

# Chapter 8

## Conclusion and Challenges



### 8.1 Summary State-of-the-Art Works and Extensions

We have discussed a research topic: graphical symbol recognition, which is considered as a challenging subfield of the research domain: pattern recognition (PR). Within the PR framework, it has been taken as a key task toward document content understanding and interpretation, and mostly architectural, engineering drawings, and elecDBLP:phd/hal/Santosh11atrical circuit diagrams. In brief, starting with its definition, the book discussed basic steps that are taken from the state-of-the-art methods, a few projects, and key research standpoints. Specifically, research standpoints are relying on the state-of-the-art works that were addressing graphics recognition [1]. For a clear and concise report, readers can take a note/message reported work [2].

At the time (around 60 and 70 s) when the resource-constrained machines did not allow complex data representation and/or recognition techniques [3], it was difficult to automate a tool that has to be dealt with big data. With the increasing demand and the evolution of more powerful machines, interactions between disciplines and new projects on data mining, document taxonomy led the progress in many ways or concepts [4]. Since the 70s, graphics recognition has a rich state-of-the-art literature [5, 6]. In the literature, the state-of-art works are grouped into the three different categories/approaches: statistical, structural, and syntactic.

In all cases (approaches, mentioned earlier), the methods have been tested in accordance with the context, i.e., defined problem that may be restricted by the industrial needs, for instance, and the provided dataset. Within this framework, the recognition problem is trivial, where two (test and model) symbols are aligned/matched to check how similar they are. The similarity, more often, relies on the computed distance between the features representing the patterns. The test symbol is said to be correctly classified as the model symbol or class from which it yields the highest similarity score. As an extension, for a retrieval task, methods are able to short-list model symbols in accordance with the order of similarity. Other methods are positioned with different applications, where the recognition of graphical elements

and/or the localization of significant or known visual parts are crucial. The latter work is referred to as symbol spotting. Symbol spotting basically user-driven, where test query can be either an isolated graphical symbol or other graphical elements (meaningful parts) that signify the common characteristics of a set of train symbols (*Ref. Chap. 2*). For evaluation, we have observed that recognition rate (accuracy), precision and recall, F-measure, ROC curve, and confusion matrices are common performance measures. It is important to note that computing the aforementioned metrics is not obvious since ground truths are uncertain and missed in case of real-world data [7]. Therefore, for such a situation, as an alternative solution, retrieval efficiency can be taken as a retrieval quality measure in case of datasets, where the number of similar symbols varies from one class to another (imbalanced but labeled ground truths). Not a surprising, it often happens in real-world project [1]. Several different techniques/approaches are found in the literature. As stated earlier, two major points: datasets and evaluation metrics, are important to make a fair comparison. This means that, in order to see, how far we have been advanced, one needs to follow the exact similar evaluation protocol. More often, the characteristics of the datasets, their availability for further researches, and the applications (or intentions) may change one's evaluation metric. Besides, one may be biased in re-implementing previously reported algorithms/techniques. As a consequence, we are unable to track researches done over several years, since results cannot be consistent as algorithms may not be tuned (i.e., parameters) as in the original references [8]. As reported in [9], document analysis and exploitation (DAE) was conceived and built around a core data model that establishes an exhaustive range of relations between document images, annotation areas, interpretations, or ground truth. It also connects the data to user interactions, experimental protocols, or program executions. In Chap. 3, more detailed discussion has been made on several different services, such as querying, up- and download, and remote execution.

Based on our review, statistical approaches are appropriate to recognize isolated symbols as they are robust to noise (of almost all types), degradations, deformations, and occlusions. Statistical signatures (shape-based signatures) are basically simple (1D feature vector) to compute with low computational cost. Several different signal-based features can be combined. Discrimination power and robustness, however, strongly depend on the selection of an optimal set of features. Integrating features are not straightforward and trivial, since appropriate fusion of classifiers is also crucial. A more detailed information can be taken from Chap. 4 and [?] for extended results.

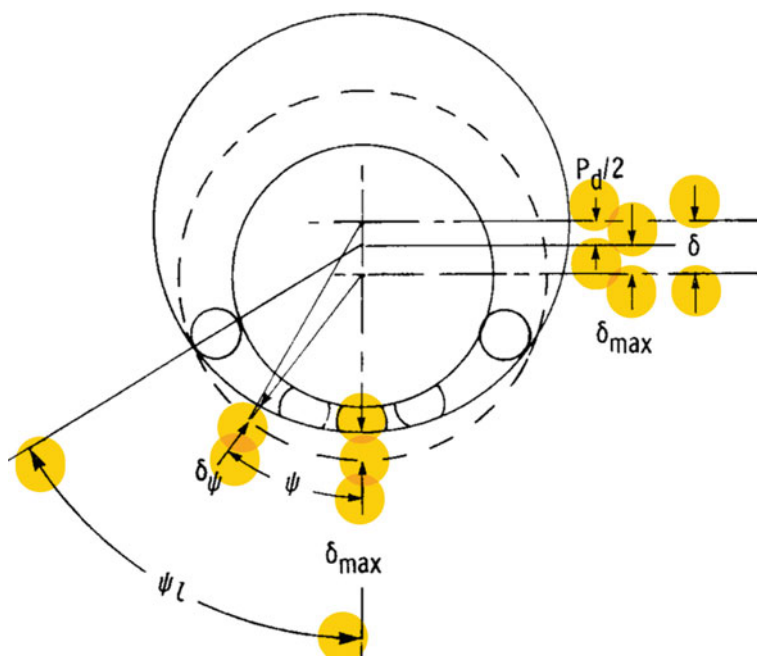
On contrary, structural approaches are particularly well suited for recognizing complex and composite graphical symbols (*Ref. Chap. 5* and previous works [10, 11]). Under this framework, graphical elements/symbols can be used for spotting/localization. For example, these techniques/algorithms are designed to recognize meaningful region-of-interest that can be a complete graphical symbol or any basic shapes representing the characteristics of any particular graphical symbol in technical documents. In structural approaches, methods are relying on symbolic data structures, such as graphs, strings, and trees. In the state-of-the-art literature, graph-based pattern representation (including matching) has been considered as a prominent technique

even if it suffers from high computational cost. Graph matching cost, i.e., computational complexity often increases when complex and composite symbols are taken for study due to the well-known problem: subgraph isomorphism. Further, due to the presence of noise and possible distortions in the studied patterns, graph sizes vary a lot. This variation is taken as one of the reasons that helps increase graph matching computational cost. In contrast to statistical approaches, structural approaches provide a powerful representation since they convey how parts are connected to each other. Such a representation preserves the technique's generality and extensibility. The term "extensibility" allows us to combine/integrate to other methods that come from different approaches.

Since not a single method (either from statistical or structural) provides a satisfactory performance, hybrid approaches (*Ref. Chap. 6*) are designed to check whether they can compliment each other. In other words, hybrid approaches try to integrate best of the two worlds: statistical and structural, for instance. In the previously reported work [?], results have been extended/advanced. Such approaches are often dedicated to the graphical symbol localization in accordance with the specific rules and are based on a set of arbitrary graphical symbols. Not to be confused, the concept of integrating descriptors and classifiers can be different than hybrid approaches. Within the framework, in visual cues/primitive selection, error-prone raster-to-vector conversion can limit the number of applications. As we are aware that primitive extraction is not generic, one can focus on those primitives that are important in that particular application. Therefore, depending on the studied samples, graphs vary. For example, graph can be either proximity graph or line graph. We observed that, often, proximity graph uses local interest points (via computer vision local descriptors) and line graph uses lines (high-level information). Researchers have shown that the line graph is appropriate for technical line drawings.

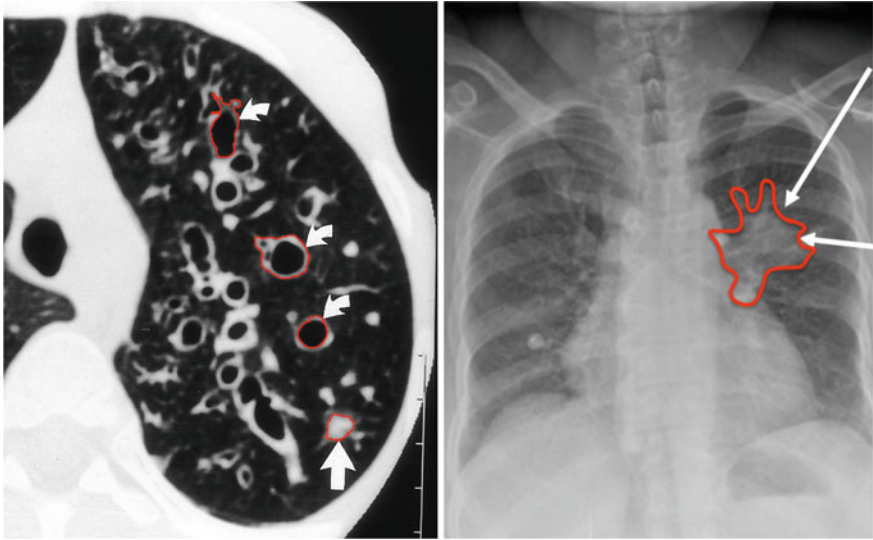
Syntactic approaches (*Ref. Chap. 7*) describe graphical symbols (or technical documents) using well-mastered grammars (rule-set, for instance). For syntactic approaches, one can use similar primitives as in structural approaches. An idea to use syntactic approaches is to make image description close to the language (first-order logic description). As reported in [12], statistical signatures to spatial predicates conversion may not carry precise information. This means that no metrical details can be found. This results syntactic approaches do not possess detailed information and the approaches may not handle complex and composite documents.

Even though we have not observed that state-of-the-art methods are generic, applications in graphical symbol recognition are not limited. Other than conventional graphics recognition tasks, arrow detection can be considered as one of the graphical symbol/elements and has several different applications. Arrow detection was initially designed for a technical document understanding, where detecting arrows (pointers, in general) can help locate quotation, measurements, and of course, meaningful regions/parts [13–15]. Figure 8.1 shows an example of it. Not a surprising, use of arrow detection can be extended to other domains as well. Arrow detection has recently been considered as an important step in biomedical images to advance the CBIR problem [16–19]. Regardless of applications, often, they aim to address regions-of-interest. Like in technical drawing, detecting overlaid arrows in medical



**Fig. 8.1** Arrow detection: another important task in graphics recognition. Arrow detection helps locate important quotations and meaningful parts. Highlighted regions (in yellow) are the detected arrows

images can help speed up in region labeling since biomedical images, by nature, tend to be very complex. Few examples are shown in Fig. 8.2. For better understanding, a complete project is demonstrated in Fig. 8.3, where a project from the US National Library of Medicine's (NLM's) entitled Open- $i^{SM}$  image retrieval search engine is provided. In brief, pointer detection can minimize the distractions from other image regions, and more importantly, meaningful regions (regions-of-interest) are often referred to article text and figure captions. It can thus help better analyze the content using other text semantics through the use of natural language processing. Further, can we use pointer location to learn regions-of-interest so that one does not require to learn all pixels (end-to-end) from the image (see Fig. 8.2)? In Fig. 8.2 (right), pointers help learn "infiltrate" without considering all pixels into account. From the machine intelligence (machine learning) viewpoint, one should not stop learning, since learning helps machine robust. This may sometime confuse decisions. Can we just avoid redundancies (via the use of pointer location) from which machines are confused? Of course, let us examine more and extend graphics recognition techniques to another level. In a similar fashion, robust circle-like element detection can help advance abnormality chest X-ray screening [20–22]. These examples can prove that graphics recognition is not just limited to technical drawings, architectural drawings, electrical



**Fig. 8.2** Illustrating the importance of arrow/pointer detection that helps locate meaningful regions. Regions (in red) are labeled as soon as we detect arrows. These regions-of-interest (in red) are automatically generated regions based on the changes in gradients (not annotated by experts)

**Fig. 8.3** Addressing the usefulness of the annotated arrow in biomedical images. Its location pointing region-of-interest (ROI) and relationship between the texts and ROI (source: US National Library of Medicine’s (NLM’s) Open-*i*<sup>SM</sup> can help advance image retrieval search engine (url: <https://openi.nlm.nih.gov>))

circuit diagrams, and other business document imaging; it can attract a large audience (up to the level of medical imaging [23]).

## References

1. K.C. Santosh, Reconnaissance graphique en utilisant les relations spatiales et analyse de la forme. (Graphics Recognition using Spatial Relations and Shape Analysis). Ph.D. thesis, University of Lorraine, France, 2011
2. K.C. Santosh, L. Wendling, *Graphical Symbol Recognition* (Wiley, New York, 2015), pp. 1–22
3. G. Nagy, State of the art in pattern recognition. Proc. IEEE **56**(5), 836–862 (1968)
4. A.K. Jain, R.P.W. Duin, J. Mao, Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 4–37 (2000)
5. H. Bunke, P.S.P. Wang (eds.), *Handbook of Character Recognition and Document Image Analysis* (World Scientific, Singapore, 1997)
6. D. Doermann, K. Tombre, *Handbook of Document Image Processing and Recognition* (Springer, New York Incorporated, 2014)
7. B. Lamiroy, T. Sun, Computing precision and recall with missing or uncertain ground truth, in *Graphics Recognition. New Trends and Challenge*, ed. by Y.-B. Kwon, J.-M. Ogier, Lecture Notes, in Computer Science, vol. 7423, (Springer, Berlin, 2013), pp. 149–162
8. B. Lamiroy, D.P. Lopresti, H.F. Korth, J. Heflin, How carefully designed open resource sharing can help and expand document analysis research, in *Document Recognition and Retrieval XVIII, Part of the IS&T-SPIE Electronic Imaging Symposium* (2011), p. 787400
9. B. Lamiroy, DAE-NG: A shareable and open document image annotation data framework, in *1st International Workshop on Open Services and Tools for Document Analysis, 14th IAPR International Conference on Document Analysis and Recognition* (2017), pp. 31–34
10. K.C. Santosh, B. Lamiroy, L. Wendling, Symbol recognition using spatial relations. Pattern Recognit. Lett. **33**(3), 331–341 (2012)
11. K.C. Santosh, L. Wendling, B. Lamiroy, Bor: Bag-of-relations for symbol retrieval. Int. J. Pattern Recognit. Artif. Intell. **28**(06), 1450017 (2014)
12. K. C. Santosh, B. Lamiroy, J.-P. Ropers, Inductive logic programming for symbol recognition, in *Proceedings of International Conference on Document Analysis and Recognition* (IEEE Computer Society, Washington, 2009), pp. 1330–1334
13. W. Min, Z. Tang, L. Tang, Recognition of dimensions in engineering drawings based on arrowhead-match, in *Proceedings of 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japan)* (1993), pp. 373–376
14. G. Priestnall, R.E. Marston, D.G. Elliman, Arrowhead recognition during automated data capture. Pattern Recognit. Lett. **17**(3), 277–286 (1996)
15. L. Wendling, S. Tabbone, A new way to detect arrows in line drawings. IEEE Trans. Pattern Anal. Mach. Intell. **26**(7), 935–941 (2004)
16. K.C. Santosh, L. Wendling, S. Antani, G. Thoma, Scalable arrow detection in biomedical images, in *Proceedings of the IAPR International Conference on Pattern Recognition* (IEEE Computer Society, Washington, 2014), pp. 1051–4651
17. K.C. Santosh, N. Alam, P.P. Roy, L. Wendling, S.K. Antani, G.R. Thoma, A simple and efficient arrowhead detection technique in biomedical images. IJPRAI **30**(5), 1–16 (2016)
18. K.C. Santosh, P.P. Roy, *Arrow detection in biomedical images using sequential classifier* (Int. J. Mach. Learn. Cybern, 2017)
19. K.C. Santosh, L. Wendling, S. Antani, G.R. Thoma, Overlaid arrow detection for labeling regions of interest in biomedical images. IEEE Intell. Syst. **31**(3), 66–75 (2016)
20. F.T. Zohora, K.C. Santosh, Circular foreign object detection in chest x-ray images, in *Recent Trends in Image Processing and Pattern Recognition, Revised Selected Papers*, ed. by K.C. Santosh, M. Hangarge, V. Bevilacqua, A. Negi. Communications in Computer and Information. Science **709**, 391–401 (2017)
21. F.T. Zohora, K.C. Santosh, Foreign circular element detection in chest x-rays for effective automated pulmonary abnormality screening. Int. J. Comput. Vis. Image Process. **7**(2), 36–49 (2017)

22. F.T. Zohora, S.K. Antani, K.C. Santosh, Circle-like foreign element detection in chest x-rays using normalized cross-correlation and unsupervised clustering, in *Medical Imaging: Image Processing, Houston, Texas, United States, 10-15 February 2018* (2018), p. 105741V
23. K.C. Santosh, S. Antani, *Automated chest x-ray screening: can lung region symmetry help detect pulmonary abnormalities?* (IEEE Trans. Med, Imaging, 2017)