IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 11, 2021, accepted January 20, 2021, date of publication January 25, 2021, date of current version February 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3054563

Bibliometric Study on the Use of Machine Learning as Resolution Technique for Facility Layout Problems

PETER BURGGRÄF, JOHANNES WAGNER, AND BENJAMIN HEINBACH
Chair of International Production Engineering and Management, University of Siegen, 57076 Siegen, Germany
Corresponding author: Benjamin Heinbach (benjamin.heinbach@uni-siegen.de)

ABSTRACT Abstract Facility Layout Problems (FLP) are concerned with finding efficient factory layouts. Numerous resolution approaches are known in literature for layout optimization. Among those, intelligent approaches are less researched than solutions from exact or approximating approaches. The recent surge of research interest in Artificial Intelligence, and specifically Machine Learning (ML) techniques, presages an increase of such techniques' usage in FLP. However, previous reviews on FLP research induce that, to date, this trend has not yet emerged. Utilizing a systematic literature review coupled with a k-Means based clustering algorithm, we analyzed 25 relevant publication full-texts from an original sample of 1,425 papers. Our findings corroborate the statement that ML techniques have attracted substantially less research interest than most other resolution approaches. While a few papers used Unsupervised Learning algorithms directly as a solution to the FLP, Supervised and Reinforcement Learning were found to be practically irrelevant. ML usage was significantly higher in FLP-adjacent planning tasks such as group technology. Drawing from experiences with other NP-hard combinatorial optimization problems in manufacturing research, we conclude that Reinforcement Learning is most promising to bridge the evident gap between FLP and ML research. Our study further contributes to FLP research by extending established classification frameworks.

INDEX TERMS Artificial intelligence, layout, machine learning, reviews, production engineering, production management, production facilities.

I. INTRODUCTION

The Facility Layout Problem (FLP) is an important research stream within production research. An FLP is defined as the search for the most efficient arrangement of departments, i.e. facilities, in a plant area subject to different constraints while attempting to satisfy one or more objectives [1]. According to [2], FLPs can be located on the tactical hierarchical layer in organizations, provoking the assumption that sound decisions on facility layouts constitute a significant contribution to economic success of manufacturing companies. Backing this notion is a general agreement on cost-saving potentials in a range of 10-30% for material handling cost-based operating expenses [3]. Furthermore, a robust layout has an immediate measurable impact on the operational performance, as measured by manufacturing lead time, throughput rate, and work in process [4]. Another important driver to market survival in today's manufacturing environment is flexibility [5], which has sparked a research stream within FLP on its own right called Flexible Manufacturing Systems. The authors of this study posit that FLPs are an important organizational topic because a) proper factory layouts are required for an efficient and effective production process, and b) the frequency of FLP encounters is expected to increase with more customized manufacturing.

In the past years, we have experienced a surge of Artificial Intelligence (AI) reports in which AI algorithms have proven to outperform humans in complex, yet very specific tasks such as Go [6], complex online multiplayer games [7], Poker [8], hide-and-seek simulations [9] or emulations of old Atari games [10]. Beyond such primarily demonstrative and trailblazing purposes, the question remains how the capabilities of AI can be used in value-adding industrial settings. It appears to date that widespread recognition of AI in practical real-life applications, especially in production or manufacturing research, is still pending. In [11], we found that

The associate editor coordinating the review of this manuscript and approving it for publication was Hiram Ponce.

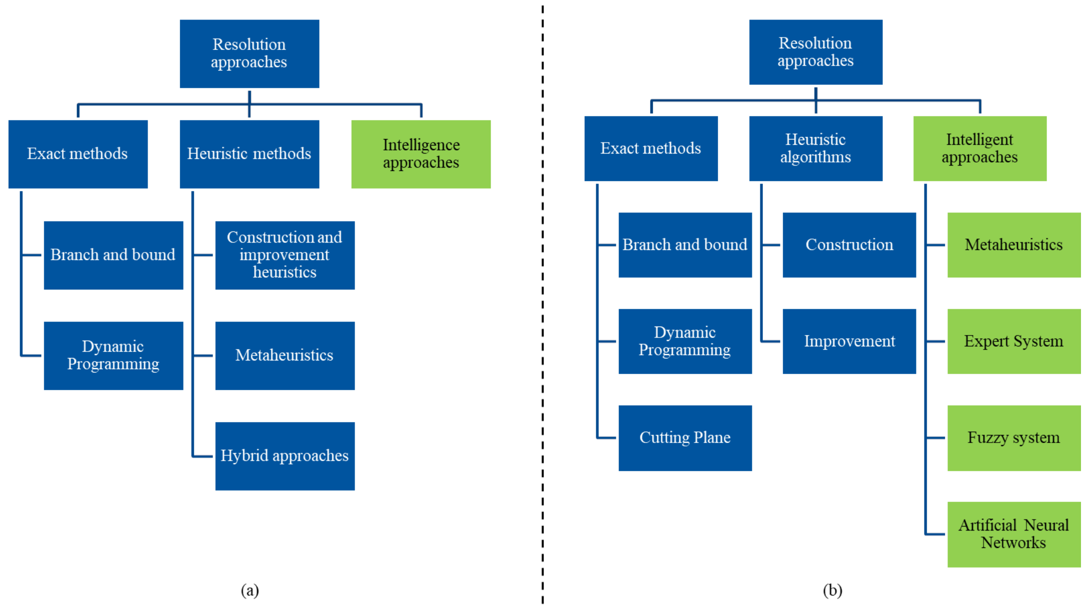

FIGURE 1. Juxtaposition of the FLP frameworks used by Drira (2007) (1a) and Moslemipour *et al.* (2012) (1b) to classify resolution approaches.

AI has predominantly been used in production for scheduling and quality control; however, we assert that AI will play a more vital role in higher-level production management planning and decision-making problems that lead to *cyber production management systems*. Several previous extensive reviews in the field of FLP have examined this research field concerning the *resolution techniques* used [2], [12]–[16]. FLP are NP-hard combinatorial problems [17] which can be solved via approximating bio-inspired meta-heuristic algorithms yielding a fair solution rather than an optimal one, e.g. Genetic Algorithms, Particle Swarm or Ant Colony Optimization, or recently Coral Reef Optimization [18]. The adoption of Machine Learning (ML) - being a subset of AI - for complex real-life problems by both researchers and practitioners has made this field a dynamic area of research [19]. Also, ML is known for its ability to handle many problems of NP-hard nature [20]. Thus, it appears counter-intuitive that the aforementioned reviews on FLP merely indicate a scarce use of *intelligent approaches* in FLP while these are widely recognized as one class of these techniques. Models and algorithms that commonly star in Machine Learning terminology have been listed as one relevant tool, yet we do see a lack of a distinct drill-down in this specific domain. This paper seeks to tackle this limitation. Employing an extensive systematic literature analysis, the objective of this paper is to investigate to which degree AI topics, and more specifically different ML techniques, have pervaded the field of FLP. This study makes three main contributions to the extant literature.

1) We provide an in-depth examination of ML usage as a resolution method for Facility Layout Problems that goes beyond the available reviews in academic literature until thus far.
2) We extend previously known taxonomies of FLP resolution approaches by different meta-heuristics and Machine Learning algorithms.
3) The study highlights unexpected white spaces in ML usage FLP research.

The remainder of this paper is structured as follows: In Section 2, we will build up the AI framework that we use for our review. Section 3 will provide a detailed description of our chosen review methodology. The descriptive results of the literature analysis are presented and discussed in detail in Section 4. Results will be discussed in Section 5. Finally, Section 6 will provide our concluding remarks.

II. BACKGROUND

A. INTELLIGENT APPROACHES IN FACILITY LAYOUT PROBLEMS

Across the previous contributions in FLP research, there is no definite agreement on the definition of intelligent techniques. A large number of publications, especially of type review, seem to go back to the seminal review by Drira [16]. While his framework includes *intelligence approaches* (cf. Fig. 1a), Drira does not specify a second level here. Interestingly, Artificial Neural Networks are introduced as a class of problem formulation techniques rather than resolution approaches. The review itself does not have an own section on these approaches, nonetheless, Expert Systems (ES) are listed as relevant tools.

The framework used by Moslemipour *et al.* [13] is notably similar to the previous one (cf. Fig. 1b) with the large exception that meta-heuristics are considered intelligent approaches here. Moslemipour *et al.* further list

Expert Systems, Fuzzy Systems, and Artificial Neural Networks (ANN) as tools [13].

Yet another picture is painted by Renzi *et al.* [15]. Here, AI is grouped under non-exact, i.e. approximating, approaches along with meta-heuristics (see Fig. 2). This taxonomy retains Fuzzy systems and ANN but drops ES. Renzi identifies four papers related to the use of ANN in reconfigurable manufacturing systems and 42 for cellular manufacturing systems [15]. However, according to their analysis, these papers deal with cell formation and scheduling problems. No papers were identified as being relevant for cell layout problems.

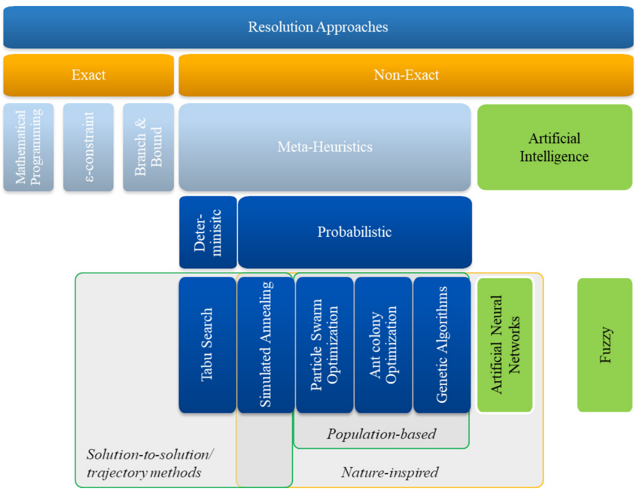

FIGURE 2. Taxonomy of CMS techniques according to Renzi (2014).

Kumar classified a total of five papers as being AI-relevant [14]. However, he includes ANN in evolutionary techniques along with Genetic Algorithms (GA), Ant Colony Optimization (ACO), and other algorithms more commonly denoted as meta-heuristics, which in turn, Kumar all groups as Artificial Intelligence.

The probably most comprehensive framework can be found in [12] (see Fig. 3) where intelligent approaches are defined as a category of resolution techniques which is further subdivided into ES and ANN. In their extensive review, they find a total of five papers of which Expert Systems make up three and ANN the remaining two. Not only is the low number surprising by itself, the authors merely name these two concepts as a sub-class and do not provide further details as to their usage.

The richness of taxonomies for intelligent approaches certainly leads to wide interpretability of results and thus to potentially competing insights. We do not intend to draft an additional and entirely new framework. Instead, due to its recency and comprehensiveness, we extend the recent classification scheme in [12] for intelligent approaches such that it becomes workable for our review on ML techniques in Facility Layout Problems.

B. A CONCEPTUAL FRAMEWORK FOR CODING INTELLIGENT APPROACHES

ML techniques belong to a problem representation framework called *sub-symbolic AI* [21]. Following a bottom-up, i.e. data-driven, approach, sub-symbolic AI attempts to create structures that can learn intelligent behavior through information-processing mechanisms [22]. The counterpart to sub-symbolic AI is *symbolic AI*, which is originally defined as a system that has sufficient means for general intelligence by symbol manipulation [23]. Here, knowledge is encoded in terms of explicit symbolic structures, and inferences are based on handcrafted rules that sequentially manipulate these structures [24]. Preparatory work on our review study, most notably the scoping review as described in section III-A, showed that the umbrella term "Artificial Intelligence" is inherently chosen over usage of the term "Machine Learning". We, therefore, pick the starting point of AI as the top-level node in our taxonomy, followed by the two branches symbolic and sub-symbolic AI. As indicated above, the three concepts that have attracted the most research interest within intelligent resolution approaches regarding their occurrence in conceptual review frameworks are Fuzzy systems, Expert Systems (ES), and Artificial Neural Networks (ANN). ES, are a well-studied member of symbolic AI [21]. However, research interest in ES has been declining continuously in operations research [11], [25], [26]. Considering ML definitions by [27], [28] and [21], a key trait in ML, is the capability to *independently* learn and be able to find solutions to previously unseen situations. As ES are a one-to-one representation of expert domain knowledge which are known to lack adaptation capabilities in new situations, we exclude ES from our synthesis scope. Nonetheless, as previous reviews have indicated, we acknowledge that ES were an important research topic in the past and therefore retain the concept for coding purposes. Regarding Fuzzy systems, we argue in line with Drira *et al.* [16] and Hosseini-Nasab *et al.* [12] that Fuzzy numbers and Fuzzy systems ought to be classified as data types for layout formulations. This means that any variable used in the FLP can either be notated as crisp, as stochastic, i.e. sampled from a distribution, or as a degree of membership, ranging from 0 to 1, in a fuzzy set. Thus, as opposed to some authors before us, we will not include Fuzzy systems in the class of intelligent resolution approaches.

Lastly, ANN have a long history in AI since being introduced in 1943 by McCulloch and Pitts [29]. While they have recently gained traction thanks to breakthroughs in some fields like image recognition [30], [31], they do not comprise the only available model in ML. In fact, a substantial amount of different models and algorithms have been designed. ML can be subdivided into the classes Supervised Learning (SL), Unsupervised Learning (UL), and Reinforcement Learning (RL) [32], [33]. *Supervised Learning* is a data-driven learning paradigm in which inputs are mapped to outputs. Methods include, among others, Artificial Neural

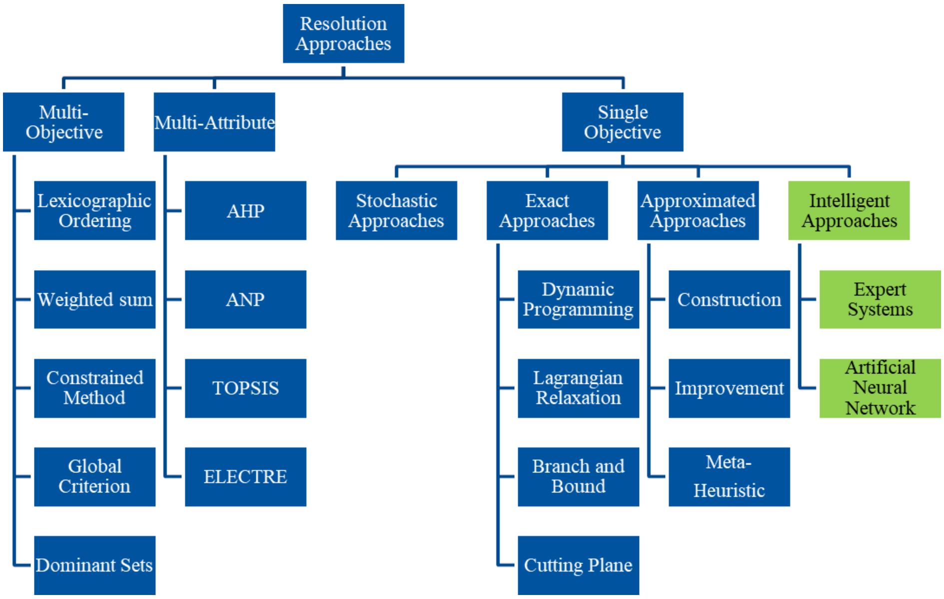

FIGURE 3. Resolution approach classification used by Hosseini-Nasab *et al.* (2018).

Networks, Logistic Regression (LR), K-Nearest-Neighbor (k-NN), Support Vector Machines (SVM), Random Forest (RF), Decision Trees, and Bagged Trees (DF) [34]. *Unsupervised Learning* on the other hand is a branch of Machine Learning method that aims at finding hidden patterns in unlabelled data instead of mapping known inputs to known outputs [35]. Typical examples of UL are clustering (e.g. via k-Means), association rules, and self-organizing maps (a specific architecture of ANN) [36]. *Reinforcement Learning* mimics human decision-making by having an agent gradually collect information from interaction with its environment. For each action performed in a distinct state, the agent collects a reward which over the course of several iterations shapes a policy that describes the optimal action to take for each state [37]. Prominent RL algorithms are Q-Learning [37], state-action-reward-state-action (SARSA) [37] or Deep-Q-Networks (DQN) [10], to name but a few. In addition to those three, *Deep Learning* (DL), uses multiple layers to progressively extract higher-level features from raw input data [38], typically using Neural Network architectures with a significant amount of hidden layers. DL can be used both in a supervised, semi-supervised or unsupervised fashion [39]. Plus, Google's DeepMind has fostered attention *in Deep Reinforcement Learning* (DRL) by outperforming human players in ATARI games [10]. Showing links to all other branches, we consider DL as a cross-sectional collection of methods. However, for ease of visualization, we grant it a separate branch in our framework and highlight the aforementioned particularity with an asterisk (*).

Artificial Neural Networks have been proposed in a multitude of architectures, a sample of which is presented below as listed by Leijnen and van Veen [40]. The assignments to learning paradigms in square brackets were added by the authors of this paper.

- Feed Forward Neural Networks (FF) and multilayer perceptrons (MLP) [SL]
- Recurrent Neural Networks (RNN) [DL]
- Long Short-Term Memory (LSTM) [DL]
- Autoencoders (AE) [UL]
- Hopfield Networks and Boltzmann Machines (HN) [UL]
- Convolutional Networks (CNN) [DL]
- Generative Adversarial Networks (GAN) [UL]
- Kohonen Networks, or self-organizing maps (SOM) [UL]

A synthesis of the above discussion is depicted in Fig. 4. This figure shows the conceptual framework for ML in FLP that we use to code retrieved articles. For our study, we adjust the definitional review scope for intelligent approaches by enriching the previously used concept of ANN, by dropping Expert Systems due to their lack of recent research focus, and by incorporating our understanding of ML approaches in the form of lowest-level tree nodes. Drawing from [15], we further added common meta-heuristic methods as concepts for a better overview.

Note that the visualization in Fig. 4 is not meant to be exhaustive, but seeks to underline the complex structure and the diverse nature of currently available and common ML concepts. Hence, only those concepts discussed above are shown in the figure. Since we made no changes to the framework in [12] regarding exact and stochastic approaches, we exclude these branches from our visualization for better legibility. Our additions to approximating and intelligent approaches are highlighted in green color. The following

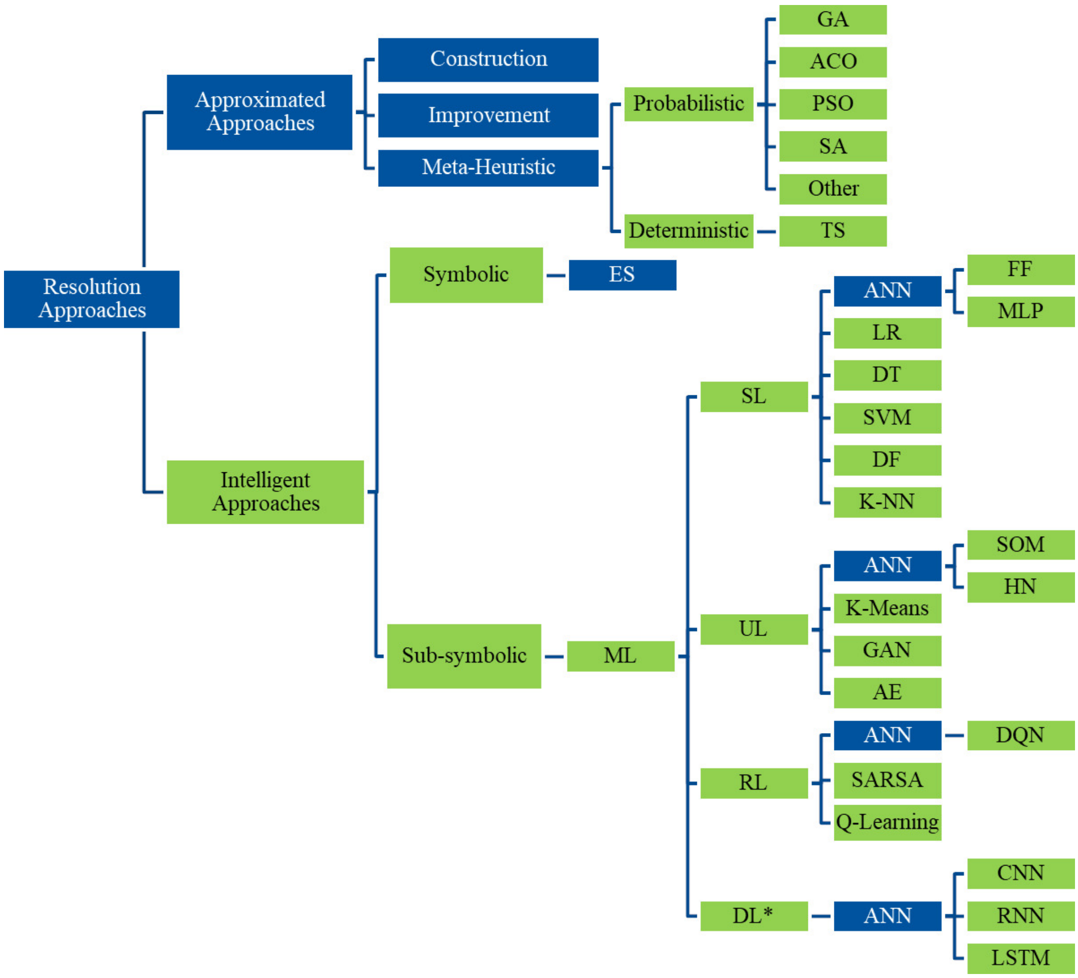

FIGURE 4. **Synthesis of the ML framework used for coding in this review: Additions to the previous framework are shown in green.**

section will present in detail how the framework is leveraged in our review.

III. METHODOLOGY

A systematic literature review is a research methodology to "review [...] research literature using systematic and explicit, accountable methods" [41], that has become the gold standard to synthesising findings from several studies investigating a similar question regardless of the discipline where it is used [42]. Systematic reviews tend to have a narrow research question, which, combined with strict quality criteria, result in a reasonable amount of studies to be included in the review whose results add up to answering the research question. This sort of review can be termed aggregative [41] and qualifies as conceptual rather than empirical research.

To provide a comprehensive overview of the state-of-the-art of research on Artificial Intelligence, and specifically Machine Learning algorithms, in factory planning, this paper adopts the methodology for conducting systematic literature reviews proposed by Boland *et al.* [42]. As it became apparent early on that the review would face a substantial number of publications, we extended this methodology by an algorithmic clustering step, see Fig. 5. The methodological details of each step are presented in the following sub-sections.

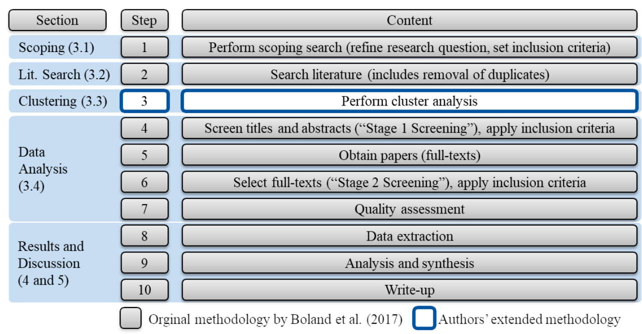

FIGURE 5. **Extended SLR methodology building up on Boland *et al.* (2017).**

Furthermore, following the advice of Gough *et al.* [41], we assembled a review team consisting of this paper's authors to increase inter-rater reliability for the data analysis phases.

A. SCOPING SEARCH

The scoping review we conducted indicated a large number of returns in the databases, ranging beyond a six-digit figure. This large number of records exceeded the research team's capacity to screen and retrieve. As proposed by Gough *et al.* [41], the scoping search motivated the authors to employ automatic clustering as a text mining approach. Therefore we included Step 3 (Cluster Analysis) in the review process. The purpose of this step is to partially automate the classification process in Stage 1 Screening (Step 4).

1) RESEARCH QUESTION

As outlined above, the purpose of this paper is to investigate the degree to which ML techniques have been used to solve Facility Layout Problems. Thus, we formulate our research question as follows:

RQ1: How have different Machine Learning algorithms been used as resolution techniques for Facility Layout Problems?

2) INCLUSION CRITERIA

To identify a body of literature that is relevant to answering the research question, we established the following quality criteria:

1) Include peer-reviewed academic work, such as books, journal articles, and conference contributions
2) Include theses and dissertations
3) Omit other grey literature (e.g. magazines, commercial websites, company white papers, patents)
4) Exclude publications not pertinent to Facility Layout Problems
5) Exclude all non-English publications
6) Include only articles newer than 1987 (review scope of [12])

For the detailed analysis, we set the following inclusion criteria (IC) to be applied to all references short-listed for Stage 1 Screening:

- *IC1:* The paper is relevant to the research area "Facility Layout Problems"
- *IC2:* The paper indicates the use of any of the ML techniques from the framework in II-A

B. LITERATURE SEARCH

1) SEARCH TERMS

Fig. 6 displays the combinatorics of terms commonly used in FLP research. The most common term is "Facility Layout Problem", as highlighted in blue. However "plant layout design" [43], "facility layout planning" [44], "site layout planning" [45]), "space layout" [46] or simply "layout problem" [47] can be found as well. This combinatorics indicates a richness in research that needs to be addressed in the systematic search. There are two important considerations for the search strings. First, a search string including three components as depicted below proves to be difficult to enter or resolve in some of the databases used in this research. A uniform way to address this is a long string of OR-combinations, which may be inapplicable due to character input restrictions. However, any database query for the former two components should pick up publications including all terms in the third component. Using all three components in Fig. 6 also increases the risk of missing an important stream not involving any of the terms. Therefore, we chose to eliminate this part from the search string and only regard the combinatorics from the left-hand side of the dashed line. Second, the inclusion of the first component is vital to address the correct research communities. Excluding terms such as "facility", "factory", "plant" and the like showed to pick up publications from hospital and operating theatre designs or integrated circuit layout design. While there may be relevant knowledge on RL applications in those adjacent fields, our initial scoping revealed an extensive literature base already for the domain of facility layouts. Therefore, we opt to exclude publications not having a distinct manufacturing focus. Yet, we acknowledge that a thorough search in these fields whilst highlighting similarities and congruencies in the approaches or logic can be a worthwhile topic for future research.

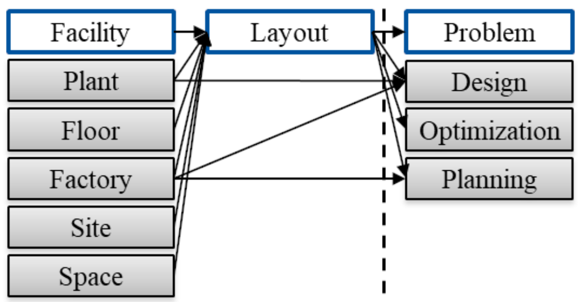

FIGURE 6. Overview of relevant search terms.

2) DATABASES

The following nine databases were queried with the search strings above: ScienceDirect, Web Of Science, JSTOR, EBSCO Host, Emerald Insight, IEEE Xplore, ProQuest, and ACM Digital Library. This selection is based on our previous good experience with these databases (see [48], [49]). Furthermore, we tackled the file-drawer bias by searching in OpenReview for potential AI-related contributions in FLP that may have failed a peer-review process.

3) DATASET PREPROCESSING

As the scoping indicated, our search strategy yielded a large number of retrievals. Therefore, in a first step, the full dataset has been cleaned from all irrelevant entries such as front or back matters, volume contents, errata, glossaries, forewords, keyword indices, and the like. Titles, abstracts, and keywords were changed to lower case to facilitate the identification and removal of duplicates. In the last step of preprocessing, we recreated a title-abstract-keyword search that was not supported by most databases. The purpose of this approach is to exclude entries which do include the search terms but without close relation to the search field. Using a short macro, we re-applied the above search string to the database fields "Title", "Abstract" and "Keywords", thus eliminating a large number of artifacts we consider false positive entries in our dataset. The result of this step is a cleaned dataset ready for Stage 1 Screening.

4) SNOWBALL SAMPLING

Given that *snowball sampling* is a powerful data collection tool in systematic reviews which can contribute up to half of the relevant references in a study [50], we applied this sampling technique to the review articles from II-A. We performed snowball sampling during a longer period between September and October 2019. The references obtained via this way were added to the preprocessed dataset.

C. CLUSTER ANALYSIS

We employ the NLP-based clustering algorithm designed by Weißer *et al.* [51]. This k-Means based algorithm uses tokenization (word separation), stop words removal and TFIDF vectorization to identify the most relevant words per cluster that describe its thematic relevance. Clustering was performed on the "Title" column of the cleaned dataset for Stage 1 Screening. From the previous reviews we synthesized 19 possible clusters consisting of different exact, meta-heuristic and intelligent techniques and one cluster for un-classifiable papers. Accordingly, we set the algorithm to creating 20 clusters. Thus, we selected a k_{Max} for the Elbow method of $k_{Max} = 20$, and an explained variance of 0.7. The algorithm was tasked to indicate the five most relevant top words in each cluster.

D. DATA ANALYSIS

1) STAGE 1 SCREENING

For the Stage 1 Screening procedure, the remaining papers in the dataset were screened manually. To increase inter-rater reliability, the final dataset for this step was split evenly and at random and distributed among the three members of the research team. Papers were presented to the researchers only via their title, abstract, and keywords. Each researcher coded their respective papers according to the framework presented in Fig. 4 (i.e. categorical coding). The concepts within the "Exact" branch of Fig. 3 were united under that term. The multi-attribute concepts describe group decision techniques and were thus coded as "qualitative". Along with the lower-level entries in this conceptual framework, the labels "unknown" or "none" (for publications, e.g. of conceptual nature, which did not clearly report on any resolution technique), "not FLP" (for papers that discuss layout issues but not within the facility problem domain, i.e. not meeting IC1) as well as "review" (for publications of type review which are assumed to refer to the majority of publications included in this dataset). We determined that for references coded with an ML-relevant concept from the above framework both inclusion criteria IC1 and IC2 applied. The respective papers thus progressed to Stage 2 Screening.

Upon completion of the first coding round, a random subset of references at a size of 20% of each researcher's dataset (i.e. approximately 65 publications each) was re-distributed to another researcher in the team for a second blind coding. Any mismatches in the coding results were mediated by the third researcher. This number of interventions was considered sufficiently low to warrant a 20% sample size as amply sized.

By the end of the above screening process, we pooled the relevant coding results from the two sources clustering and manual screening.

2) OBTAIN PAPERS

All papers flagged relevant for Stage 2 Screening could be obtained as full-text. Each paper was given a unique two-digit ID for further coding processes.

3) STAGE 2 SCREENING

In Stage 2 Screening, initial coding results from Stage 1 Screening were verified. Full-texts were screened by the lead researcher. In a separate slide deck, a specific slide was maintained for each publication. Full-text snippets were collected in support of or in opposition to both inclusion criteria and new codes were applied where necessary. As a result of an emergent low-precision pattern, two different final codes were applied to the publications in review: papers to which IC1 and IC2 applied were assigned the code "Prio 1". To enrich the final synthesis sample, papers which reported on a distinct ML technique from the framework, albeit not meeting IC1, were labeled "Prio 2". For these two classes, the publications were labeled according to the Machine Learning paradigm in question, i.e. Supervised Learning, Unsupervised Learning, or Reinforcement Learning. The final subset containing all "Prio 1" and "Prio 2" papers was used for the synthesis of our findings.

IV. DESCRIPTION OF ANALYSIS RESULTS

This section reports the results of each review step from section III according to the PRISMA method [52]. The numbers reported hereafter can be traced using Fig. 7.

The literature searches were performed in late September 2019. We identified a total of 11,851 publications across all nine databases. Only OpenReview did not incur any results. 1,903 irrelevant entries were removed. Another 191 entries were removed for being duplicates. In data preprocessing, we excluded an additional 6,811 publications through the recreated title-abstract-keywords search. Furthermore, before entering the Stage 1 Screening process, we examined all titles, and where necessary the abstracts, too, as to the relevance to FLP (IC1). IC1 did not apply to 1,601 entries, leading to their removal. This resulted in a final dataset for screening consisting of 1,303 references.

The dataset was first fed into a clustering algorithm which assigned 327 papers into clusters relevant for further analysis: The algorithm identified a cluster for the words "artificial",

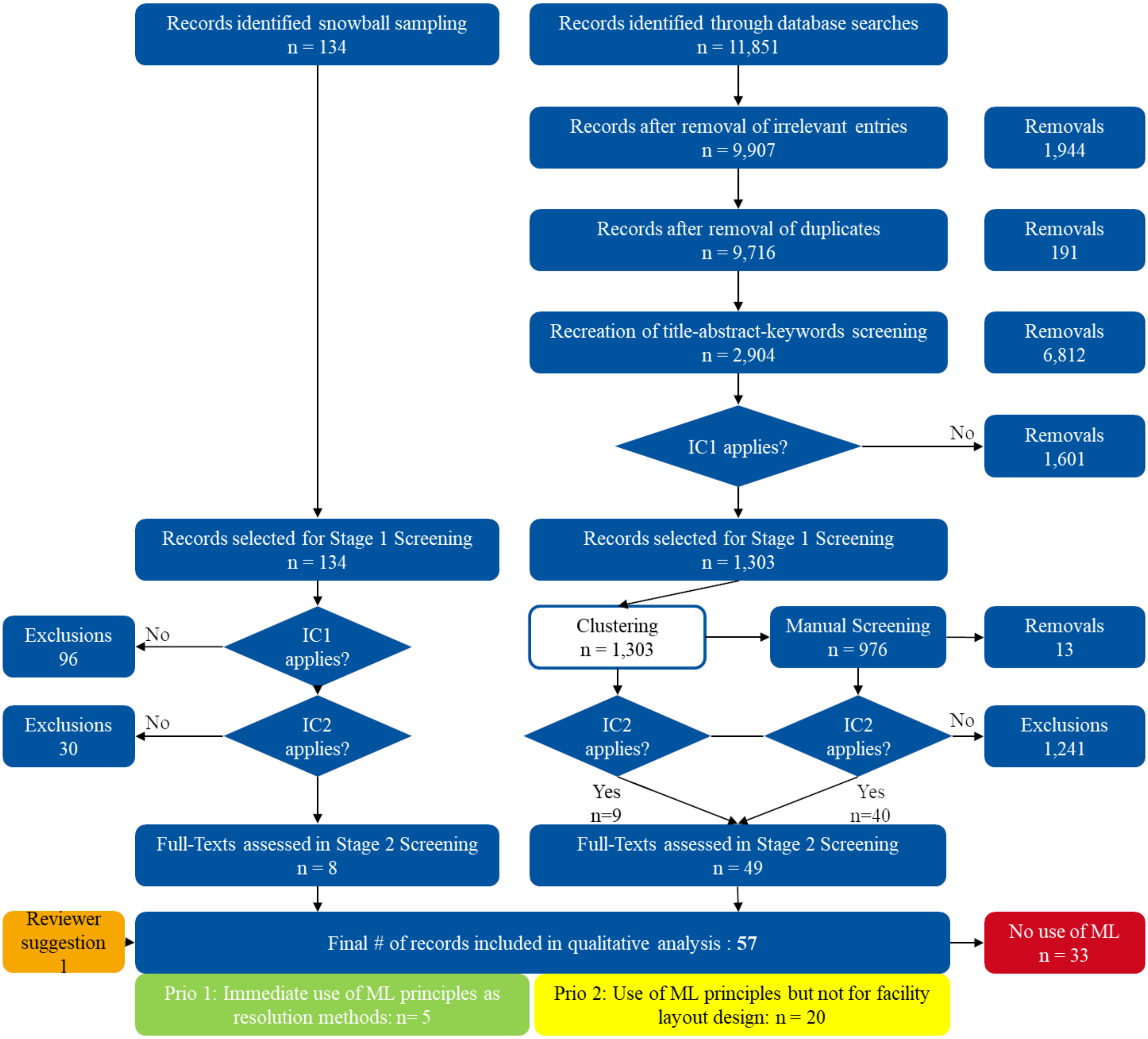

FIGURE 7. Stepwise report for all removals during the review process.

"intelligence", "neural", "network", "system", containing a total of 20 papers. Out of the 20 papers considered relevant by the clustering, nine progressed further to Stage 2. Three papers did have an ML focus but were excluded for being review articles. The remaining eight papers were re-coded according to the title and abstract information. The algorithm grouped five further clusters which could be unambiguously mapped to the meta-heuristic concepts known in FLPs: Genetic Algorithms, Ant-Colony Optimization, Particle Swarm Optimization, Tabu Search, and Simulated Annealing. A total of 307 papers were assigned to these clusters. These were screened again manually to verify the inclusion choices made by the clustering algorithm. The classification could be affirmed for 177 papers, the remainder was re-coded accordingly. As a result of the clustering, 976 papers remained for manual assessment by the research team.

In Stage 1 Screening, another 13 references were removed manually: eight were duplicates, two were notices of articles being retracted, and for the remaining three no abstracts or keywords could be retrieved. On top of the nine ML-relevant references picked up by the clustering, Stage 1 Screening yielded another 40 papers that the research team considered potentially relevant. The 1,241 papers excluded for not satisfying IC2 were coded with a label describing the resolution method used.

From the snowball sampling screening of the previous reviews, depicted on the left-hand side of Fig. 7, we obtained a total of 134 additional papers, 27 in the backward direction (cited articles) and 107 in the forward direction (citing articles). In the forward search sample, we identified three new papers that were potentially relevant. For backward sampling, we revisited the papers that were used in the previous analyses. The paper of Renzi *et al.* [15] contained 17 potentially relevant papers. Although the original study concluded that none of their reviewed publications dealt with layout generation, we marked two papers for further examination in Stage 2 Screening. As outlined before, [14] used a diverging AI definition. Based on our framework, we found no ML-related papers in this sample. Hosseini-Nasab *et al.* [12] reported five AI-related papers which dealt with intelligent approaches such that they progressed to Stage 2. On closer scrutiny of their references 186 and 133, we cannot confidently assign the content to either ES or ANN and thus contest the above number of five AI-related papers, leading us to only relay three papers from this sample to Stage 2 Screening. Thus, snowball sampling yielded a total of eight papers we considered relevant for full-text analysis.

In Stage 2 Screening, we assessed a total of 57 papers regarding IC1 and IC2. We found five papers that reported on ML techniques used as a resolution method for FLPs (Prio 1). Given that we considered this a relatively low number for a proper narrative analysis, we flagged another 19 papers which dealt with ML techniques, yet were used primarily as support functions for layout problems (Prio 2). The papers not labeled with any of these two categories we re-coded according to the actual resolution technique used, if any. In the end, we obtained 24 papers for the synthesis in the discussion section.

Fig. 8 shows the final coding results of our review study. Not shown in this figure are references coded as "reviews" and those excluded for not meeting IC1. The data show that the cumulative figure for intelligent approaches, marked as the green bar, is marginal compared to all other approaches. In line with the review study conducted by Hosseini-Nasab *et al.* [12], the total number for approximating approaches is by far greater than that for exact ones, followed by stochastic and lastly intelligent approaches.

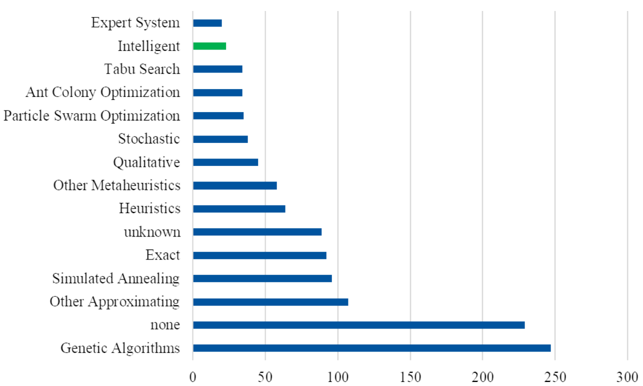

FIGURE 8. **Aggregated results for resolution methods after Stage 2 screening (n = 1212).**

The dominant role of genetic algorithms to solve FLPs is uncontestable.

As shown in our framework, we have purposefully excluded Expert Systems from our scope of analysis for not being relevant to Machine Learning. Our analysis shows that, albeit by a narrow margin, Machine Learning approaches have overtaken Expert Systems in terms of cumulative numbers around the year 2010 (see Fig. 9). ES stagnation is in line with the aforementioned decline in research interest. The data in Fig. 9 also show that intelligent approaches were outnumbered by qualitative, stochastic, and several approximating approaches in the early 2010s as well.

V. RESULTS

A. STUDY FINDINGS

In the following, we will discuss our findings grouped by the learning paradigms from our research framework in Fig. 4. Deep Learning is treated as a cross-sectional paradigm and will thus be presented accordingly within the other three, where applicable.

1) UNSUPERVISED LEARNING

Within Unsupervised Learning, we found three predominant concepts in FLP research: clustering techniques, self-organizing maps (SOM), also referred to as Kohonen Networks and Hopfield Networks.

In [53] and [54], a potential fields modeling method is used to realize self-organization in semiconductor manufacturing. Machines are represented by the nodes of the neural network and assigned machine properties such as "idle", "in process", "malfunction", "maintenance" and "buffer with product". Then, attraction and repulsion fields, the latter type to avoid machine overlap, are generated for the machines according to these properties. Attraction field attributes for transporters and buffers are modified by the requirements of the product. As a result of their simulation, the total travel distance for the product being produced has been minimized such that the most frequented machines in the manufacturing process were orbited by their respective support functions.

Similarly, Tsuchiya *et al.* [55] employ an artificial two-dimensional maximum neural network to solve a QAP where N facilities are to be placed on an N^2 location array. The number of nodes in the ANN matches the number of locations. The ANN provides a gradient descent method to minimize the fabricated energy function. All nodes' activation functions are step functions between 0 and 1, to be activated for the node that has the strongest weight at a given time step. Weight strength is determined via Manhattan distance, i.e. shorter distances between facilities, and thus lower transport cost, yield higher weights.

Zhang *et al.* [56] propose the use of Hopfield networks for construction site layout planning. The function of the Hopfield network is to assign temporary establishments (such as tower cranes on construction sites) to pre-defined locations available for layouts, using the knowledge passed from an

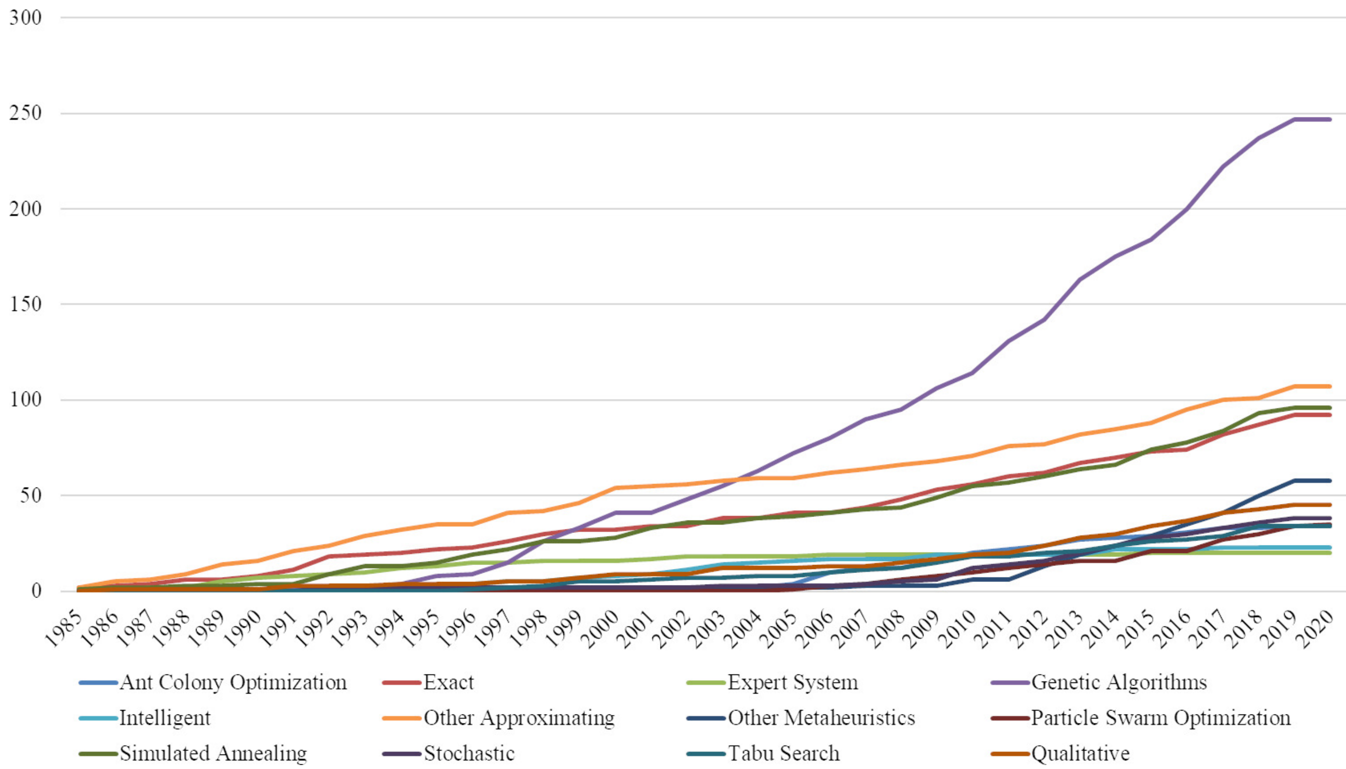

FIGURE 9. **Time series analysis showing cumulative results for resolution methods.**

Expert System and info input by a user. The construction sites are mapped to the network nodes and the energy function consisting of distance metrics is being minimized via forward computing until convergence is achieved and an optimal layout thus found. However, the paper does not provide a clear description of how the above-mentioned assignment is performed in detail. It appears that the framework is run once; algorithmic support, such as Simulated Annealing, to find a stable state through adjustment of network weights is proposed yet not explained further.

Such algorithmic support is described by Yeh [57]: using a hospital case study with 28 facilities formulated as quadratic assignment problem (seven sites on four floors), the authors extend previous approaches to an Annealed Neural Network which combines characteristics of the Simulated Annealing algorithm and the Hopfield Neural Network. The purpose is to maintain the rapid convergence capabilities of the Neural Network while preserving the solution quality afforded by Simulated Annealing. This way, the Hopfield Network's major disadvantage of not being able to find the global energy minimum in a single run, but settling with the nearest local minimum instead, is overcome. The Hopfield Network recomputes the permutation matrix which describes the strongest links of assignment of facilities to sites. Such an assignment is determined by a local energy minimum, calculated as a function of layout preference (space requirements), interactive preference (adjacency), and constraint violation (penalty). According to the authors, this approach makes it independent of the size of the search space, unlike many other approaches.

The study of Vakharia and Wemmerlöv [58] analyses the impact of dissimilarity measures and clustering techniques on the quality of solutions in the context of cell formation. Instead of SOM, they propose a clustering algorithm called "set merging" where the smallest entity considered during the merging process is the cluster rather than the objects within each cluster. As with the examples mentioned before, layout resolution is not within the scope of their contribution.

In contrast to the examples provided above where self-organizing maps (SOM) can be regarded as resolution technique by creating layouts, this type of algorithm has found wide application as a support function in the above-mentioned group technology. This type of application is within the scope of Prio 2 papers. Examples of this can be found in [44], [59]–[62] and [63]. For the same purpose, [64] provide an extension to their previous SOM by using an Adaptive Resonance Theory-based clustering which yielded better grouping results. Another approach that reports the use of SOM is presented by [65]. Here, SOM is used in conjunction with ALDEP, a member of the class construction algorithms. While they employ ALDEP for the actual layout resolution (and the paper has been coded by the researchers as such), SOM is used to tackle problem size issues as described in the introduction section. As the ALDEP algorithm can only handle problems with up to 14 facilities, SOM groups facilities into clusters, and each cluster's individual layout is created via ALDEP.

2) SUPERVISED LEARNING

The most remarkable and unexpected finding regarding Supervised Learning approaches is that we found none which satisfies both our IC1 and IC2. For those satisfying IC2, we found evidence for supervisors to act either as classifiers (e.g. an expert rating surrogate [66]) or as regressors (e.g. in the case of hoisting times [67]) without a distinct tendency towards one of these two problem classes. For those that do not satisfy IC1, contributions either dealt with cell formation problems in cellular manufacturing research, quite similar to our analysis on Unsupervised Learning approaches, or relied on other algorithms as resolution techniques.

A prominent approach in ML research for FLPs is to combine Neural Network capabilities with meta-heuristic approaches. We found that for these approaches, the key functionality of the ANN used is the capability to represent non-linear relationships to compute input parameters for the construction or optimization algorithms in use. Tam and Tong [67] and Azimi and Soofi [68] both present solutions that make use of Genetic Algorithms and Neural Networks. In [67], a simple back-propagation network (five input nodes, one fully connected hidden layer, one output node) predicts hoisting times (supply and return) of tower cranes on construction sites. The network is trained with 1,000 cases from six projects whereas the test datasets consisted of other data points from a further project. The GA in the setup creates a layout based on a population and evaluates their fitness by computing total traveling times using the predicted hoisting times.

Azimi and Soofi [68] make use of ANNs to estimate the make-span of production jobs in different layout scenarios: a number of different layouts is generated and the make-span is determined using discrete-event simulation. These input-output pairs, with the input being a vector of machine locations and transporter allocations, are used to train an ANN to approximate the objective function. However, a seemingly low number of 24 layout simulations was used for training. As in [67], the layouts are designed by a GA, NSGA-II in this case. The authors claim that the ANN serves as chromosomes fitness function evaluator, but details of how the mechanism to evaluate the material handling cost-based fitness function works are not provided.

A GA is employed to generate layouts in an optimization phase in [66] as well. The resulting set of layout solutions is then presented to an expert for evaluation according to the factors material flow, adjacency requirements, and aspect ratios using a five-point Likert Scale. The evaluation results are then used to train an ANN which will act as a surrogate for the expert primarily to avoid fatigue. The training and validation datasets used here seem sufficiently large (365 and 181 cases for two tests at an approximate 50% split for training and test data). When subjected to test data or new layout instances, the trained model will predict the mark that the expert would have given to that layout by applying a softmax function on the five output classes (corresponding to the 5-point Likert evaluation) at an accuracy ranging between approx. 80% and approx. 91%, depending on which expert evaluates the ANN predictions. Thus, the key of this contribution is to incorporate human knowledge into the FLP evaluation phase.

The usage of expert ratings in combination with Machine Learning can be found in [69] as well. Here, a pool of layouts is generated using SA, Clonal Selection Algorithm, and Firefly Algorithm. Relevant criteria are identified by submitting them to a rating by company experts. Then, an Unsupervised Learning approach leveraging Principal Component Analysis (PCA) is used to obtain a subset of uncorrelated criteria. Following an efficiency analysis using Data Envelopment Analysis, the authors use Supervised Learning based on training linear and logistic regression models to predict layout rankings. Model optimization is performed with k-fold cross-validation.

Further, Deep Learning systems that include expert knowledge are also presented by Chung [70], [71]. Here, neuro-based Expert Systems using Bidirectional Associative Memory (BAM) Neural Networks are proposed. BAM Neural Networks are a form of Recurrent Neural Networks, a member of the Deep Learning class, that have a simple architecture, consisting of two fully connected layers, and need fewer data points than back-propagation. Each neuron in a layer stands for one abstract concept (assertion) of a layout rule. The key is to translate verbal if-then rules from layout requirements into a matrix reformulation of weights (i.e. the memory) and match the rules to layouts. A variety of training examples can be generated from empirical simulations, historical data, or layout software programs. The BAM system analyses the layout clues, including closeness relationship, relative position, area size, and site constraints, to identify patterns and relationships that may subsequently lead to rules for laying out facilities. While this system can aid a planner by classifying a layout according to specified input requirements, i.e. estimate the layout's fitness, this approach does not in itself create layout alternatives other than a generalization from training examples. The layout design is accomplished by construction algorithms such as CORELAP.

Group technology problems without immediate layout considerations have also played a role in Supervised Learning. Chen and Sagi [72] design a Neural Network that predicts certain manufacturing cell design parameters such as layout strategy, number of operators needed, part mix, and production sequence. The Neural Network design consists of two cascaded networks: a design Neural Network takes performance measures as input and outputs cell configurations. These, in turn, are fed into a control Neural Network which will output cell control function ratings. Training occurs with pairs of performance measures known a priori and cell configurations as well as complexity requirements for control functions respectively. Layout strategy is mentioned as one predictor, yet this does not refer to layout types (i.e. transfer line, loop, warehouse, etc.) but to a binary classification of machines grouped vs. using two separate lines for the examined products to be manufactured. Layout design is not within the scope of the paper.

Rao and Gu [73] use a combination of Unsupervised and Supervised Learning to solve a cell configuration problem. First, a cluster center-seeking algorithm similar to k-Means takes a production flow analysis chart as input and identifies cluster centers which are most disparate from each other according to the target number of clusters specified by the user. The cluster centers are input to a Neural Network. While the authors state that the network uses "reinforced learning", as "vigilance values are modified during the training process to inflict punishment (negative feedback) for an erroneous classification" (p. 1058), the description of top-down and bottom-up weight adaptations imply that this refers to backpropagation. Moreover, it is stated that the network then acts as a classifier that maps new input patterns to the existing exemplars. While these reflections allow us to classify this contribution as Supervised Learning, it is not a resolution approach for an FLP.

3) REINFORCEMENT LEARNING

Our sample further yielded a small number of papers that could be attributed to Reinforcement Learning. We surprisingly found no publications that directly use Reinforcement Learning as a resolution technique. The remainder of this sample is discussed below.

Tarkesh *et al.* [74] present an approach to FLP based on a multi-agent system in which the plant layout is generated by the interactions of the agents. Each agent corresponds to a production unit with inherent characteristics, emotions, and a certain amount of money, which together describe its utility function. The available credit of an agent is adjusted during the training phase, while each agent tries to adjust its utility function to minimize its total layout costs in competition with others. Since each agent's pricing decision has a strong influence on the pricing proposals of other agents in later decision iterations, and since each agent's pricing proposal is directly related to the last system state, this procedure can be considered a Markov chain and thus modeled as a Markov Decision Process (MDP). This modeling is not discussed in their paper. However, the iterative adjustments of the amount of money available to the agent are represented by a modified Q-Learning notation. Nonetheless, this paper does not satisfy our IC2, as it uses a multi-agent system as a resolution approach and is thus excluded from our Prio 1 pool.

In [75], a self-organization approach for cellular manufacturing systems based on Reinforcement Learning is proposed. The FLP is in essence represented as a quadratic assignment problem in the sense that cells are available for placing equipment on a grid. The reward is computed as travel time between cells and operation times, based on a certain operation sequence. An underlying MDP formulation cannot be distinguished as the notation differs from what is currently predominant in RL: in this paper, the action space is referred to as "tactics". A definition for the state space is not given. The authors further do not specify whether they employ any well-known RL paradigm, such as Q-Learning as above. However, the tactic selection probability strongly resembles typical epsilon-greediness from RL. Furthermore, it becomes clear that this FLP is designed as an episodic task, as the pallet unit moves back to the entrance after obtaining a reward. While this is an interesting early approach to RL in FLP, the goal in this paper is to choose one out of three different configurations (line, flexible, or island) for the manufacturing system. Layout planning, on the other hand, is explicitly mentioned as a future research topic effectively putting this paper beyond our research scope (i.e. Prio 2).

The paper presented by Ono *et al.* [76] uses tree search and a backtracking algorithm with time-reducing heuristics to make up a layout and schedule synchronously. Some characteristics closely resemble standard RL concepts: for rearrangement of the floor plan, the equipment can move along a grid (i.e. the action space is top, down, left, right, rotations, and idle). The reward, or penalty, is computed according to the number of workers needed to fulfill a given action. The number of turns taken in a simulation corresponds to the training epochs in RL. Plus, the mechanism to award bonus points for favorable layouts resembles rewards structures of MDPs, e.g. "if an action reduces the distance (...) add points to its weight". While time constraints point to an episodic design, spatial constraints, such as the freeness of collision, provide a hint to a finite state space. However, the construction of the tree to be searched is not specified, such that no concluding statements about the state space can be deduced. The search can occur in two modes: search alone or in cooperation with a user. In the latter mode, an interface enables users to give instructions to the system to improve the quality of the solution or search speed. Whereas the paper's authors note that in their system, "actions are decided one by one" which may satisfy the Markov property, their analysis shows that the system needs user help in problems larger than 20 objects which corresponds approximately to the performance decay reported for exact approaches. As points like this lack the learning component critical to RL, we consider this paper a dynamic programming approach rather than an RL one and thus code it as such.

B. DISCUSSION OF FINDINGS

Fig. 10 below summarizes the results of our analysis in absolute numbers per learning paradigm[1]. We found a total of nine publications belonging to Supervised Learning, none of which were used directly as the FLP resolution method (IC1). One explanation that stands to reason is that the lack of labeled data shift strongly drives their use as auxiliary functions or preparatory planning tasks. As Chung [70] argues, Facility Layout Problems are always ill-structured and their information is noisy, uncertain, or incomplete, making it hard to obtain a proper supervisor. The well-known drawback of

[1]Note that the numbers only add up to 23, as we added two papers for synthesis which did not use ML concepts.

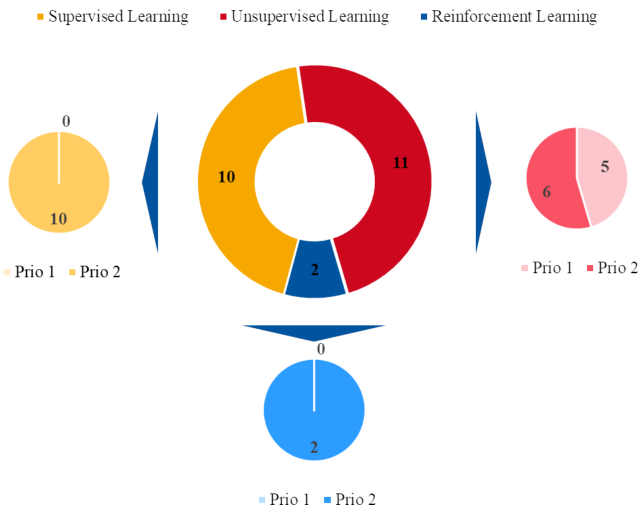

FIGURE 10. Distribution of occurrence of learning paradigms in final dataset.

FLPs is that their NP-hard nature prohibits optimal solutions at problem sizes larger than around 15. Layouts labeled through expert judgment are of a rather subjective quality and disqualify as suitable training data with respect to layout *optimization*. The quality of approximating approaches is per their definition sub-optimal, albeit workable. Using layouts generated by GA, ACO, construction algorithms or the like instead as training data might allow a Neural Network to generalize over a variety of feasible solutions. With Machine Learning methods being statistical tools [77], it is debatable whether these abstractions could produce layouts with higher quality on their own rather than replicating approximated solutions. It remains to be seen whether SL based resolution techniques will find their way into FLP research.

We further identified 11 contributions from Unsupervised Learning, roughly half of which were used directly to solve FLPs. All but one relevant publications employ Kohonen or Hopfield Neural Networks. The underlying motivation for this appears to be a low computational load compared to meta-heuristics, especially for larger-sized problems, as well as faster and less labor-intensive implementation than the traditional (i.e. manual) Systematic Layout Planning method. The remaining papers deal with a major research stream named group technology, which can be considered an upstream planning process in production planning which is not *primarily* concerned with layout design. Nonetheless, approaches using generative adversarial networks (GAN) in architectural research [78] or variational autoencoders in (surface) layout designs in material science [79] demonstrate ongoing research potential and might deserve further attention in FLP research.

Lastly, we found a low number of two papers using Reinforcement Learning approaches, yet only one pertinent to IC1. While we could identify evidence of reward or penalty formulations and actions space definitions, we did not find any cue for the use of Markov Decision Processes or the discussion of problem formulations satisfying the Markov property. Instead, problems were solved via exact approaches or multi-agent systems where the reward earning mechanism in Q-Learning merely played a minor role and was not further discussed. The one publication that most resembled an RL application did not have layout design as its focus (thus violating IC1) yet mentioned this application as future research scope. A forward search starting from this paper does indicate that this has been fulfilled to date. The pervasion of RL techniques, such as Q-Learning, in operational research is not overtly widespread, as another recent review investigated in which these techniques account for 1.5% of the papers [80]. Apart from this, we observed that some of the methods proposed for FLPs originate from other practical problems. Yeh [57] notes that SA has been effective in solving the traveling salesman problem (TSP), with respect to solution quality rather than to computing times. Similarly, [60] states that Hopfield networks had previously been used as TSP solutions. Given that Reinforcement Learning applications can be found in job-shop scheduling problem literature [81]–[84], AGV routing [85] and have further been proposed as a future research topic on material handling system location planning [86], adopting RL as resolution technique for FLP, based on re-formulating classic layout formulations as Markov Decision Process, certainly stands to reason.

Fig. 11 shows the temporal distribution of the papers mentioned above. One can observe that there was no research activity on the topic of ML in FLP before 1995, followed by a surge of activity until the year 2004. Contributions after this point in time are marginal. We attribute this to the end of the AI winter in 1997 [77], which may have led to fresh funding for ML research for different practical domains. Pursuing this theory by analyzing whether this pattern is mimicked by other disciplines goes beyond the scope of our review. Hence, we leave this observation to further speculation.

Interestingly, 86% (19 out of 22) of the papers examined make use of ANN, for both supervised and unsupervised approaches. As these papers do not discuss their decision regarding other algorithms from Supervised Learning as shown in Fig. 4 we cannot credibly conclude on reasons. However, taking into account the research interest in ANN, exemplarily displayed for the database Sciencedirect in Fig. 12, we can observe that the peaks we see for the years 1995 and 2014 in Fig. 11, correlate with high relative changes in general research interest. Fig. 12, on the other side, does not explain the peak in 1999 and merely provides a slight hint for the one in 2014. Yet, this leads us to the assumption that model selection was primarily influenced by zeitgeist rather than consideration of concurrent algorithmic alternatives.

In summary, we find that there is a distinct number of publications reporting on using ML techniques, mainly from the field of Unsupervised Learning, where said techniques are not directly employed to generate layouts, i.e. as resolution methods. These publications from the field of "group technology" aim at assigning machines to a certain number of "cells" with similar machines to minimize material flow crossings

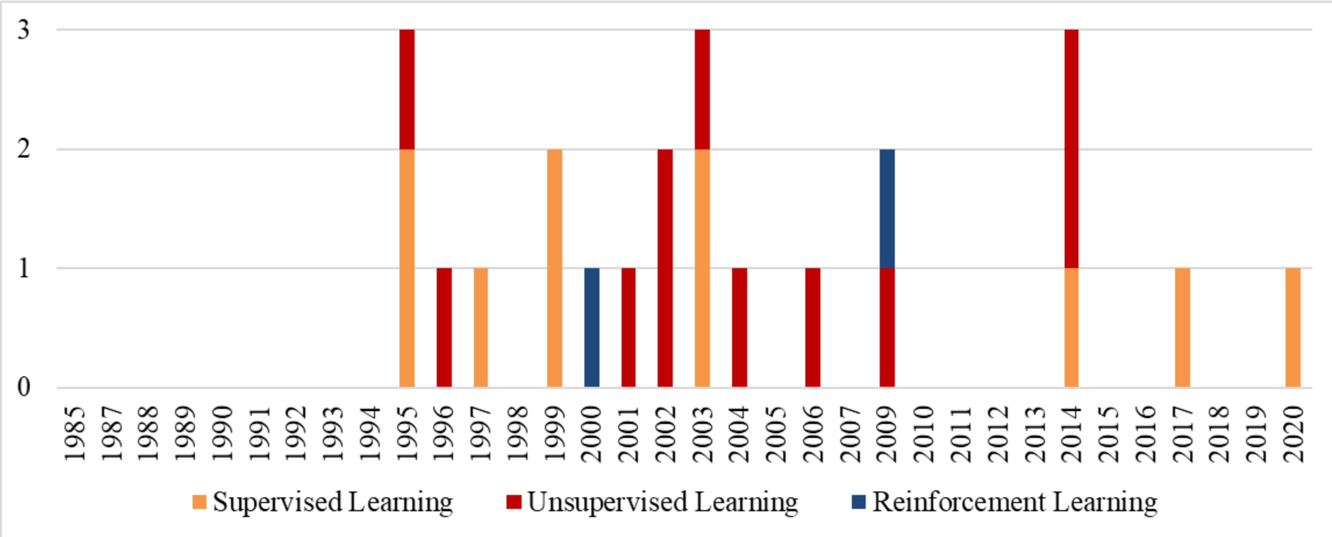

FIGURE 11. **Time series analysis of ML-related publications in FLP between 1985 and 2020.**

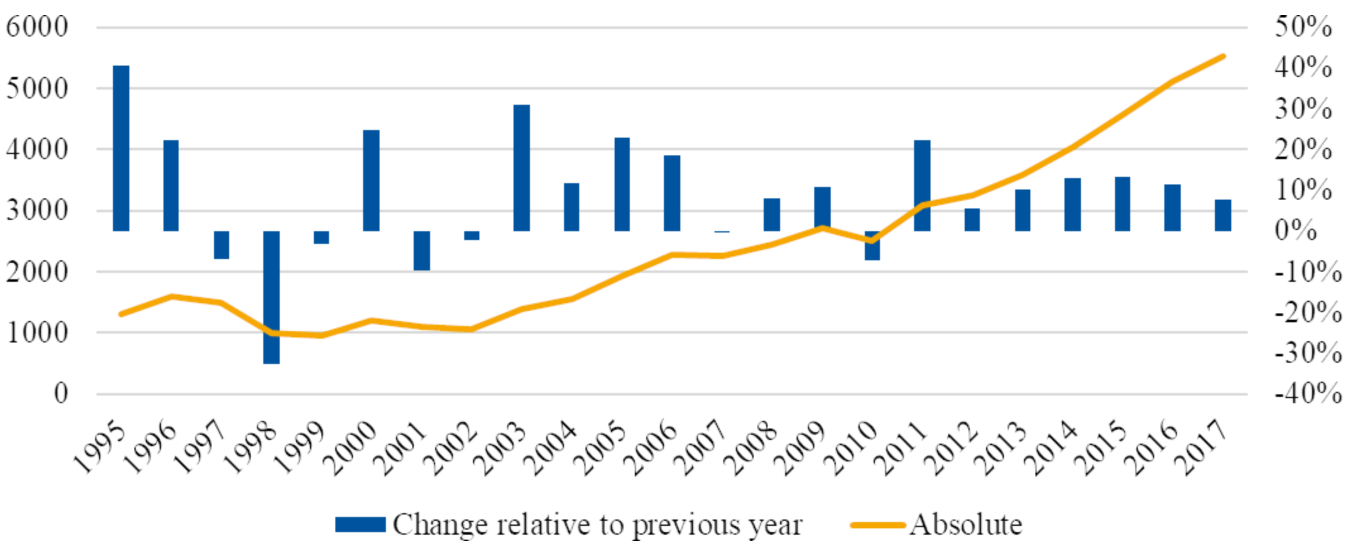

FIGURE 12. **Time series analysis of ANN-related papers on Sciencedirect between 1995 and 2017.**

between cells. As such, these publications/techniques can be said to generate the internal contents of blocks in block layouts, but they do not address the FLP in a way that determines the position of cells in a plant. This finding is in line with the analysis performed by Renzi *et al.* [15] who indicated that AI techniques are primarily harnessed for grouping/cell formation and scheduling issues and not for layout design.

The results presented in the previous section have highlighted that intelligent approaches using any of the three Machine Learning paradigms Supervised, Unsupervised, and Reinforcement Learning plus the cross-sectional paradigm Deep Learning are under-represented in FLP literature relative to other approaches. Within our final sample used for our synthesis, we found a relatively even spread between Supervised and Unsupervised Learning techniques. However, their importance for FLP problems varies as we identified no paper where Supervised Learning is used for FLPs while we found five for Unsupervised Learning. Reinforcement Learning is even less represented, with a total of two papers, none of which satisfy both of our inclusion criteria.

C. METHODOLOGY

1) SYSTEMATIC REVIEW

We have conducted a review with a *sensitive* search strategy meaning that the goal is to identify as much relevant material as possible [87]. We deliberately chose not to limit the search strategy to avoid the risk of missing relevant studies. Instead, we employed a clustering algorithm to aid us in coding papers. The results and implications of this are discussed hereinafter.

Our search yielded a raw dataset containing a total of 11,851 papers. Considering that we had 1,290 potentially relevant papers after all pre-processing and clean-up steps, we can ascertain our *precision* to be 0,11. In other words, before even entering the screening stages, we have a dropout rate of 89%, resulting in a significant upfront effort and thus a costly analysis. Concerning the selection of papers for the narrative synthesis, our review has a disproportionally low precision of 0.002 (24 in 11,851 papers). Our analysis regarding papers of which databases were still contained in the final study sample (see Fig. 13 - OpenReview is excluded as it provided 0 results) shows that some of the databases we used suffer from a dropout rate next to 100%. This finding indicates that reducing the number of databases used is not necessarily with detriment to study quality, but might increase precision and decrease review cost. As this result is always specific for the search and review study in question, our analysis does not warrant a generalization of good practice. We nonetheless propose to invest a substantial amount of effort into the scoping search (Step 1 in the review framework used herein) to fine-tune the selection of databases.

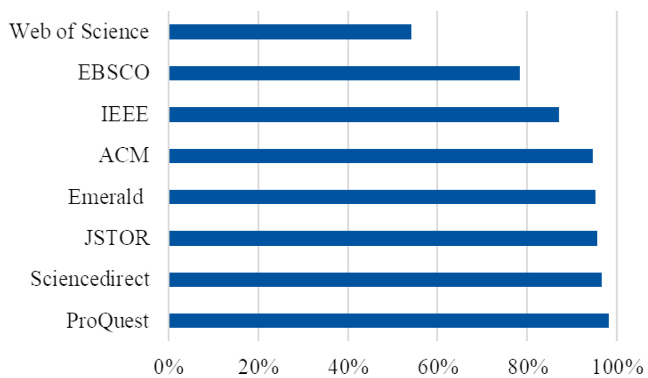

FIGURE 13. Percentages of publications dropped per database along the review process before reaching screening stages.

Our second proposition focuses on reducing losses in screening processes. We employed two distinct methods to support the analysis process: 1) a search algorithm to re-create title-abstract-keyword searches to filter out irrelevant contributions (these with facility layout problems merely being a side-notion) and b) a review team to increase inter-rater reliability. While any search automation is indispensable to sift through tens of thousands of references, special care must be exercised as to the results. We double-checked all exclusions suggested by the search algorithm and occasionally found contributions we considered relevant and thus re-included them manually. As a result of this multi-level filtering, there was still a body of 1,290 papers left to manually screen. We thus further recommend following the good practice as given by e.g. [41] and to assemble a review team to reduce screening time and avoid bias in the analysis.

2) CLUSTERING ALGORITHM

At the level of Stage 1 Screening, the clustering algorithm identified eight papers relevant to the research question while the manual screening found 49, yielding an accuracy of 14%. This number suffers from the fact that some papers used a different AI framework which considers meta-heuristics as an AI technique. Including these false positives increases the accuracy to 16%. Regarding the complete dataset, the algorithm correctly classified 185 out of 1,290 papers, likewise yielding an accuracy of 14%. Considering only the subset of papers relevant to the clusters, i.e. all papers manually assigned to GA, ACO, PSO, TS, SA, or AI, the classification rate increases to 37%. Based on these results, we conclude that using the clustering algorithm as a stand-alone technique is insufficient with respect to accuracy. However, this review can serve as an indicator for potential time-savings in lengthy review processes of 10-20%, depending on the level of trust put into the algorithm and thus the amount of rework required.

VI. CONCLUSION

Our review set out to analyze the extent to which Machine Learning techniques, herein defined as a subset of Artificial Intelligence, have been used in scientific literature as resolution approach for Facility Layout Problems.

Through a systematic literature review, we collected a set of 1,290 contributions in a sample of 11,851 papers retrieved from nine different databases plus another 134 contributions from snowball sampling. Analyzing these contributions with respect to two inclusion criteria, i.e. the relevance to the field of FLP (IC1) and reports of using an ML technique as a resolution approach to solving FLP (IC2), yielded only 23 contributions that had a distinct AI focus. Two more papers were included as they showed traits pertinent to ML techniques.

As the essence of our in-depth full-text analysis of 25 papers, we can answer our research question, *How have different Machine Learning algorithms been used as resolution techniques for Facility Layout Problems?*, as follows: while there is some evidence for all three ML learning paradigms (Supervised, Unsupervised and Reinforcement Learning), their focus on FLP resolution approaches is scarce.

None of the papers found using SL were relevant to FLP resolution approaches. Instead, SL was primarily used as an adjunct optimizer for other resolution approaches or employed in planning tasks preceding that of layout design. We infer that the under-representation of SL could be driven by a general lack of structured, labeled data.

For UL, we identified the most prominent algorithms to be SOM and Hopfield Networks, respectively. Their use is motivated by a lower computational load for large problems than is required for similar meta-heuristics. However, we also detect a downswing in research interest in UL since a peak in the early 2000s.

Additionally, we analyzed that despite a wide variety of available ML algorithms, ANN were strikingly prominent.

As a consequence, following SOM, backpropagation networks rank second in terms of occurrence. As the papers visited do not provide unambiguous evidence for the selection of ANN over other alternatives, using a time series analysis of research interest in ANN, we infer that the choice of ANN was likely to be a result of the spirit of the time rather than the inaptness of other algorithms.

Lastly, we found only two contributions in RL yet none of these satisfied IC1 and IC2. This finding indicates that RL has not yet found its way into FLP as a resolution approach. This is remarkable, as a distinct number of studies make use of simulations, requiring defined environments which could also be used to model MDPs.

In summary, this analysis contributes to the body of knowledge on the topic of facility layout problems by refining previous taxonomies. Specifically, to date, there existed no classification that decomposed intelligent approaches into the branches symbolic and sub-symbolic as well as the Machine Learning paradigms and associated algorithms. These updates allow future researchers to classify their work in more detail and encourage comparative studies on the effectiveness and efficiency of alternative algorithms on certain problem sets. We further contribute by highlighting white spaces, such as the non-existence of several algorithms in our classification in FLP to date. Thus, as future research work, we encourage the investigation of both SL and RL approaches. For SL, deep learning-based solutions may prove to be valid topics. Deep Learning models trained on known layout alternatives may be able to predict or rate new layouts simply by visual input, such as raw images, by being able to identify certain objects in the layout. By leveraging computer vision techniques, this could become an opportunity for on-the-fly layout evaluation in group decision-making settings. The idea to use RL as a resolution approach in future FLP research draws from insights that are being generated in other complex problem domains in production research. Research work in the area could include formal MDP notations as a foundation for using RL, insights on the modeling prerequisites to make RL algorithms trainable, or studies on the usability of various available algorithms (e.g. when will tabular Q-Learning approaches reach their limits? Are there differences in convergence or training speeds for different algorithms, for instance, those using discrete action spaces compared to those with continuous actions spaces).

REFERENCES

[1] S. S. Heragu and S. R. Kakuturi, "Grouping and placement of machine cells," *IIE Trans.*, vol. 29, no. 7, pp. 561–571, Jul. 1997.

[2] I. Maganha and C. Silva, "A theoretical background for the reconfigurable layout problem," *Procedia Manuf.*, vol. 11, pp. 2025–2033, Jan. 2017.

[3] J. Tompkins, J. A. White, and Y. A. Bozer, *Facilities Planning*, 4th ed. Hoboken, NJ, USA: Wiley, 2010.

[4] S. Benjaafar, "Modeling and analysis of congestion in the design of facility layouts," *Manage. Sci.*, vol. 48, no. 5, pp. 679–704, May 2002.

[5] D. Raman, S. V. Nagalingam, and B. W. Gurd, "A genetic algorithm and queuing theory based methodology for facilities layout problem," *Int. J. Prod. Res.*, vol. 47, no. 20, pp. 5611–5635, Oct. 2009.

[6] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.

[7] O. Vinyals, I. Babuschkin, J. Chung, M. Mathieu, M. Jaderberg, W. M. Czarnecki, and T. Ewalds. (2019). *Alphastar: Mastering the Real-Time Strategy Game Starcraft II*. [Online]. Available: https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii

[8] N. Brown and T. Sandholm, "Superhuman AI for heads-up no-limit poker: Libratus beats top professionals," *Science*, vol. 359, no. 6374, pp. 418–424, Jan. 2018.

[9] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, and I. Mordatch, "Emergent tool use from multi-agent autocurricula," 2019, *arXiv:1909.07528*. [Online]. Available: http://arxiv.org/abs/1909.07528

[10] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[11] P. Burggraf, J. Wagner, and B. Koke, "Artificial intelligence in production management: A review of the current state of affairs and research trends in academia," in *Proc. Int. Conf. Inf. Manage. Process. (ICIMP)*, Jan. 2018, pp. 82–88.

[12] H. Hosseini-Nasab, S. Fereidouni, S. M. T. Fatemi Ghomi, and M. B. Fakhrzad, "Classification of facility layout problems: A review study," *Int. J. Adv. Manuf. Technol.*, vol. 94, nos. 1–4, pp. 957–977, Jan. 2018.

[13] G. Moslemipour, T. S. Lee, and D. Rilling, "A review of intelligent approaches for designing dynamic and robust layouts in flexible manufacturing systems," *Int. J. Adv. Manuf. Technol.*, vol. 60, nos. 1–4, pp. 11–27, Apr. 2012.

[14] S. P. Leo Kumar, "State of the art-intense review on artificial intelligence systems application in process planning and manufacturing," *Eng. Appl. Artif. Intell.*, vol. 65, pp. 294–329, Oct. 2017.

[15] C. Renzi, F. Leali, M. Cavazzuti, and A. O. Andrisano, "A review on artificial intelligence applications to the optimal design of dedicated and reconfigurable manufacturing systems," *Int. J. Adv. Manuf. Technol.*, vol. 72, nos. 1–4, pp. 403–418, Apr. 2014.

[16] A. Drira, H. Pierreval, and S. Hajri-Gabouj, "Facility layout problems: A survey," *Annu. Rev. Control*, vol. 31, no. 2, pp. 255–267, Jan. 2007.

[17] A. Kusiak and S. S. Heragu, "The facility layout problem," *Eur. J. Oper. Res.*, vol. 29, no. 3, pp. 229–251, 1987.

[18] L. Garcia-Hernandez, L. Salas-Morera, J. A. Garcia-Hernandez, S. Salcedo-Sanz, and J. Valente de Oliveira, "Applying the coral reefs optimization algorithm for solving unequal area facility layout problems," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112819.

[19] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," *J. Phys., Conf. Ser.*, vol. 1142, no. 1, 2018, Art. no. 012012.

[20] L. Monostori, J. Hornyák, C. Egresits, and Z. J. Viharos, "Soft computing and hybrid AI approaches to intelligent manufacturing," in *Tasks and Methods in Applied Artificial Intelligence* (Lecture Notes in Computer Science), vol. 1416, A. P. del Pobil, J. Mira, and M. Ali, Eds. Berlin, Germany: Springer, 1998, pp. 765–774.

[21] P. Gentsch, *Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business—Konzepte und Best Practices*, 2nd ed. Wiesbaden, Germany: Springer, 2019.

[22] A. Turing, *Intelligent Machinery*. Berlin, Germany: Springer-Verlag, 1948.

[23] A. Newell, "Physical symbol systems," *Cognit. Sci.*, vol. 4, no. 2, pp. 135–183, 1980.

[24] R. Miikkulainen, "Integrated connectionist models: Building AI systems on subsymbolic foundations," in *Proc. 6th Int. Conf. Tools Artif. Intell.*, Los Alamitos, CA, USA, Nov. 1994, pp. 231–232.

[25] K. A. H. Kobbacy, S. Vadera, and M. H. Rasmy, "AI and OR in management of operations: History and trends," *J. Oper. Res. Soc.*, vol. 58, no. 1, pp. 10–28, Jan. 2007.

[26] K. A. H. Kobbacy and S. Vadera, "A survey of AI in operations management from 2005 to 2009," *J. Manuf. Technol. Manage.*, vol. 22, no. 6, pp. 706–733, Jul. 2011.

[27] R. Bellman, *Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 1957.

[28] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM J. Res. Develop.*, vol. 3, no. 3, pp. 210–229, Jul. 1959.
[29] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, Dec. 1943.
[30] Q. Hardy. (2016). *Reasons to Believe the AI Boom is Real*. [Online]. Available: https://Www.nytimes.com/2016/07/19/technology/reasons-to-believe-the-ai-boom-is-real.html
[31] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Piscataway, NJ, USA, May 2013, pp. 8595–8598.
[32] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben, "Machine learning in manufacturing: Advantages, challenges, and applications," *Prod. Manuf. Res.*, vol. 4, no. 1, pp. 23–45, Jan. 2016.
[33] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Found. Trends Mach. Learn.*, vol. 11, nos. 3–4, pp. 219–354, 2018.
[34] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proc. 23rd Int. Conf. Mach. Learn.*, W. Cohen, Ed. New York, NY, USA, 2006, pp. 161–168.
[35] G. Hinton, T. J. Sejnowski, and T. A. Poggio, *Unsupervised Learning: Foundations of Neural Computation* (Computational Neuroscience). Cambridge, MA, USA: MIT Press, 1999.
[36] C. Sammut and G. I. Webb, *Encyclopedia of Machine Learning* (Springer Reference). New York, NY, USA: Springer, 2011.
[37] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
[38] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014.
[39] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[40] S. Leijnen and F. V. Veen, "The neural network zoo," *Proceedings*, vol. 47, no. 1, p. 9, May 2020.
[41] D. Gough, S. Oliver, and J. Thomas, *An Introduction to Systematic Reviews*. Los Angeles, LA, USA: Sage, 2012.
[42] A. Boland, G. Cherry, and R. Dickson, *Doing a Systematic Review: A Student's Guide*. Newbury Park, CA, USA: Sage, 2017.
[43] C. Chin-Sheng and K. Kengskool, "An autocad-based expert system for plant layout," *Comput. Ind. Eng.*, vol. 19, nos. 1–4, p. 299, 1990.
[44] P. Potocnik, T. Berlec, A. Sluga, and E. Govekar, "Hybrid self-organization based facility layout planning," *Strojniski Vestnik-J. Mech. Eng.*, vol. 60, no. 12, pp. 789–796, 2014.
[45] A. W. A. Hammad, A. Akbarnezhad, and D. Rey, "A multi-objective mixed integer nonlinear programming model for construction site layout planning to minimise noise pollution and transport costs," *Autom. Construct.*, vol. 61, pp. 73–85, Jan. 2016.
[46] D. Lobos and D. Donath, "The problem of space layout in architecture: A survey and reflections," *Arquitetura Revista*, vol. 6, no. 2, pp. 136–161, Dec. 2010.
[47] J. Martens, "Two genetic algorithms to solve a layout problem in the fashion industry," *Eur. J. Oper. Res.*, vol. 154, no. 1, p. 304, 2004.
[48] B. Koke and R. C. Moehler, "Earned green value management for project management: A systematic review," *J. Cleaner Prod.*, vol. 230, pp. 180–197, Sep. 2019.
[49] P. Burggräf, J. Wagner, and T. Weißer, "Knowledge-based problem solving in physical product development—A methodological review," *Expert Syst. Appl., X*, vol. 5, pp. 1–14, Apr. 2020.
[50] T. Greenhalgh and R. Peacock, "Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources," *BMJ*, vol. 331, no. 7524, pp. 1064–1065, Nov. 2005.
[51] T. Weißer, T. Saßmannshausen, D. Ohrndorf, P. Burggräf, and J. Wagner, "A clustering approach for topic filtering within systematic literature reviews," *MethodsX*, vol. 7, pp. 1–10, Jan. 2020.
[52] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The prisma statement," *Ann. Internal Med.*, vol. 151, no. 4, pp. 9–264, W64, 2009.
[53] M. Kobayashi, T. Makita, S. Matsui, M. Koyama, N. Fujiii, I. Hatono, and K. Ueda, "Floor layout planning method based on self-organization," in *Proc. IEEE Int. Symp. Semiconductor Manuf. (ISSM)*, Oct. 2001, pp. 381–384.
[54] K. Ueda, N. Fujii, I. Hatono, and M. Kobayashi, "Facility layout planning using self-organization method," *CIRP Ann.*, vol. 51, no. 1, pp. 399–402, 2002.
[55] K. Tsuchiya, S. Bharitkar, and Y. Takefuji, "A neural network approach to facility layout problems," *Eur. J. Oper. Res.*, vol. 89, no. 3, pp. 556–563, Mar. 1996.
[56] G. W. Zhang, S. C. Zhang, and Y. S. Xu, "Research on flexible transfer line schematic design using hierarchical process planning," *J. Mater. Process. Technol.*, vol. 129, nos. 1–3, pp. 629–633, Oct. 2002.
[57] I.-C. Yeh, "Architectural layout optimization using annealed neural network," *Autom. Construct.*, vol. 15, no. 4, pp. 531–539, Jul. 2006.
[58] A. J. Vakharia and U. Wemmerlöv, "A comparative investigation of hierarchical clustering techniques and dissimilarity measures applied to the cell formation problem," *J. Oper. Manage.*, vol. 13, no. 2, pp. 117–138, Aug. 1995.
[59] T. Berlec, P. Potočnik, E. Govekar, and M. Starbek, "A method of production fine layout planning based on self-organising neural network clustering," *Int. J. Prod. Res.*, vol. 52, no. 24, pp. 7209–7222, Dec. 2014.
[60] B. Ateme-Nguema and T.-M. Dao, "Quantized hopfield networks and tabu search for manufacturing cell formation problems," *Int. J. Prod. Econ.*, vol. 121, no. 1, pp. 88–98, Sep. 2009.
[61] J. Inho and J. Rhee, "Generalized machine cell formation considering material flow and plant layout using modified self-organizing feature maps," *Comput. Ind. Eng.*, vol. 33, nos. 3–4, pp. 457–460, Dec. 1997.
[62] T. Tateyama, S. Kawata, and H. Ohta, "Self-organizing map for group technology oriented plant layout planning," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E86A, no. 11, pp. 2747–2754, 2003.
[63] T. Tateyama, S. Kawata, and H. Ohta, "A conditional clustering algorithm using self-organizing map," in *Proc. SICE Annu. Conf.*, vol. 3, 2003, pp. 3259–3264.
[64] T. Tateyama and S. Kawata, "Plant layout planning using art neural networks," in *Proc. SICE Annu. Conf.*, vol. 2, 2004, pp. 1863–1868.
[65] M. Sahragard and M. Bashiri, "Layout design for large scale problems with a hybridized clustering based heuristic method," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2016, pp. 1610–1614.
[66] L. García-Hernández, M. Pérez-Ortiz, A. Araúzo-Azofra, L. Salas-Morera, and C. Hervás-Martínez, "An evolutionary neural system for incorporating expert knowledge into the UA-FLP," *Neurocomputing*, vol. 135, pp. 69–78, Jul. 2014.
[67] C. M. Tam and T. K. L. Tong, "GA-ANN model for optimizing the locations of tower crane and supply points for high-rise public housing construction," *Construct. Manage. Econ.*, vol. 21, no. 3, pp. 257–266, Apr. 2003.
[68] P. Azimi and P. Soofi, "An ANN-based optimization model for facility layout problem using simulation technique," *Scientia Iranica*, vol. 24, no. 1, pp. 364–377, Feb. 2017.
[69] A. Tayal, U. Kose, A. Solanki, A. Nayyar, and J. A. M. Saucedo, "Efficiency analysis for stochastic dynamic facility layout problem using meta-heuristic, data envelopment analysis and machine learning," *Comput. Intell.*, vol. 36, no. 1, pp. 172–202, Feb. 2020.
[70] Y. K. Chung, "A neuro-based expert system for facility layout construction," *J. Intell. Manuf.*, vol. 10, no. 5, pp. 359–385, 1999.
[71] Y.-K. Chung, "Application of a cascade BAM neural expert system to conceptual design for facility layout," *Comput. Math. Appl.*, vol. 37, no. 1, pp. 95–110, Jan. 1999.
[72] F. F. Chen and S. R. Sagi, "Concurrent design of manufacturing cell and control functions: A neural network approach," *Int. J. Adv. Manuf. Technol.*, vol. 10, no. 2, pp. 118–130, Mar. 1995.
[73] H. A. Rao and P. Gu, "A multi-constraint neural network for the pragmatic design of cellular manufacturing systems," *Int. J. Prod. Res.*, vol. 33, no. 4, p. 1048, 1995.
[74] H. Tarkesh, A. Atighehchian, and A. S. Nookabadi, "Facility layout design using virtual multi-agent system," *J. Intell. Manuf.*, vol. 20, no. 4, p. 347, 2009.
[75] S. Kondoh, Y. Umeda, T. Tomiyama, and H. Yoshikawa, "Self organization of cellular manufacturing systems," *CIRP Ann.*, vol. 49, no. 1, pp. 347–350, 2000.
[76] S. Ono, Y. Hamada, S. Yamamoto, K. Mizuno, and S. Nishihara, "Rearrangement of floor layouts based on constraint satisfaction," in *Proc. IEEE Int. Conf. Syst., Man Cybern. e-Syst. e-Man Cybern. Cyberspace*, vol. 4, Oct. 2001, pp. 2759–2764.
[77] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Harlow, U.K.: Pearson, 2016.
[78] W. Huang and H. Zheng, "Architectural drawings recognition and generation through machine learning," in *Proc. 38th Annu. Conf. Assoc. Comput. Aided Design Archit.*, 2018, pp. 156–165.

[79] Y. Zhang and W. Ye, "Deep learning–based inverse method for layout design," *Struct. Multidisciplinary Optim.*, vol. 60, no. 2, pp. 527–536, 2019.
[80] T. Weitzel and C. H. Glock, "Energy management for stationary electric energy storage systems: A systematic literature review," *Eur. J. Oper. Res.*, vol. 264, no. 2, pp. 582–606, Jan. 2018.
[81] L. M. Gambardella and M. Dorigo, "Ant-q: A reinforcement learning approach to the traveling salesman problem," in *Machine Learning Proceedings 1995*. Amsterdam, The Netherlands: Elsevier, 1995, pp. 252–260.
[82] T. Gabel and M. Riedmiller, "Distributed policy search reinforcement learning for job-shop scheduling tasks," *Int. J. Prod. Res.*, vol. 50, no. 1, pp. 41–61, Jan. 2012.
[83] N. Stricker, A. Kuhnle, R. Sturm, and S. Friess, "Reinforcement learning for adaptive order dispatching in the semiconductor industry," *CIRP Ann.*, vol. 67, no. 1, pp. 511–514, 2018.
[84] P. Burggräf, J. Wagner, B. Koke, and M. Bamberg, "Performance assessment methodology for ai-supported decision-making in production management," *Procedia CIRP*, vol. 93C, pp. 884–889, Jan. 2020.
[85] A. Malus, D. Kozjek, and R. Vrabič, "Real-time order dispatching for a fleet of autonomous mobile robots using multi-agent reinforcement learning," *CIRP Ann.*, vol. 69, no. 1, pp. 397–400, 2020.
[86] S. Lee, H.-G. Kahng, T. Cheong, and S. B. Kim, "Iterative two-stage hybrid algorithm for the vehicle lifter location problem in semiconductor manufacturing," *J. Manuf. Syst.*, vol. 51, pp. 106–119, Apr. 2019.
[87] A. Harden, R. Weston, and A. Oakley, "A review of the effectiveness and appropriateness of peer-delivered health promotion interventions for young people," in *Database of Abstracts of Reviews of Effects (DARE): Quality-Assessed Reviews [Internet]*, A. Harden, R. Weston, and A. Oakley, Eds. U.K.: Centre for Reviews and Dissemination, 1999. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/nbk67890/

PETER BURGGRÄF was born in 1980. He received the Dr.Ing. and M.B.A. degrees. He studied mechanical engineering in Aachen and London. His wrote the Ph.D. thesis at the Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, in the field of factory planning. From 2011 to 2017, he was the Chief Engineer of the Chair of Production Engineering, WZL. Since 2013, he has been the Managing Director of StreetScooter Research GmbH. Since 2017, he has been holding the Chair of International Production Engineering and Management (IPEM), University of Siegen, Germany.

JOHANNES WAGNER was born in 1986. He received the Dr.Ing. and M.B.A. degrees from RWTH Aachen University and the M.Sc. degrees in industrial engineering and Tsinghua University, Beijing. He wrote the Ph.D. thesis at the Laboratory for Machine Tools and Production Engineering (WZL), Chair of Production Engineering, RWTH Aachen University, in the field of disruption management in low-volume assembly. Since 2017, he has been the Chief Engineer of the Chair of International Production Engineering and Management (IPEM), University of Siegen. He is also the Founder and the CEO of the Smart Demonstration Factory Siegen (SDFS).

BENJAMIN HEINBACH was born in 1988. He received the M.Sc. degree from the Chalmers University of Technology and the M.Sc. degree from Northumbria University. He is currently pursuing the Ph.D. degree with the Chair of International Production Engineering and Management (IPEM), University of Siegen. He is writing a Ph.D. thesis in the field of automated factory planning based on artificial intelligence. Since 2017, he has been working at the Chair of IPEM, University of Siegen. He has been managing the group factory planning for medium-sized businesses at the aforementioned chair since 2019.