until 1998 are available at the SUB Göttinger Digitalisierungszentrum, providing us with more than 130,000 pages of mathematical text spanning the twentieth century and more. What we will refer to as *the object detector* or *the model* is thus the implementation of the framework *plus* a given very large matrix of weights (approximately 100 MB) which represent the parameters of the model.

It is a real feat of the training process that we need not give a single exhaustive definition of a mathematical diagram as such a definition is incredibly difficult to come up with. The standard definitions include aspects such as 1. being essentially two-dimensional [10], and possibly, 2. being intended to provide certain types of cognitive aid in mathematical reasoning [8]. Mathematical practice, however, does not follow such rules consistently. Matrices are, for instance, generally not considered to be diagrams although they are two-dimensional, whereas Dynkin diagrams are considered to be diagrams even in cases where they are one-dimensional. For pragmatic reasons we combined these criteria in our code-book and considered (roughly speaking) a diagram to be a two-dimensional representation generally considered to be a diagram by mathematicians.

As is always the case with supervised training in machine learning, the quality of the detector is dependent on the quality of the training set. And thus, the practice-near definition of mathematical diagrams features into our detector through the code-books which were used in tagging the training set.

During the training, the detector went through a number of iterations, refining models through exposure to both *true positives* and *false positives* (see Fig. 1). Training by true positives provides the detector with input of (ideally) varied examples of what counts as a mathematical diagram. This is provided by human tagging of diagrams in selected parts of the corpus. For the various iterations we picked out subsets of the corpus X_i , picked out all the pages on which diagrams were found, and identified the rectangles bounding the diagrams P_i . To balance the identification of diagrams by ruling out false positives (here called *background*), we implemented a bootstrapping mechanism sometimes referred to as *negative mining*: If we let the model perform predictions on all pages in X_i for which there is no true positive identified in P_i , we know that any box identified as a diagram is a *false positive*. These boxes collected as N_i can then be fed into the training of the next model as background. Thus, the training of a model builds upon the weights of the previous model and sets of boxes of true and false positives.

After we obtained Model 3, results were sufficiently good that we could apply a different method of training, which is a variant of the process known as *active learning*, where predictions made by the model are fed to an *oracle* (a human) who will classify them as true or false positives. Running model 3 on the entire corpus from Crelle’s Journal (all 130,000 pages) provided predictions of 8,700 boxes which were inspected, labeled and corrected where needed by a human agent. Together with the previously tagged true positives, these provided the training set for Model 4, which is the present culmination of our training process.

3 Results

The model was implemented under Linux Ubuntu 18.04, building on `python` and `TensorFlow` and run on a computer with an AMD Phenom(tm) II X3 720 CPU and an NVIDIA GeForce GTX 750 Ti GPU. Run-time was a real bottleneck, both in training new iterations of the model and in running predictions on large

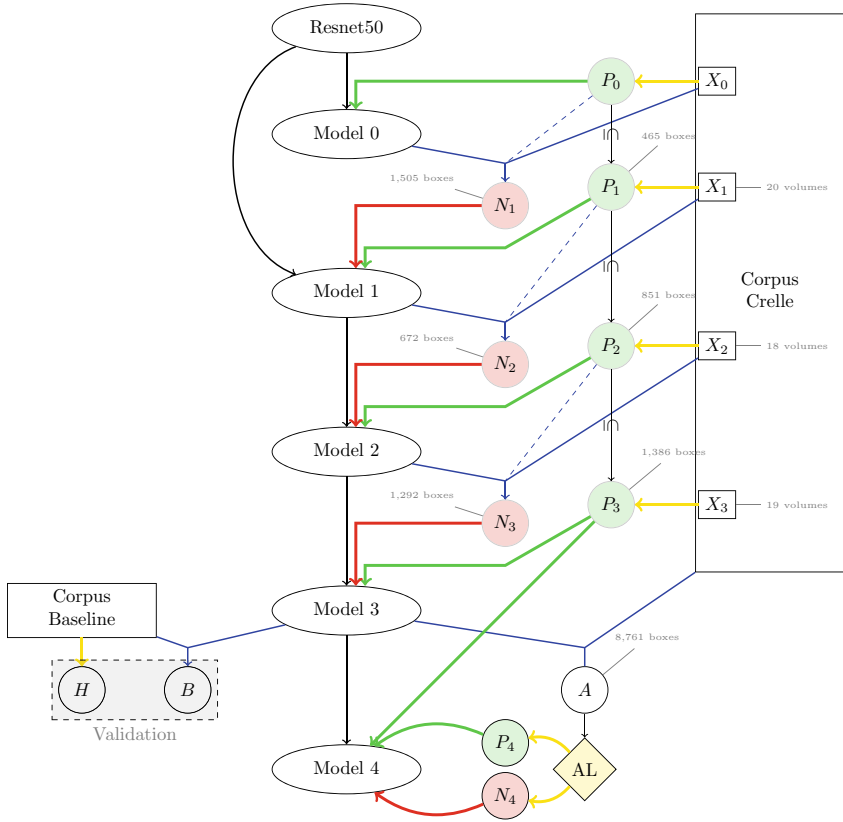


Fig. 1. Illustration of the process of training and validating the models. Boxes are corpora, ellipses are models, circles are sets of boxes (either green true positives or red false positives), diamond is active learning, yellow indicates human interaction, blue indicates prediction. H is hand-tagged, and the comparison of B and H amounts to the validation process discussed below. (Color figure online)

corpora of texts. As the system is small and somewhat dated, this could be mitigated by using more modern and larger hardware.

When we ran predictions by Model 3 on the remaining corpus from Crelle’s Journal which was not used in training Model 3, we were quite surprised at the accuracy of true positives and true negatives; in other words, the detector was surprisingly efficient in predicting diagrams precisely when they were indeed present (see Fig. 2).

However, we also encountered all the kinds of mistakes that we would expect: false positives, false negatives, and wrong partitionings. We found various types of false positives, i.e. predictions which do not correspond to diagrams. These included identifying library stamps or indented multiline formula but also some kinds of matrices and continued fractions which could rightly be considered

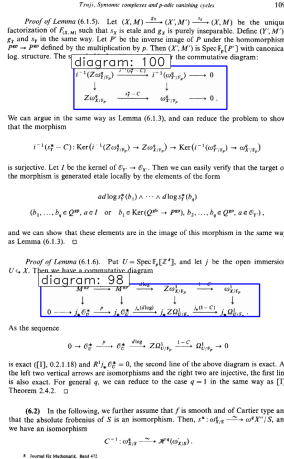


Fig. 2. One example of running the detector (model 3) on the entire corpus from Crelle’s Journal, i.e. those papers not used in training the model. For this particular page, it correctly identified two true positives.

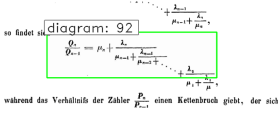
diagrams on many definitions [10]. We also found various types of false negatives, in particular some triangular commutative diagrams which were not identified as diagrams by the detector. Another special kind of false negatives came from tableaux pages with many diagrams, especially when the bounding rectangles of different diagrams would overlap; this is thought to be a side-effect of the model chosen. Furthermore, we found instances, where the detector would identify sub-rectangles of a diagram as independent diagrams.

After training our models, and to assess their quality, we ran the detector against a baseline of 677 hand-tagged articles from three journals (*Bulletin of the AMS*, *Acta Mathematica* and *Annals of Mathematics*) which are outside the training set and were tagged for another project [9]. These articles spanned 23,500 pages and contained a total of 5,271 diagrams. Different measures exist for evaluating this type of machine classification, and the best choice of measure should be based on the concerns of the application. To measure the performance of our detector on such an asymmetric set (many more negatives than positives, higher price of false negatives than of false positives), we chose to balance recall (R) and precision (P) through the F1-score:

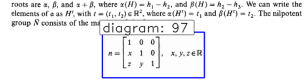
$$F1 = \frac{2 \times R \times P}{R + P}, \quad \text{where} \quad R = \frac{TP}{TP + FN} \quad \text{and} \quad P = \frac{TP}{TP + FP}. \quad (1)$$

If no true positives are found in an article, the F1-score is undefined for that article. As can be seen from the equations, R measures how many positives are picked up and classified correctly, whereas P measures the degree to which those diagrams are identified are indeed true diagrams.

The F1-score for Model 4 against the entire baseline corpus was found to be 0.90777, which is significantly improved from 0.7198 for Model 3. This is a very



On some accounts continued fractions would be considered diagrams.



On some accounts matrices would be considered diagrams.

waves and applications

By Re diagram: 90

Abstract. This paper presents a spectral analysis of large wavelength periodic trav-

Sometimes it would also mis-classify horizontal lines as diagrams.

Here $\binom{a}{b}$ denotes the quadratic Legendre symbol. The constant $c_n(f)$ is $\sqrt{3}$ or 3 depending on whether $\binom{a}{b}$ is ± 1 or ± 2 respectively. diagram: 99. Similarly, $\binom{a}{b}$ is ± 1 or ± 2 if the highest degree irreducible factor of f over \mathbb{Q}_n is linear, quadratic, or cubic, respectively. The basic conventions established in the section on notation in Part I will also be used here.

There were many referees to both authors during the creation of this work. First, we hope this work will demonstrate the arithmetic richness of the late T. Shantzer's path-breaking work on this subject. Secondly, we should like to thank Professor S. J.

And more frequently, space between lines were mis-classified as diagrams.

Fig. 3. Examples of false positives and wrong partitionings produced by running the detector (model 3) on the corpus from Crelle's Journal against which it was not trained.

Keep Z as a Sheffer map. 149
(see the discussion below Theorem 1.1.1) diagram: 100

one obtains a finite state covering $\mathcal{X}^* \rightarrow \mathcal{X}^*$. One checks easily that the induced map $\sigma^* : \mathcal{X}^* \rightarrow \mathcal{X}^*$ coincides with the spectral covering $\pi : \mathcal{X}^* \rightarrow \mathcal{X}^*$. The σ^* -factorizations of $\sigma_{\omega_1} \sigma^* : \mathcal{X}^* \rightarrow \mathcal{X}^*$ coincides with $\sigma_{\omega_1} \sigma^* : \mathcal{X}^* \rightarrow \mathcal{X}^*$ and $\sigma^* \sigma^* \sigma^*$ factors through $\sigma_{\omega_1} \sigma^*$. Consequently, to prove Theorem 1.1 we only need to consider big g .

Lemma 1.1. (1) Suppose g is of Type B and big. Then there exists a non-zero pluri-canonical form $\tau \in \Gamma(X, K_X^2)$ for some integer $\tau \geq 1$. Let ω_g be the zero class of τ . Then the pull back $\sigma^*(\omega_g)$ contains the ramification divisor R of σ .

(2) Suppose that $rk = rk_g G$. Then $\Gamma(X, K_X^2) \neq 0$.

Proof. Since $\sigma : \pi_1(X) \rightarrow G(K_X)$ is π -unbounded, the equivariant pluriharmonic maps into the building of $G(K_X)$



are not constant. Therefore, the complex differential of σ gives rise to non-zero holomorphic 1-forms $\omega_1, \dots, \omega_n$ on X^* . Since g is big, by the above factorization theorem the dimension of the span $\langle \omega_1, \dots, \omega_n \rangle$ in $H^0(X^*, \mathcal{O}_{X^*}(1))$ is maximal, and is equal to $\dim X^* = n$.

Considering the differentials of the coordinate functions $\{d\beta_1, \dots, d\beta_n\}$ on an apartment $\mathcal{A} \subset \mathcal{D}_i$, for each $d\beta_j = d\beta_j^i$, whose corresponding $\omega_j = \omega_j^i \neq 0$, we may find $n-1$ of $d\beta_1^i, \dots, d\beta_n^i$ from $\{d\beta_1^i, \dots, d\beta_n^i\}$, such that the corresponding wedge product

$$\langle \omega_1 - \omega_1 \rangle \wedge \omega_2^i \wedge \dots \wedge \omega_n^i \neq 0,$$

since the dimension of the span $\langle \omega_1, \dots, \omega_n \rangle$ in $H^0(X^*, \mathcal{O}_{X^*}(1))$ is equal to $n = \dim X^*$.

The form of the following symmetric tensor product on \mathcal{A}

$$\bigotimes_{i=1}^{n-1} \omega^*(d\beta_i - d\beta_i^i) \wedge d\beta_n^i \wedge \dots \wedge d\beta_n^i$$

is clearly \mathbb{R} -invariant, thus, we may piece those forms on all apartments together and obtain a single valued form on Δ . Pulling back this form to X^* , its (1,0)-part is a non-zero holomorphic form. Since it is $\pi_1(X)$ -invariant, it descends to a form $\tau \in \Gamma(X, K_X^2)$, where $\tau \neq 0$. Since $R = \bigcup_{i=1}^n \langle \omega_1 - \omega_1 \rangle$, there $\tau = \omega_1$ is considered as a form from $\Gamma(X^*, K_X^2)$ and $\sigma^*(\omega_g) = \bigcup_{i=1}^n \langle \omega_1 - \omega_1 \rangle$ via its construction, thus, $R \subset \sigma^*(\omega_g)$.

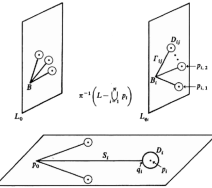
110 *Likhtcher, Homology type of the complement to algebraic curves*

Now substitution of the expression for y^2 into the third relation and elimination of y^2 (i.e. substitution of second relation in the third) gives the presentation $\mathcal{A}W$. Finally elimination of x^2 in $\mathcal{A}W$, i.e. substitution of the last relation into the first one and elimination of x^2 and x^3 (i.e. dropping second and third relations) gives Artin's presentation.

2. The presentation of the fundamental group of the complement to plane curves

Let C be an algebraic curve in \mathbb{C}^2 of degree d which we shall assume is in general position relative to the line at infinity in $\mathbb{C}P^2$.

Let \mathcal{D} denote a linear projection $\mathbb{C}^2 \rightarrow \mathbb{C}^2 = \mathcal{D}$ and p_1, \dots, p_n be the points in \mathcal{D} such that the fibres of σ over p_i are either tangent to C or contain singular points of C . These tangency points or singular points of C will be called singularities of C relative to the chosen pencil. For simplicity we shall assume that e is chosen in such a way that at each tangency point the curve C and the line $\sigma^{-1}(p_i)$ can be locally given by the equations $y = x^2$ and $y = 0$ respectively. Also we shall assume that $\sigma^{-1}(p_i)$ contains at most one singular point of C and that $\sigma^{-1}(p_i)$ does not belong to the tangent cone of C at this singular point. Let p_0 be any point in \mathcal{D} distinct from p_1, \dots, p_n and $L_0 = \sigma^{-1}(p_0)$. We shall fix a base point B in L_0 and consider a system of non-intersecting loops in $L_0 = L_0 \cap C$ representing a basis $\alpha_1, \dots, \alpha_n$ of $\pi_1(L_0 \cap C, B)$. The natural projection $\pi : \sigma^{-1}(\mathcal{D} - \bigcup_{i=1}^n p_i) \rightarrow \mathcal{D} - \bigcup_{i=1}^n p_i$ defines a locally trivial bundle. Let D_i denote a small disk in \mathbb{C} centered in p_i and let \mathcal{D}_i be a system of non-intersecting paths connecting p_i with a point q_i on the boundary of D_i (see fig. 6).

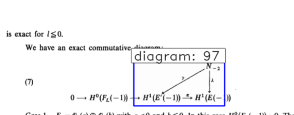


Here it picked out only one of the two commutative diagrams; triangular ones are less frequent and sometimes missed.

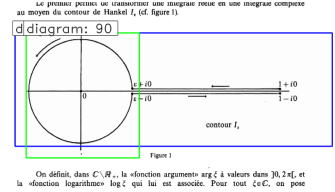
Some specific diagrams would also sometimes be missed, in particular on pages with overlapping diagram boxes.

Fig. 4. Examples of false negatives produced by the detector (model 3) on the corpus from Crelle's Journal against which it was not trained.

good score for training Model 4 on a relatively small set of tagged images and testing the model against a corpus from different mathematical and typographical traditions. It also shows that Model 4 has succeeded in eliminating many of the false predictions made by Model 3.



Not capturing the entire diagram; in particular 1-dimensional ‘appendices’ sometimes allude the detector.



Sometimes, it would pick out both the entire diagram and a part of it.

Fig. 5. Examples of wrong partitioning from running the detector (model 3) on the corpus of Crelle’s Journal on which it was not trained.

4 Discussion

Our efforts to build a mathematical diagram detector have been successful to such a degree that we now have a tool that can provide large-scale quantitative background for historical and philosophical investigations of the use of diagrams in mathematics. This background is important for several different reasons. With the detector (and its subsequent improvements) it is possible to build large corpora of diagrams spanning many journals, periods, and sub-disciplines. This will allow a more grounded approach to the investigation of the function of diagrams as large samples that better represent the diversity in the types and uses of diagrams, can easily be accessed.

Furthermore, mathematicians do not only use diagrams (and other representations) as a way to convey mathematical content. Diagrams and other representations also play a major role in the heuristic phases of the mathematical work practice and during idea and concept development. Consequently, changes in the frequency and type of the diagrams being published not only reflects aesthetic and stylistic preferences, but may also indicate underlying changes in cognitive style and epistemic values among the practitioners. The precise understanding of the changes in diagram use over time or between different sub-disciplines of mathematics is thus not only of interest in and by itself, but may also be used to identify specifically interesting periods or publications for further historical or philosophical investigation of the role of diagrams.

Finally, the fact that it is at all possible to build and train a model capable of detecting mathematical diagrams is, in itself, an interesting philosophical result. As pointed out above, it is quite easy to point to many different examples of mathematical diagrams, but difficult to give clear definitions of the concept in terms of necessary and sufficient conditions. Despite this difficulty, the detector is largely capable of mirroring human judgement concerning whether or not something is a diagram (and some of the ‘mistakes’ made by earlier iterations of the detector even reflect the inconsistencies of the concept as when it classified a continued fraction as a diagram). Although a full explanation of the concept is beyond us, it simply seems that the prototypes embedded in the examples which we provided to the detector are strong enough to allow a reasonably clear concept to form from its actions.

Acknowledgement. The images from Crelle’s Journal i Figs. 2, 3, 4 and 5 are used with permission of Walter de Gruyter and Company; permission conveyed through Copyright Clearance Center, Inc.

References

1. Bardool, K., et al.: Keras-frcnn (2019). <https://github.com/kbardool/keras-frcnn>
2. Carter, J.: Diagrams and proofs in analysis. *Int. Stud. Philos. Sci.* **24**(1), 1–14 (2010). <https://doi.org/10.1080/02698590903467085>
3. Carter, J.: Graph-algebras: faithful representations and mediating objects in mathematics. *Endeavour* **42**(2–3), 180–188 (2018). <https://doi.org/10.1016/j.endeavour.2018.07.006>
4. De Toffoli, S.: ‘Chasing’ the diagram: the use of visualizations in algebraic reasoning. *Rev. Symb. Log.* **10**(1), 158–186 (2017). <https://doi.org/10.1017/S1755020316000277>
5. De Toffoli, S., Giardino, V.: Forms and roles of diagrams in knot theory. *Erkenntnis* **79**(4), 829–842 (2013). <https://doi.org/10.1007/s10670-013-9568-7>
6. Giaquinto, M.: Crossing curves: a limit to the use of diagrams in proofs. *Philos. Math.* **19**(3), 281–307 (2011). <https://doi.org/10.1093/philmat/nkr023>
7. Girshick, R.: Fast R-CNN. In: *International Conference on Computer Vision (ICCV)*, pp. 1440–1448 (2015)
8. Johansen, M.W., Misfeldt, M., Pallavicini, J.L.: A typology of mathematical diagrams. In: Chapman, P., Stapleton, G., Moktefi, A., Perez-Kriz, S., Bellucci, F. (eds.) *Diagrams 2018*. LNCS (LNAI), vol. 10871, pp. 105–119. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91376-6_13
9. Johansen, M.W., Pallavicini, J.L.: Beyond the valley of formalism: trends and changes in mathematicians’ publication practices 1885–2015 (2020, manuscript)
10. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* **11**(1), 65–100 (1987). <https://doi.org/10.1111/j.1551-6708.1987.tb00863.x>