

画像認識における説明可能なAIに関する研究

Research on Explainable AI in Image Recognition

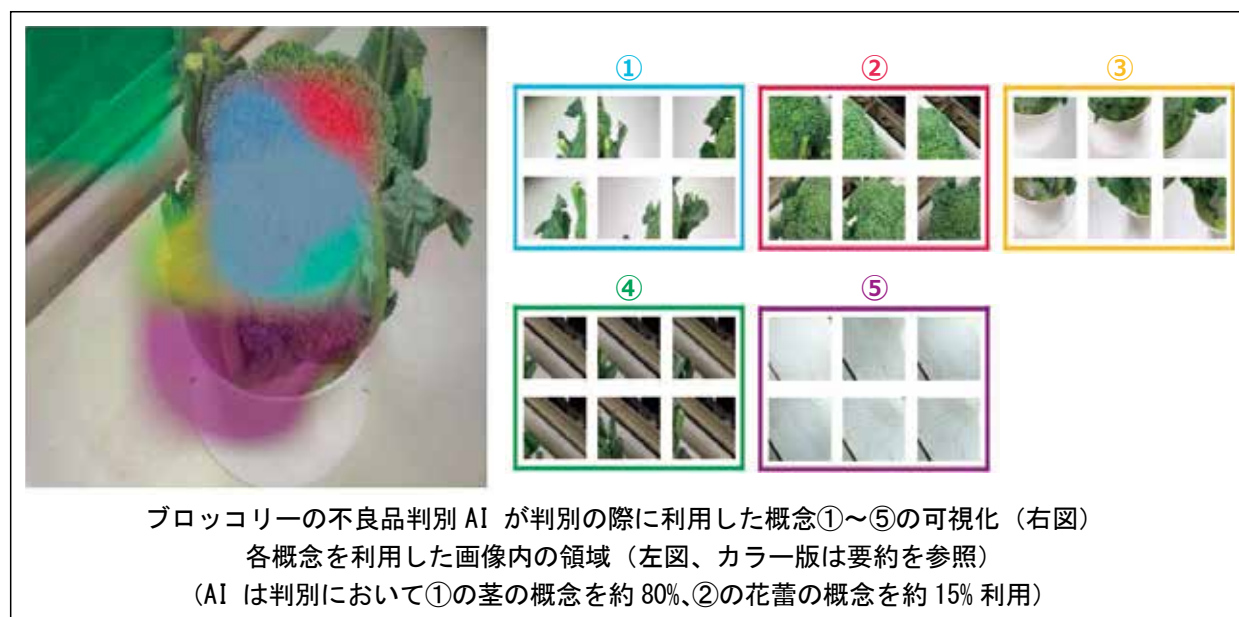
産業システム部 全 慶樹・本間 稔規
企画調整部 近藤 正一

■研究の背景

深層学習をはじめとするAIモデルは膨大な量のパラメータから構成されているため、内部構造を分析しても認識や予測の根拠を説明することは困難です。そのためAIを用いた各種技術開発に取り組む道内企業から、「なぜAIがそのような認識や予測をしたのか」などのAIの判断根拠を提示する技術を開発してほしいとの相談が当場に寄せられています。そこで本研究では、画像認識AIの判断根拠を可視化する最新の手法を既存のAIモデルへ適用するプログラムを開発しました。開発したプログラムを農作物の画像から不良品を判別するAIモデルへ適用し、その有用性を確認しました。

■研究の要点

1. AIの判断根拠を可視化する最新手法をAIモデルへ適用するプログラムの開発
2. 農作物画像から不良品を判別するAIモデルへの適用
3. AIモデルの妥当性の確認



■研究の成果

1. 畳み込みニューラルネットワークによる画像認識AIの判断根拠を可視化するプログラムを開発しました。
2. 開発したプログラムをブロッコリーの不良品を判別するAIモデルへ適用した結果、AIの判断根拠が可視化され、AIモデルの問題点を明らかにすることができました。
3. AIモデルの妥当性を確認できるようになり、より信頼性の高いAIの開発が可能になりました。