

NTTコミュニケーションズ
イノベーションセンター テクノロジー部門 担当課長

田良島周平 Shuhei Tarashima

NTTにおけるメディアAI技術の ハブとなるチームの構築を めざして

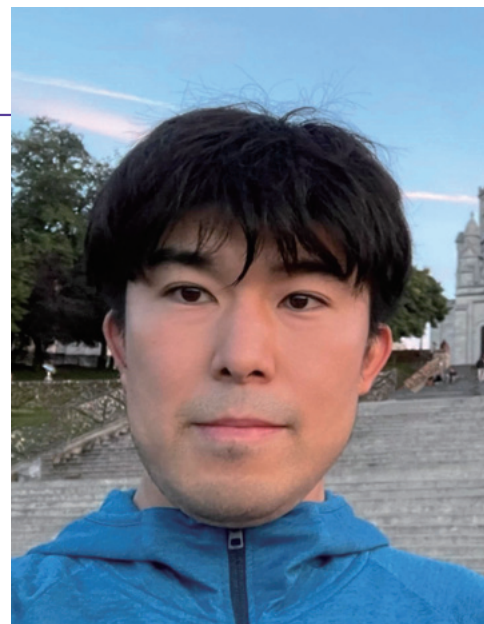
2023年はLLM (Large Language Model) をはじめとする生成AI (人工知能) が大きな注目を集めました。これに連動するように、AI技術全体への期待も今まで以上に高まってきています。一方で、AI技術の新陳代謝は加速度的に進んでおり、先進的なAI技術をお客さまが求めるサービスやソリューションとして提供することの難易度も上がり続けています。こうした環境の中、NTTコミュニケーションズ イノベーションセンターで、画像や映像・言語などメディアを対象とするAI技術 (メディアAI技術) の開発に取り組む田良島周平氏に、プロジェクトのこれまでの取り組みと今後のチャレンジについて伺いました。

「CoE」と「R&D」の両輪でメディアAIの技術開発に取り組む

現在の業務の概要をお聞かせください。

NTTコミュニケーションズ イノベーションセンター テクノロジー部門で、画像や映像・言語などメディアを対象とするAI (人工知能) の技術開発に取り組むプロジェクト (メディアAIプロジェクト) のリーダーを担当しています。2024年2月現在、メディアAIプロジェクトは創設3年目で、本務のメンバー7名を中心にメディアAIに関する「CoE (Center of Excellence)」と「R&D (Research & Development)」に取り組んでいます。CoE活動では、NTTコミュニケーションズでビジネスを推進している、あるいは新たに立ち上げようとしている部署からの要望を受け、それにこたえる技術検証、システム開発、カスタマイズを行っています。一方R&D活動では、プロジェクトメンバーがエンジニア・研究者の視点で取り組むべきテーマを選定し、数年先等の事業貢献を見据えた技術調査、研究開発、そしてそれらのPR活動を推進しています。

テクノロジー部門で取り組むICT技術群の中で、メディアAIはもっともアプリケーションに近い技術領域の1つです。NTTコミュニケーションズが手掛ける広範なビジネス領域で求められるメディアAIの機能は多岐にわたるのですが、技術の新陳代謝の極めて速いAI分野で、事業部やお客さまのニーズにいかにか正確かつスピーディにこたえていくかは私たちの大きなチャレンジとなっています。そのためには事業部ひいてはお客さまのニーズに直接こたえる



「CoE」活動はもちろんのこと、いわゆるシーズに基づく「R&D」活動を並行して進めていくことが極めて重要であると私は考えています。

「CoE」活動について、いくつか事例を紹介いただけますか。

「CoE」に関する最近の事例として、「COTOHA Takumi Eyes 向け新機能開発・既存機能高度化」と「除雪業務効率化に向けたメディアAI機能の開発」を紹介します。

まず1番目について、COTOHA Takumi Eyes⁽¹⁾は、防犯やマーケティング、おもてなしなどの領域に応用可能なNTTコミュニケーションズの映像解析AIソリューションです。COTOHA Takumi Eyesはさまざまな映像解析機能を有していますが、新規顧客への提案を行う中で「人物全身のうち特定の領域 (例えば靴等) にフォーカスし、その領域における対象の見え方に基づく画像検索を行いたい」といった、従来にない機能に関するニーズと、ユーザー体験向上のためには「既存の映像解析機能の推論の高速化が必要」といった要件が浮き彫りになりました。これら要望を受けメディアAIプロジェクトで技術検証を行い、前者のニーズにこたえる技術として、画像中で注目すべき領域を特定する機構を導入した人物画像検索技術 (図1) を開発し、従来技術に対し約20%の検索性能改善を達成しました。また後者についてはAI推論高速化 (TensorRT) のノウハウを活用し、認識性能を劣化させることなく2倍以上の高速化を実現させました。両機能とも最新のCOTOHA Takumi Eyesに搭載されており、お客さまから高い評価を受けています。特に前者の機能は特許出願中であり、類似

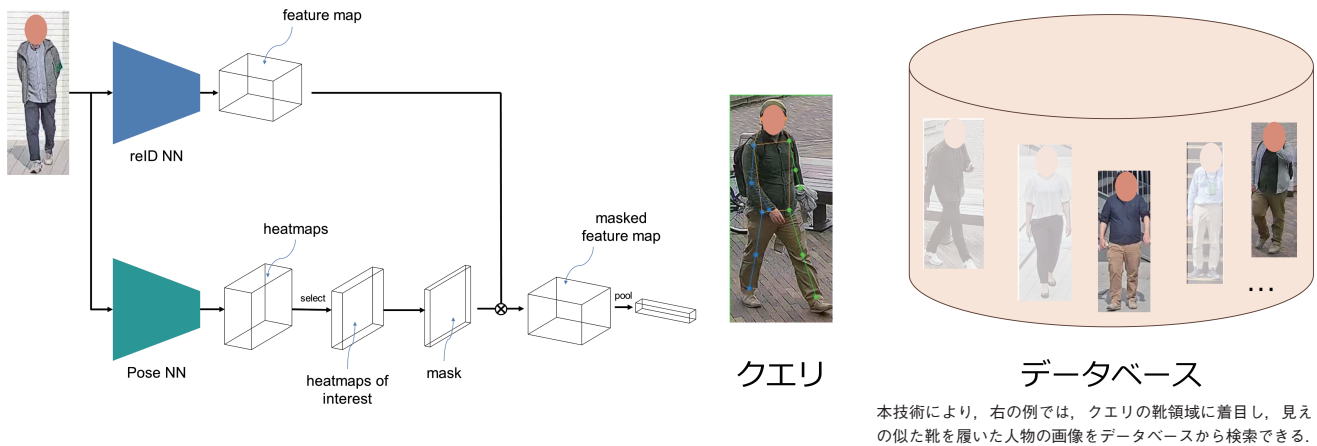


図1 画像中で注目すべき領域を特定する機構を導入した人物画像検索技術

ソリューションに対するCOTOHA Takumi Eyesの差別化機能として大きく注目されています。

2番目では、NTTコミュニケーションズが自治体に提供している積雪状況可視化プラットフォーム⁽²⁾に除雪業務の効率化に資するようなAIを追加すべく、地方自治体のお客さまとの共同実験として機能の具体化と開発に取り組んでいます。ドライブレコーダや車載カメラから得られる映像の解析技術の一種といえますが、「除雪業務の効率化に資するようなメディアAI」に関する既存の取り組みは存在せず、取り組み開始時点ではそもそもAIでどのようなことができればよいのかすら自明ではありませんでした。私たちは事業部と連携しつつお客さまと複数回にわたって議論を重ね、関係者全員が納得できる機能を特定してきました。現在は、特定した機能を精度良く実現すべく、学習に用いられるデータが希少な状況下でも正確な判定が行えるようモデルや学習方法を検証しており、ポジティブな結果が得られつつあります(図2)。豪雪地帯では、積雪は交通機関や電気などの生活インフラに多大な影響を及ぼしており、自治体では除雪関連業務に多大なコストと労力をかけています。今後は本機能の改善に加え積雪状況可視化プラットフォームへの導入や他自治体様への導入支援にも取り組み、除雪関連業務のコストと労力の削減に、より貢献していきたいと

考えています。

これら以外にも、AI空調制御クラウド⁽³⁾サービスに資するメディアAI機能を提供したり、データセンタ向け新ソリューションや宇宙関連新事業の創出にメディアAIの観点で貢献したりするといった取り組みにも着手しています。現状社内でのクローズドな取り組みが多いのですが、遠くない未来に広く外部にもPRできるよう、プロジェクトメンバーと各案件を成功に導いていきたいと考えています。

「R&D」活動ではどのような取り組みを行っているのでしょうか。

まず、メディアAIプロジェクトで特に着目している技術の1つにNeRF (Neural Radiance Fields) があります。NeRFはAIで高品質な自由視点画像・映像を生成する技術として近年特に注目されており、すでに実用化されているフォトグラメトリベースの方法と比較して光学現象を精緻に再現できるというメリットがあります。NeRFの処理は大きくモデル学習とその推論に分けることができますが、メディアAIプロジェクトではその両者についてノウハウ蓄積や技術研鑽を行っています。まずモデル学習につい

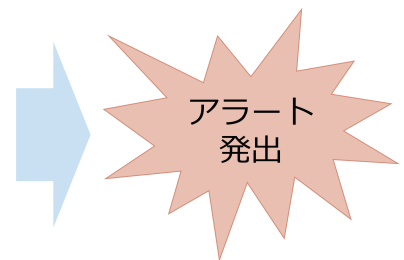


図2 除雪業務の効率化に資するようなメディアAI機能の開発

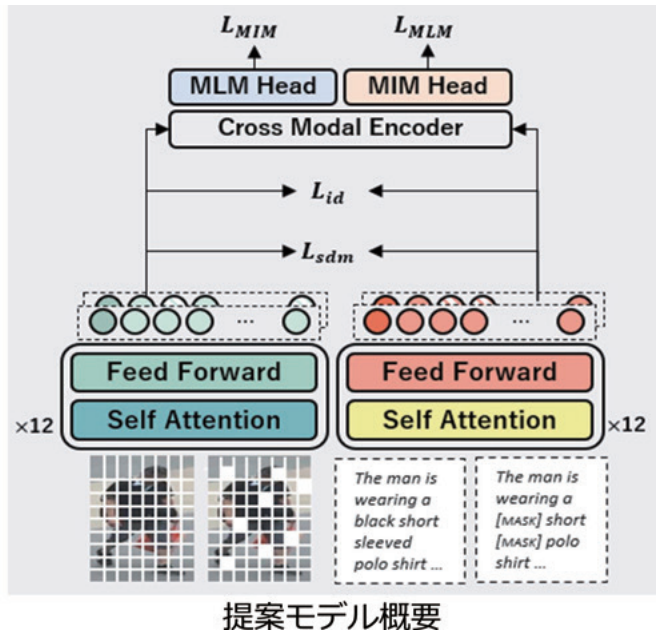


図3 クロスモーダル人物検索技術

では、モデル学習にあたって入力画像・映像間の相対位置を決定する必要があり、これは各画像から抽出される局所特徴点のマッチング（画像マッチング）で実現されることが一般的です。この技術を磨くため、当該タスクの正確さを競う国際コンペ「Image Matching Challenge 2023⁽⁴⁾」にプロジェクトで参加し、試行錯誤の結果入賞（494チーム中94位、Bronzeメダル）を果たしました。コンペ参加を通じて培った技術は、NeRFだけでなく前述の「CoE」活動へも応用を進めています。次に推論について、さまざま存在する課題のうち私たちは特に「推論に時間がかかる」点に着目し、その時間の短縮につながる手法の開発を進めています。こちらについても、2023年に発表された技術に対し約5倍程度の高速化を達成しており、その成果を含む内容を2023年の映像情報メディア学会冬季大会内で招待講演として発表しました⁽⁵⁾。

上記のような成果が出つつあるものの、NeRFの実行には依然GPU（Graphics Processing Unit）等の高性能なAI計算リソースを必要とします。今後は、エンドユーザーの手元にそのようなリソースがない状況でも遅延なくNeRFアプリケーションを実行可能とするため、NTTドコモのモバイルエッジコンピューティング（docomo MEC⁽⁶⁾）やNTTのオールフォトニクス・ネットワーク（IOWN APN⁽⁷⁾）との連携にも取り組んでいきたいと考えています。

R&Dの結果新しい技術の創出につながった際には、特許出願や学会などの対外イベントへの投稿・発表にも積極的に取り組んでいます。2023年には、コンピュータビジョン分野最高峰の国際会議の1つであるICCV 2023の併設ワークショップにて、テキストをクエリとしてその内容に合致する人物画像をデータベースから検索するクロスモーダル人物検索技術（図3）の発表⁽⁸⁾や、BMVC 2023という同分野難関国際会議にて、東京都立大学との



共同研究成果として、スポーツの映像におけるボールをはじめとする、小さな物体の検出追跡技術（図4）の発表⁽⁹⁾等を行いました。特許出願や対外発表につながる技術は「CoE」活動に取り組む中で着想が得られるものも少なくありません。今後はこれらR&Dの成果をCoEへも還元できるよう、技術の応用可能性を模索していきたいと考えています。

NTTコミュニケーションズ・NTTドコモグループ・NTTグループにおけるメディアAI技術のハブとなり得るチームをつくり上げることがミッション

プロジェクトでの技術研鑽にはどのように取り組んでいますか。

まずメディアAIプロジェクト内では、毎週の定例ミーティングの中で、メンバー持ち回りで勉強会を開催しています。トピックについて限定はしておらず、論文調査を行うメンバーもいれば分野のビジネス動向を調査するメンバーもいますし、中にはUnreal Engine5等のゲームエンジン上でウォークスルー可能なバーチャルオフィスを構築して紹介してくれるメンバーもいます。また部署を横断した勉強会も定期的で開催しており、例えば前述のCOTOHA Takumi Eyesの開発チームとは、コンピュータビジョン分野のトップカンファレンスであるCVPR 2023に共に参加し社内外で参加報告を実施しました。これらの勉強会を開催する中で意識していることの1つは、勉強調査する中で得られた知識を外部的向けアウトプットするという点です。幸いNTTコミュニケーションズには開発者ブログという外部公開の場があるので、外部イベント参加の報告や勉強会資料は積極的にブログへ投稿させて



緑：正解 赤：予測

図4 小さな物体の検出・追跡

もらっています⁽¹⁰⁾。

NTTコミュニケーションズでの業務の魅力を教えてください。

私は、NTTに入社以来約7年間NTT研究所に在籍していたのですが、その後NTTコミュニケーションズに異動することで、事業部あるいはお客さまの声をより直接的に聞くことができるようになったと感じています。この声にこたえるアプローチはさまざま考えられ、成熟した技術で対応可能である場合にはそれを迅速に適用すればよいですし、該当する技術がないのであれば既存の技術を組み合わせたり、あるいはゼロからつくり上げたりしていくことになります。このように、ニーズと必要な技術に応じて適切なアプローチを主体的に選択できるという点は、NTTコミュニケーションズで私が所属している組織における業務の魅力の1つだと感じています。

業務で今後取り組みたいことを教えてください。

先ほどさまざまなアプローチを主体的に選択できると申し上げましたが、それは取り組むべき業務の範囲が極めて広範であることも意味しています。現時点でメディアAIプロジェクトに所属するメンバーは決して多いわけではなく、実際には聞こえてくるニーズの中で十分対応しきれていない「CoE」施策や、取り組みたいのに手が付けられていない「R&D」トピックが存在しています。メディアAIプロジェクトでの技術開発をよりスケール・加速させられる体制を構築することが、今後の私の最も大きなミッションの1つです。

アプローチの1つは、プロジェクトの拡大と各メンバーの成長促進です。私たちの取り組みの発信を通じて仲間になってくださる方をNTT内外から集めることと、メンバーの成長機会を創出することに積極的に取り組んでいきたいです。

また、もう1つのアプローチとして、プロジェクト外の組織と連携を拡大していくことも重要だと考えています。私はNTT研究所で研究活動に取り組む中で、研究開発に長けたNTT内外の方々とのヒューマンネットワークを得ることができました。今後はこのヒューマンネットワークを、プロジェクトの「CoE」と「R&D」によりよく活かす方法を模索していきたいと思っています。典型

的な方法はNTT研究所の技術を成果提供で受け取ることだと認識していますが、特に技術の新陳代謝が極めて速いAI領域では、これ以外の方法も模索すべきではないかと考えています。例えば研究者・エンジニアの交流の場を設けたり、事業会社で技術に精通した私たちから研究所に取り組みでほしいテーマを提案したりするなど、新しい連携のあり方の創出にも取り組んでいきたいです。

これらの取り組みを通じて、NTTコミュニケーションズ・NTTドコモグループ・NTTグループにおけるメディアAI技術のハブとなり得る組織をつくり上げていきたいと考えています。

社内外の技術者へメッセージをお願いします。

NTTグループ・NTTドコモグループ・NTTコミュニケーションズの中でもAI技術への注目度は非常に高くなっています。現在私たちは本務7名でメディアAIの技術開発に取り組んでいますが、今後はその取り組みをより拡大・加速させていく予定です。私たちと一緒にその活動に取り組んでいただける方をNTTグループ内外から募集しております⁽¹¹⁾、⁽¹²⁾ので、もしご興味を持っていただけでしたらぜひご応募・ご検討いただけると嬉しいです。よろしくお願いたします。

参考文献

- (1) <https://www.ntt.com/business/solutions/enterprise-application-management/takumieyes.html>
- (2) <https://www.ntt.com/about-us/press-releases/news/article/2022/0204.html>
- (3) <https://www.ntt.com/business/services/aiaccontrol.html>
- (4) <https://www.kaggle.com/competitions/image-matching-challenge-2023>
- (5) <https://www.ite.or.jp/winter/2023/program/feature/#S5>
- (6) <https://www.mec.docomo.ne.jp/index.html>
- (7) <https://www.ntt.com/about-us/technology-strategy.html>
- (8) https://openaccess.thecvf.com/content/ICCV2023W/CLVL/html/Fujii_BiLMa_Bidirectional_Local-Matching_for_Text-based_Person_Re-identification_ICCVW_2023_paper.html
- (9) <https://proceedings.bmvc2023.org/310/>
- (10) <https://engineers.ntt.com/archive/category/%E3%83%A1%E3%83%87%E3%82%A3%E3%82%A2A1>
- (11) NTTグループ公募 NTTコミュニケーションズの募集ポスト【IC】メディアAI技術の開発検証【イノベーションセンター】
<https://n-jobboard.bnccloud.net/jcs/Pages/A001.aspx>
- (12) <https://hrmos.co/pages/nttcom0033/jobs/1928706488764108823>