

主催：一般社団法人 日本ロボット学会

協賛：計測自動制御学会、産業技術連携推進会議 医療福祉技術分科会、システム制御情報学会、情報処理学会、人工知能学会、精密工学会、電気学会、電子情報通信学会、土木学会、日本感性工学会、日本機械学会、日本シミュレーション学会、日本神経回路学会、日本設計工学会、日本時計学会、日本人間工学会、日本バーチャルリアリティ学会、日本ロボット工業会、農業食料工学会、バイオメカニズム学会（以上 20 団体予定）



第 152 回 ロボット工学セミナー

ロボットのための LLM・VLM 利活用

日時：2024 年 5 月 23 日（木）9:50～17:20（9:30 開場）

会場：日立製作所中央研究所 日立馬場記念ホール およびオンライン配信

（詳細はお申込み後にご案内します）

住所：東京都国分寺市東恋ヶ窪一丁目 280 番地（JR 中央線 国分寺駅下車北口 徒歩約 10 分）

定員：現地 90 名、配信 180 名（申込締切詳細は案内ページを参照ください）

参加費（税込）：現地/オンライン共通

※ お支払の際、別途システム手数料「220 円」を頂戴致します。

当学会及び協賛学会の正会員（個人）／8,500 円、会員外（一般）／20,000 円

当学会及び協賛学会の学生会員（個人）／3,000 円、会員外（学生）／10,000 円

当学会賛助会員 招待券ご利用／無料、優待券ご利用／3,000 円、左記サービス券なし／20,000 円

特別優待券使用の場合：学生（RSJ 会員非会員問わず）／無料、学生以外／3,000 円

オンラインのみ：団体申込み（*）～10 名まで／40,000 円、～25 名まで／100,000 円

（*）団体申込みは、申込者は RSJ 会員であること、また申込者以外の聴講者は学生であることが条件です。

口上：近年発展を続けている LLM (Large Language Models) 及び VLM (Vision Language Models) は、ロボットシステムでも活用されるようになってきました。ロボットシステムへの LLM・VLM 利用は、例えばロボットの利用者視点から見ると、ロボットの操作が容易に利用できるといった利点を持ちます。一方で実利用時は、ロボットの低レイヤーの制御や基本的な認識技術をどこまで用意する必要があるか、どのような入力が有用かなど、実際に開発する際に検討を要する項目もあります。本セミナーでは専門家の方々に講師を迎えて、LLM・VLM の概要から様々なロボットへの LLM 応用までをご紹介します。本セミナーを通じて、LLM・VLM をロボットに利用する知見を共有します。

オーガナイザー：坂井 亮（日立製作所）

WEB サイト：「日本ロボット学会>ロボット工学セミナー」よりご確認ください。

<https://www.rsj.or.jp/seminar>

講演内容：

09:30 <開場>

09:50-10:00 <開会挨拶・講師紹介>

10:00-11:00 第 1 話 自己教師あり学習によるビジョン基盤モデルの事前学習

中部大学 藤吉 弘亙

自己教師あり学習は、ラベルなしデータを用いた事前学習法である。自己教師あり学習では、データから正

解ラベルを自動で作成できるプレテキストタスクにより事前学習を行い、様々な下流タスクに効果的な特徴表現を獲得する。対照学習に基づく手法が教師あり事前学習法と同程度以上の学習効果が示されて以降、対照学習の派生手法が数多く提案された。また、Vision Transformer (ViT) の台頭以降は、ViT の構造に合わせた方法として Masked Image Modeling が提案され、現在に至るまで盛んに研究が行われている。本チュートリアルではまず自己教師あり学習の概要と代表的な評価方法について紹介し、対照学習と Masked Image Modeling 手法について解説する。

11:00-11:05 <休憩>

11:05-12:05 第 2 話 ロボティクスにおける大規模言語モデル・視覚言語モデルの利活用

慶応義塾大学 杉浦 孔明

大規模言語モデル (LLM) や視覚言語モデル

(VLM) の急速な発展は社会全体に影響を与えており、ロボティクスも大きな転換期を迎えている。複雑な環境の理解、自然な対話、長期の動作計画、等を統合したロボットシステムの構築には複数分野に亘る専門性が必要であったが、LLM や VLM は各機能の汎用性を向上させるとともに、それらの統合を容易にした。本講演では、LLM や VLM の概要、視覚言語情報の利用方法、マルチモーダル言語処理の応用、ロボティクス応用における注意すべきポイントと展望について述べる。

12:05-13:00 <休憩(昼食)>

13:00-14:00 第3話 LLM・VLMの実ロボット応用例とその分類

東京大学 河原塚 健人

LLM・VLMに代表される基盤モデルは、言語や画像、音声、深度などの多様な入出力を扱うことができ、すでにロボットの認識・制御・動作計画などの様々な領域に応用が広がっている。本講演ではまず、実ロボットへの適用という観点から、多様な入出力を持つ基盤モデルを分類し、その特徴を整理する。次に、ロボットの認識・制御・動作計画のそれぞれについて、どのような形で基盤モデルを適用することができるのか、世界中の研究例、そして我々の事例を紹介する。

14:00-14:05 <休憩>

14:05-15:05 第4話 ロボット基盤モデルの構築に向けて

東京大学 松嶋 達也

近年、GPTやCLIPなどの大規模言語モデル(LLMs)や大規模視覚言語モデル(VLMs)を、認識やプランニングのモジュールとして活用し、実世界の中でタスクや環境の変化に頑健な汎用性の高いロボットシステムの研究開発が盛んになっている。一方で、ロボットが実世界で行動することにより蓄積されるデータを活用して、ロボットの制御とより密接に結

びついた基盤モデル(いわばロボット基盤モデル)を構築しようとする研究も進行している。本発表では、これらのロボット基盤モデルの構築を目指した、モデルやデータセットのスケール化に向けた取り組みを解説する。

15:05-15:15 <休憩>

15:15-16:15 第5話 大規模マルチモーダル物理基盤モデルとしてのロボティクス

早稲田大学 尾形 哲也

大規模言語モデル(LLM)のロボティクス利用が活発となってきている。しかし言語は世界の抽象表現であり、今後、マルチモーダル情報を基盤とした世界モデル(能動的なインタラクションモデル)が重要となるだろう。本講演では、これらの概念の基盤となる能動的推論についての基本的な考え方を概説し、実ロボットへの応用事例を紹介する。またムーンショットプロジェクトを例に今後の展望についても議論する。

16:15-16:20 <休憩>

16:20-17:20 パネルディスカッション
<閉会>

内容は予告無く変更される場合がございます。最新情報は学会 WEB ページにてご確認ください。

<申込方法/問合せ先>

各セミナーの詳細を学会 HP (<https://www.rsj.or.jp/seminar/>) からご確認くださいの上、お申し込み下さい。

お問合せ先: RSJ 事務局セミナー係: TEL. 03-3812-7594 E-mail. seminar@rsj.or.jp

※ 賛助会員招待券/優待券および特別優待券(複数回受講特別優待制度)の詳細は下記 WEB ページをご参照ください。

<https://www.rsj.or.jp/event/seminar/ticket/>

<注意事項>

1. 会場、講師、日時等は都合により変更になる可能性がございますのでご了承下さい。最新の情報は学会ロボット工学セミナーHP (<https://www.rsj.or.jp/seminar/>) に掲載されます。
2. 警報発令時のセミナー開催中止判断については右記 WEB ページをご確認ください。 (<https://www.rsj.or.jp/event/seminar/cancel/>)
3. 参加者の理解を深めるため電子テキストを配布致します。この電子テキストは、諸事情により講演に使用されるスライド資料から修正・抜粋がされている場合がございます。ご了承下さい。また、電子テキストの後日販売は行いません。
4. 参加者のセミナー会場内での撮影・録音行為は禁止させていただきます。なお、撮影・録音を含む取材をご希望の場合は必ず事前に学会事務局までお問い合わせ下さい。