

一般論文

AI技術を活用した図面検索ツールの構築

Construction of Drawings Retrieval Tool Using Artificial Intelligence Technology

森本 充* 竹原 輝 巳*
M. Morimoto T. Takehara

概要

近年、ディープラーニング（深層学習）や強化学習をはじめとした人工知能（AI）技術の発展は目覚ましく、産業分野においても幅広く活用されている。当社においても、AI技術を新製品や社内業務への適用を積極的に進めている。本稿では、社内業務への適用例として、設計業務の工数削減を目的に開発し、運用を開始した“図面検索ツール”について紹介する。

Synopsis

In recent years, artificial intelligence (AI) technology such as deep learning and reinforcement learning has developed remarkably, which is widely used in the industrial field. We positively take on application of AI technology to new products and internal works. In this article, as an example of the application to internal works, we will introduce "drawings retrieval tool" which was developed and is currently used in order to reduce man-hour in design work.

1. はじめに⁽¹⁾

当社製品は、顧客のニーズに丁寧に対応した個別仕様が多い。そのため、標準図面で対応できない場合、流用設計を行っており、参照可能な過去の製作図面を探し出す作業が重要となる。ただし、新規の受注製品の仕様と完全に一致する場合は稀で、通常は部分一致もしくは類似する過去の仕様図面を、設計者の経験やノウハウに基づいて探し出しており、工数を要している。また、このような状況で適切な過去の仕様図面を探し出すために、キーワード検索やパターンマッチングのような一般的な検索による自動化方法はあるが、仕様を絞り込むためにノウハウが必要となってくる。

そこで、仕様が一定レベル以上類似すれば検索が可能となるAI技術を活用した“図面検索ツール”の開発に取り組んだ。

ここで、今回開発した図面検索ツールの導入前後における図面検索作業の比較を図1に示す。図1（導入前）において、経験やノウハウを活かすことができる熟練設計者は、時間をかければ、新規の受注製品の仕様と部分一致もしくは類似する過去の仕様図面を探し出せる。しかし、経験やノウハウが不足している経験年数の少ない設計者にとっては、適切な過去の仕様図面を探し出すことが難しい。そのため、図1（導入後）のように、経験やノウハウを問わず、適切な過去

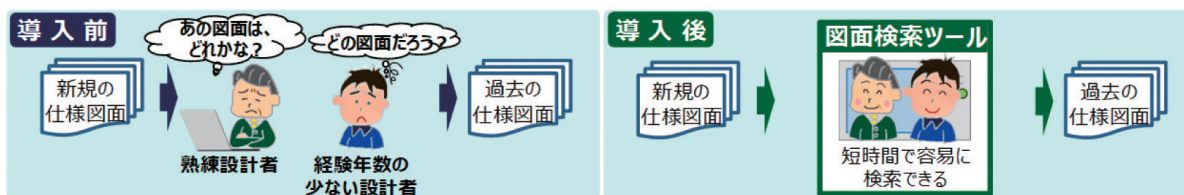


図1 図面検索ツールの導入前後における検索作業のイメージ

* 研究開発本部

の仕様図面を短時間かつ容易に検索できるようにし、探し出す作業の工数の削減を目指した。なお、実際の設計業務で活用するにあたり、図面検索精度の目標を70%以上に設定した。

2. 図面検索ツールの全体像

今回、図面検索ツールの構築にあたり、機械学習を活用した。機械学習とは、『手持ちのデータに対して反復的に“学習”を行うことで、未知のデータを“予測”できる技術』であり、主要なAI技術の一つである。

次に、図面検索ツールの技術的要点を図2に示す。

学習では、学習データ（過去の仕様図面）を用いて、過去の仕様とその仕様図面の関係性を機械学習のクラスタリングとロジスティック回帰を活用して、グループ毎に学習し、その結果として学習モデルを生成する。予測では、新規の受注製品の仕様を、その仕様と類似する学習モデルに入力し、過去の仕様図面を検索する。以下に学習および予測の手順を示す。

<学習>

- ①まず、クラスタリングにより学習データをグループ化する（学習データをN個のグループに分割する）。
- ②次に、グループ毎にロジスティック回帰を行い、学習モデルを生成する。

<予測>

- ③新規の受注製品の仕様と最も類似するグループを選択し、そのグループの学習モデルを選択する。
- ④選択した学習モデルを用いてロジスティック回帰を行い、新規の受注製品の仕様と類似する過去の仕様図面を検索する。

検索結果の出力イメージを表1に示す。検索結果には、新規の受注製品の仕様毎に、過去の仕様図面の番号を複数候補が表示される。あわせて、検索した過去の仕様図面の番号毎に正確度も表示することにより、設計者が検索結果を活用しやすい仕組みとした。

3. 図面検索ツールの開発

AI技術を活用するための開発手順を図3に示す。



図3 AI技術を活用するための開発手順

ここで、『データ準備・データ前処理』、『設計』、『評価』の各手順について説明する。

3.1 データ準備・データ前処理

図面検索ツールの学習データ（過去の仕様図面）は、イラストと文字情報で構成された図面である。今回、学習データとして、図面内の文字情報やそれらの出現回数を選定し、整理した。

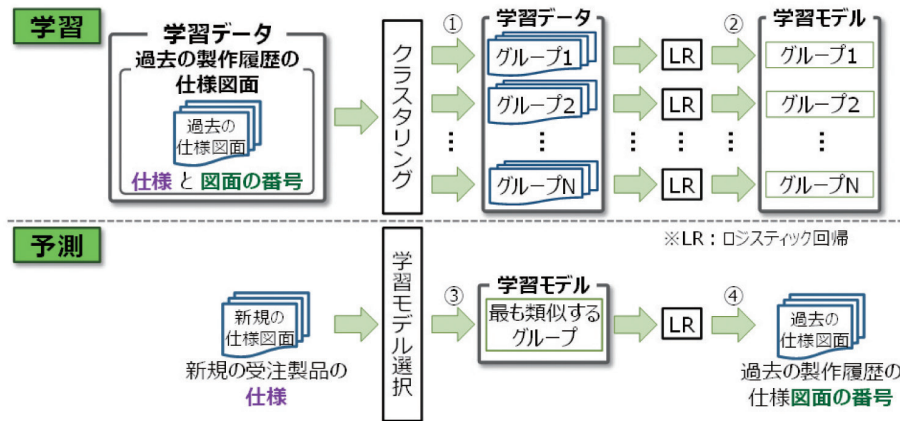


図2 図面検索ツールの技術的要点

表1 図面検索ツールの出力イメージ

	候補1		候補2		...	候補3	
	図面番号	正確度	図面番号	正確度		図面番号	正確度
新規の仕様図面1	過去の仕様図面A	95.3%	過去の仕様図面B	1.02%	...	過去の仕様図面F	0.102%
新規の仕様図面2	過去の仕様図面G	79.4%	過去の仕様図面H	13.2%	...	過去の仕様図面J	0.836%
新規の仕様図面3	過去の仕様図面K	86.9%	過去の仕様図面L	5.81%	...	過去の仕様図面O	0.215%
⋮			⋮				

3. 2 設計

図面検索ツールに活用した“クラスタリング”、“ロジスティック回帰”および“図面検索精度向上への取組み”を説明する。

3. 2. 1 クラスタリング

クラスタリングは、類似するデータをグループ化する技術であり、マーケティングの戦略立案や製品ポジショニングの分析などに活用されている。

過去の仕様図面は多種多様であるため、まず過去の仕様図面に対してクラスタリングを行い、お互いに類似する仕様をグループ化することにより、図面検索精度の向上を図った。しかし、過去の仕様図面は、明確にグループ化できるとは限らないため、複数のグループに帰属できる“fuzzy c-means”を採用した。

fuzzy c-meansは、各グループに帰属する確率（帰属度）をデータ毎に演算することにより、特定のグループに分割しにくいデータを複数のグループに帰属させることができる手法である。ここでは例として、特徴量（因子） j_1, j_2 を持つ複数のデータがfuzzy c-meansによって3グループに分割されるイメージを図4に示す。各グループの範囲が重なる部分に存在するデータは、複数のグループに帰属することを表している。

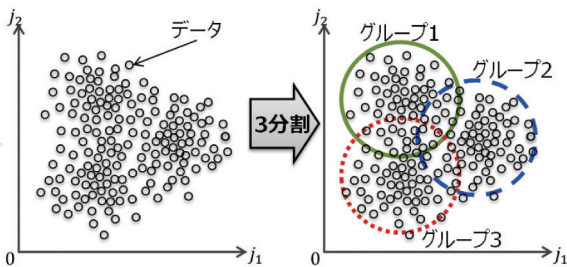


図4 fuzzy c-meansによるグループ化のイメージ

fuzzy c-meansでは、分割したグループの中心座標 c_k と帰属度 $\omega_k(x_i)$ を次のように表せる⁽²⁾。

$$c_k = \frac{\sum_{i=1}^n \{\omega_k(x_i)^m x_i\}}{\sum_{i=1}^n \omega_k(x_i)^m}$$

$$\omega_k(x_i) = \frac{1}{\sum_{p=1}^N \left\{ \frac{d(x_i, c_k)}{d(x_i, c_p)} \right\}^{\frac{2}{m-1}}}$$

c_k : k 番目のグループの中心座標

($1 \leq k \leq N$)

N : グループの個数

x_i : i 番目のクラスタリング対象のデータ

m : 曖昧度合い ($m > 1$)

※1に近づくとも曖昧さが小さくなる

$\omega_k(x_i)$: k 番目のグループに対する x_i の帰属度

j : 特徴量 (因子)

$d(x_i, c_k)$: x_i と c_k 間のユークリッド距離

$$d(x_i, c_k) = \sqrt{\sum_j (x_{ij} - c_{kj})^2}$$

3. 2. 2 ロジスティック回帰

ロジスティック回帰は、複数の特徴量（因子）の組合せから、ある事象の発生確率を予測する機械学習の手法であり、気象観測データを用いた土砂災害発生率の予測、患者の検査値を用いた発症率の予測などに活用されている。

今回、複数の特徴量（因子）の組合せから新規の受注製品の仕様に類似する過去の仕様図面を検索可能で、かつ、比較的少ない学習データでも予測できるため、図面検索ツールの主要エンジンとして採用した。

以下にロジスティック回帰の一般式⁽³⁾と、例として3項 ($N = 3$) ロジスティック回帰のイメージを図5に示す。

$$P_Y(X) = \frac{e^{\beta_Y X}}{1 + \sum_{k=1}^{N-1} e^{\beta_k X}} \quad (1 \leq Y \leq N-1)$$

$$P_Y(X) = \frac{1}{1 + \sum_{k=1}^{N-1} e^{\beta_k X}} \quad (Y = N)$$

N : 出力データ数

X : 入力データ (x_1, x_2, \dots, x_j)

Y : 出力データ ($1 \leq Y \leq N$)

$P(Y)$: 入力データから出力データが生じる確度

β_k : 入力データに掛かる係数 (重み)

($\beta_1, \beta_2, \dots, \beta_j$)

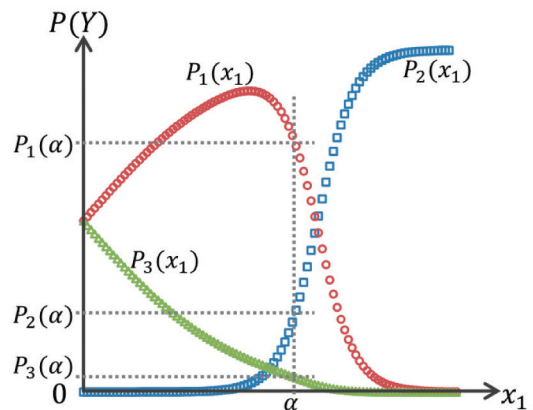


図5 3項 ($N = 3$) ロジスティック回帰のイメージ

図5に示すそれぞれの曲線は、入力データから出力データが発生する確率を表している。 x_1 軸にある値 a を3つの曲線に入力したとき、3つの出力データが発生する確率は、 $P_1(a)$, $P_2(a)$, $P_3(a)$ のように予測され、 $P_1(a)+P_2(a)+P_3(a)=100\%$ になる。 $P_1(a)$ が最も高い確率であることから、ある値 a を入力した場合、出力データ1が発生しやすいことが分かる。

3. 2. 3 図面検索精度向上への取組み

まず1点目の工夫として、クラスタリング (fuzzy c-means) のパラメータであるグループ数を、クラスタリングの結果を定量評価する指標である“PseudoF”を用いて調整した。PseudoFは、分母にグループ内のデータの分散値、分子にグループの中心の分散値で構成されており、『グループ内のデータが密集する』または『グループの中心がばらつく』ほど、大きい値をとる。つまり、PseudoFが大きい時、特徴量 (因子) の近いデータは同一グループに、または、特徴量 (因子) が離れているデータは別グループに帰属していることを指す。この評価指標を利用して、最適なグループ数を設定した。

次に2点目の工夫として、ロジスティック回帰の前段階で主成分分析を行った。主成分分析は、データの次元圧縮 (データ中の多くの情報を表現できる新たな変数を作成すること) を行う手法である⁽⁴⁾。学習データ (過去の仕様図面) の特徴量 (因子) に対して主成分分析を行い、少ない変数で再現することにより、比較的少ない学習データでも効率良い学習を可能とした。

最後に3点目の工夫として、新規の受注製品の仕様に類似する過去の仕様図面を出力するためには、複数のグループ (学習モデル) から最も適したグループを選択する必要がある。そこで、新規の受注製品の仕様に類似する度合いをユークリッド距離で評価した。

3. 3 評価

複数の新規の受注製品の仕様図面を対象に図面検索を実施した結果、当初目標の70%以上の精度で新規の受注製品の仕様に類似する過去の仕様図面を検索可能であること確認したため、現在設計業務にて活用を開始している。

4. まとめ

AI技術の活用により、新規の受注製品の仕様に類似する過去の仕様図面を検索する図面検索ツールを完成した。図面検索ツールを実際の設計業務へ導入することにより、経験年数の少ない設計者の設計工数の約10%を占めていた図面検索時間が大幅に削減できた。さらに、付随効果として、適切な過去の仕様図面をツール操作のみで検索可能となるため、設計者間の負荷平準化および設計品質の向上も期待できる。

引き続き、図面検索ツールの水平展開を図るとともに、AI技術の新製品や社内業務への適用を幅広く推進していく予定である。

5. 謝辞

図面検索ツールの開発にあたり、終始ご指導を賜った京都工芸繊維大学情報工学・人間科学系寶珍輝尚教授に深謝する。

参考文献

- (1) 「2018年の技術と成果」、日新電機技報、Vol.64、No.1、pp.2-5(2019)
- (2) 元田、栗田 他：「パターン認識と機械学習 下」、丸善出版、pp.140-144(2012)
- (3) 荒木：「フリーソフトではじめる機械学習入門」、森北出版、pp.94-98(2018)
- (4) 元田、栗田 他：「パターン認識と機械学習 下」、丸善出版、pp.277-285(2012)

執筆者紹介



森本 充 Mitsuru Morimoto
研究開発本部
技術開発推進センター
博士(工学)



竹原 輝巳 Terumi Takehara
研究開発本部
技術開発推進センター
主幹