

第24回 愛媛大学 DS 研究 セミナー

愛媛大学データサイエンスセンター(CDSE)は、AI・統計解析・機械学習等の広義でのデータサイエンスと接点のある研究者、実務家、教育者を学内外から招聘し、講演していただくデータサイエンスセミナーを開催していきます。

日時 2023年
4月28日(金) 16:30-18:00

定員 オンライン開催 Zoom先着300名(YouTube同時配信)

「MetaFormers：Transformerによるコンピュータビジョンの変革」



【講演者】 瀧 雅人氏
立教大学・准教授

深層学習の広範な分野での成功・高い性能の秘密は、単に深いニューラルネットを使うということではなく、ネットワーク構造のデザインにより様々な帰納バイアスを柔軟に実現できる点にある。その典型的な例は、画像タスクにおける畳み込みニューラルネット(CNN)である。この10年の画像タスクにおける深層学習の大きな成功も、CNNの持つ帰納バイアスをベースにしていると言って良い。ところが2020年になり、畳み込み構造に一切依存しない深層学習Transformerによって、SOTA CNNに匹敵する性能を実現できることが判明した。この新しいモデルはVision Transformer(ViT)と呼ばれる。ViTは自然言語処理分野において発展したself-attentionを使うことで、非局所的なパターンも動的に抽出することが可能なアーキテクチャである。また翌年にはMLP-Mixerと呼ばれる手法で、self-attentionを使わずとも、多層パーセプトロン(MLP)をベースにしたモデルだけでも同等の性能の画像認識が可能であることが判明した。このように、現在CNNパラダイムとは全く別の画像認識の方向性が明らかになりつつある。このViT、及びMLP-Mixerの発見は、データセットサイズがスケールアップした深層学習では、必ずしもCNNのような強い帰納バイアスは必要でなくなることを実証している。また訓練データセットがスケールしていくとやがてViTがCNNを凌駕することから、場合によっては強い帰納バイアスがモデルの性能に限界をもたらすことも具体的にわかってきた。このようなpost-CNNコンピュータビジョンモデルは現在、MetaFormerという統一的な枠組みで研究されている。このセミナーでは、MetaFormerの現状とこれまで判明している興味深い事実、およびその欠点について議論したい。

参加費
無料

お申込み

参加を希望する方は右側Googleフォームからお申込みください。もしくは下記URLよりお申し込みください。

<https://forms.gle/x47eD2TG4K7Ve77E9>

締切：4月26日(水)12:00

※資料は事前配布のみです。後日の配布は致しませんので、ご了承ください。



主催：愛媛大学データサイエンスセンター



CDSE
Center for Data Science, Ehime University

お問い合わせ／愛媛大学データサイエンスセンター（石川、越智）
E-mail: cdse@stu.ehime-u.ac.jp
H P : <https://www.cdse.ehime-u.ac.jp/>