

浄水処理における深層学習を用いた凝集後濁度予測手法の比較

鈴木 昭弘[†] 川上 敬[†] 山村 寛[‡] 根本 雄一[§] 大江 亮介[†]
 北海道科学大学[†] 中央大学[‡] 前澤工業[§]

1 緒言

浄水場の浄水プロセスにおいて「凝集プロセス」は重要である。このプロセスは原水に凝集剤(PAC)を注入し攪拌することで、原水中の懸濁質を集塊したフロックと呼ばれる塊を生成する。このとき形成されるフロックの良し悪しは凝集剤の注入量によって制御するため、適切な注入量の決定が必要である。

浄水場では一般的に近似式を用いて注入量を決定しているが、安全のために過剰注入の状態にあると言われている。さらに台風などにより濁度が非常に高まった場合には近似式を用いることができない。そこでジャーテストとよばれる、原水に対して複数パターンの量の凝集剤を注入しフロックを形成させ、良好な結果が得られるまで繰り返すテストが行われる。ジャーテストの様子を fig.1 に示す。しかし、これは所要時間が長い問題がある。そして凝集剤の注入量と形成されるフロックの関係は複雑であり、pH、温度、アルカリ度など様々な要素で変動するため予測が困難であると言われている。

凝集剤の注入量を予測するために機械学習を用いて pH やアルカリ度といった化学的なパラメータから予測する手法が提案されているが、汎化性能が低いことが報告されている¹⁾。

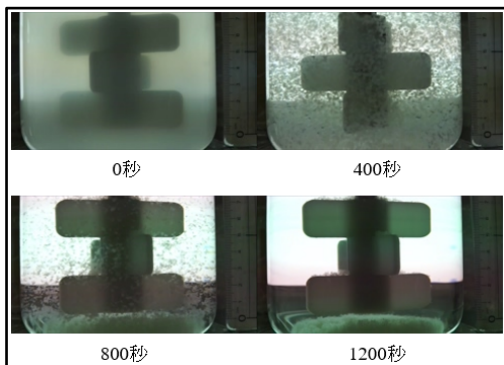


fig.1 ジャーテストの様子

そこで本研究では深層畳み込みニューラルネットワーク(DCNN)を用いることにより、凝集剤を注入しフロックが形成され始めた初期のフロック画像から最終的な凝集後濁度を予測するモデルの作成を試みている。これにより常に過不足の無い適切な量の凝集剤の注入を制御する仕組みである「凝集プロセス制御システム」を提供することが最終的な目的である。

著者らは VGG-16 を用いて凝集中の画像から凝集後濁度の予測を行ったところ、凝集開始後 200 秒経過後の画像を用い、0.88 程度の精度が得られることがわかった²⁾。そこで本稿では提案手法の有効性を検証するために、化学的なパラメータを用いて機械学習を行い、比較実験を行った。

2 システム概要

凝集プロセス制御システムの概要を fig.2 に示す。フロック形成池にて、①フロックが形成され始める初期の水を抽出し、そのフロックの様子を画像として撮影し、②現在実験および開発中のモデルを利用して凝集後濁度を予測する。そして、予測に基づき③適切な量の凝集剤を注入するよう制御するシステムを目指している。

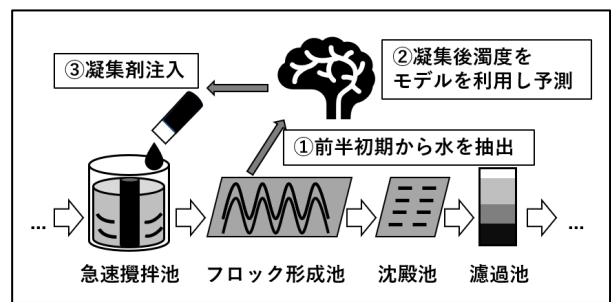


fig.2 凝集プロセス制御システム

3 実験

3.1 DCNN モデルと比較手法

本研究では凝集中画像から後述する3つの濁度区間のカテゴリを予測する。そのための DCNN として VGG-16³⁾を利用した。このモデルを利用し、これまでの実験では 0.88 程度の結果を得ている。最適化手法には Adam を使用し、損失関数には Softmax Cross Entropy を使用した。学習はバツ

A Study of Turbidity Prediction Method after Flocculation in Water Purification Process by Machine Learning
[†]Hokkaido University of Science
[‡]Chuo University
[§]Maizawa Industries

Table 1 データセット

名称	凝集後濁度	教師データ数	テストデータ数
カテゴリ 1	0.0-0.5 NTU	35	11
カテゴリ 2	0.5-1.0 NTU	32	10
カテゴリ 3	>1.0 NTU	45	15

チサイズを 500, エポック数を 200 にて実施した。

さらに、画像を用いる DCNN との比較対象として、原水の pH やアルカリ度といった化学的なパラメータを用いて、線形重回帰分析、非線形重回帰分析、SVM、Random Forest、MLP の 5 つの手法を用いた。これらをベースライン手法と呼ぶ。

化学的パラメータによる実験は Scikit-learn を用い、ハイパーパラメータは標準値のほか、GridSearchCV を用いて各パラメータの組について総当りで最大の精度を探索した。

3.2 データセット

河川から取水しジャーテストを行った様子を撮影した動画 148 本を用意した。各動画を Table 1 のように凝集後濁度別に 3 つのカテゴリに分類した。学習では濁度区間を教師ラベルとして用いる。カテゴリごとに教師データとテストデータは 7.5:2.5 に分けた。教師データは合計 112 本、テストデータは合計 36 本である。

動画は 5fps で画像化し、ビーカー内でフロックが明瞭に確認できる範囲を 200×200px の画像に切り出して使用した。また、フロックが形成され始める 200 秒から 300 秒区間の画像のみを利用した。最終的な教師データの画像は 56,000 枚であり、テストデータの画像は 18,000 枚である。

比較に用いる化学的パラメータは、凝集剤注入量、pH、原水濁度、水温、アルカリ度、UV-254、全有機体炭素 (TOC)、溶存有機炭素 (DOC) を用いた。これらは各動画を撮影する際に計測しており、動画数と同数の 148 のデータがある。

3.3 実験結果

ベースライン手法と提案手法の比較結果を Table 2 に示す。ハイパーパラメータのチューニング後の結果はアスタリスク付きで示した。ハイパーパラメータのチューニング項目は、SVM では C, gamma, kernel である。Random Forest では n_estimators, max_depth である。MLP では隠れ層のノード数、レイヤー数である。非線形重回帰分析では次数を変えながら実験を行った。

実験の結果、化学的パラメータを用いた手法では、多くの場合で精度は 0.7 前後に落ち着くことがわかった。いずれも提案手法の精度で

Table 2 ベースライン手法との比較

	手法	精度
ベースライン手法	線形重回帰分析	0.532
	非線形重回帰分析 (次数=2)	0.404
	非線形重回帰分析 (次数=3)	0.702
	非線形重回帰分析 (次数=4)	0.723
	SVM	0.570
	SVM*	0.680
	Random Forest	0.698
	Random Forest*	0.723
	MLP	0.718
	MLP*	0.744
提案手法		0.882

ある 0.882 に及ばなかった。また線形重回帰分析や次数 2 の非線形重回帰分析では精度が特に低く、凝集の様子が非線形であることが再確認できた。

以上のことから重回帰分析や一般的な機械学習手法では凝集後濁度の予測が難しいことがわかった。これは化学的パラメータのみでは凝集後濁度を把握するための説明変数が足りていないことを示唆していると考えられる。

一方で本研究が用いるフロック画像には凝集後濁度を予測するための特徴を有している可能性が高いと考えられる。さらに DCNN を用いることによって非線形な命題について、適切に各カテゴリの境界線を適切に引くことが可能になったと考えられる。以上のことから提案手法の有効性を確認できたと考えられる。

4 結言

本稿では、浄水処理の凝集処理中の画像から DCNN を用いて凝集後濁度の予測を行う提案手法とベースライン手法との比較実験を行った。

実験の結果、フロック画像には凝集後濁度を予測するための特徴が含まれていることを明らかにするとともに、提案手法の方がいずれのベースライン手法よりも精度が高く、有効性が高いことを明らかにした。今後はより予測精度を高めるための実験を行っていく。

参考文献

- Chan Moon Kim, Manukid Parnichkun: MLP, ANFIS, and GRNN based real-time coagulant dosage determination and accuracy comparison using full-scale data of a water treatment plant, Journal of Water Supply: Research and Technology-Aqua, 66-1, 49/61, 2017
- 鈴木昭弘, 川上敬, 山村寛, 根本雄一, 大江亮介: 浄水プロセス制御のための深層学習による凝集濁度予測, 第 21 回計測自動制御学会システムインテグレーション部門講演会, 2G2-02, 2020
- Karen Simonyan, Andrew Zisserman: Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv, vol. 1409.1556, pp.1-14, 2015