

画像認識技術が変える検索の世界

～リクルートの人工知能 API「A3RT」の導入事例から～

舟木 類佳

株式会社リクルートテクノロジーズ
IT ソリューション統括部データテクノロジーラボ部
データテクノロジープロダクト開発グループ

1 はじめに

弊社、株式会社リクルートテクノロジーズは A3RT (Analytics & Artificial intelligence API via Recruit Technologies の略称。「アート」と読む。) という人工知能のブランドを展開しています。これは Deep Learning などの機械学習ロジックを一つのブランドとして整備し、リクルートグループの各事業会社へ展開するためにプロジェクト化されたソリューションの総称です。A3RT では図 1 のような画像解析や、文章分類、チャットボット、文章生成、文章校閲などの様々なプロダクトが含まれており、社内の様々なサービスで活用されています。

これまでは社内での利用に限定されていた A3RT の API ですが、それらの一部を、2017 年 3 月に外部公開¹しました。公開されている API を用いることで、リクルートで活用されている技術を体験することができるようになっています (図 2)。

私達は日頃リクルートの提供するサービスを機械学習を用いてより良くするべく取り組んでいますが、その中でもとても重要と考えているのが検索体験の改善です。リクルートは様々な領域の Web サービスを展開しており、カスタマーはそれらを利用することで仕事、住宅、結婚式場、レストラン、ヘアサロン、宿、車などを探します。これらの Web サービスにおいてカスタマーが目的のものを見つけるためには検索機能は核となるものであり、私達はそこに機械学習技術を適用することでカスタマーがより容易に探しているものにたどり着くことができるような仕組みを実現しています。

本講演では、そうした検索の中でも特に画像を用いた検索に焦点を当て、「画像認識技術がカスタマーの検索体験をどのように変えていくのか」ということを A3RT を用いた取り組みの事例からご紹介いたします。

2 検索における画像認識技術の適用事例

画像は人にとって理解しやすい情報であり、非常に直感的です。そして、直感的で分かりやすいことは検索においてとても重要です。画像の直感的であるという特性を検索にも活用できれば、より目的のものを見つけやすくなると考えられます。私達は画像認識技術を検索と画像情報を橋渡しする鍵として捉えており、複数の検索改善プロジェクトに画像認識を利用してきました。

リクルートのサービスにおいて、画像認識技術がどのように検索で応用されているのかを順にご紹介いたします。着目していただきたいのは、画像認識技術が検索における課題を解決する手段となっているということです。今回事例を大きく下記の 3 つのタイプに分けました。

1. 「探しにくい」を解決する仕組み
2. 「そもそも探せない」を解決する仕組み
3. 「探したけれど結果に納得感がない」を解決する仕組み

これらの仕組みを、それぞれ順に説明していきます。

2.1 「探しにくい」を解決する仕組み

カスタマーが検索において目的のものを探す際に、より探しやすいするために画像認識技術が活躍しています。

2.1.1 ホットペッパービューティにおけるネイルレコメンデーション機能

まず第一にホットペッパービューティでの事例をご紹介します。ホットペッパービューティは美容院、

¹A3RT: <http://a3rt.recruit-tech.co.jp>



図 1: A3RT の社内向けプロダクト一覧

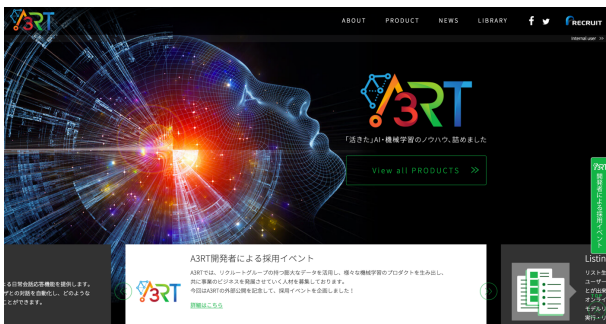


図 2: A3RT の Web ページ

美容室、ヘアサロン、ネイルサロン、エステサロン、リラクゼーションサロンを検索、予約できる Web サイトです。

課題 ホットペッパービューティにはネイルサロンの検索機能がありますが、思い描くイメージにあったネイルにたどり着けるか、という部分が重要となってきます。そこで、雑誌のネイルカタログをめくるように、カスタマーが良いと思ったネイルを行き来しながら自分の欲しいネイルに到達できるようにすることで、気になるネイルの発見機会の向上、回遊率の向上を実現したいというニーズがありました。

解決策 これを実現するために、私達はネイルのレコメンドーションを行うことを考えました。ネイルデザインを探す際にはまさに見た目が重要となってきます。そこで現在見ているネイルに対して類似しているデザインの画像を他の候補としてレコメンドする、ということを行いました (図 3)。これにより、今見ているネイルに類似したネイルにより簡単に遷移することがで



図 3: ホットペッパービューティにおける類似ネイルレコメンドーション

きるようになります。

技術 ネイルの画像は背景を含んでしまっているので単純にはいかず、爪の領域を切り出す必要があります。そのために画像から HOG (Histograms of Oriented Gradients) 特徴量を用いて AdaBoost により認識を行いました。そこから、畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) による画像分類を行い、どのようなデザインなのかのラベルを認識します。それと同時に、カラーヒストグラムから色情報を取り出し、デザイン情報とあわせてコサイン類似度を取り、似ているデザインを探し出します (図 4)。

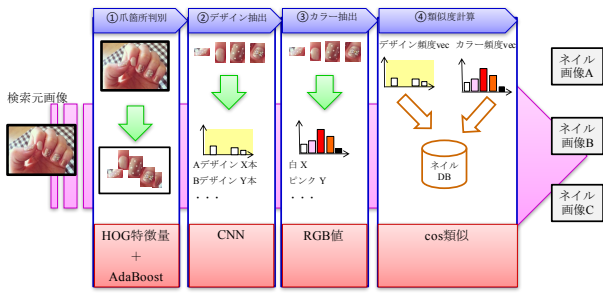


図 4: ネイルレコメンデーションの仕組み



図 5: 内装色判別機能

2.1.2 カーセンサーにおける内装色判別機能

リクルートが運営している中古車情報のサービスである「カーセンサー」での事例をご紹介します。

課題 カーセンサーにおけるカスタマーアンケートの結果、内装色で検索できなくて困っているということが分かりました。車を検索する際に、検索条件に内装(シート)色で検索する機能はありませんでした。当時の現状では内装色で探したいというユーザーはまず検索して、その後詳細ページに置かれている画像から一つ一つ目で内装色がどのような色なのかを確認する必要がありました。

解決策 そこで、内装色が何色なのかということを予め判定して色ラベルを付与することによって、検索時にフォームから探したい内装色を選べるようにしました。自動で内装色を絞り込むことができるため、探しやすさが向上します。

技術 一つの車に対して複数の画像はあるものの、内装画像がどの画像かという情報はありません。そこで、まず「外装」「トランク」「ダッシュボード」「内装」「その他」などを予め CNN により画像分類しました。そして内装と判定された画像から広く利用されている色をカラーヒストグラムとして抽出することで、



図 6: カーセンサーアプリの車種判別機能

最も広範囲に利用されている色ラベルを付与するというを行いました。

2.2 「そもそも探せない」を解決する仕組み

検索をする際に検索自体ができなくて離脱してしまうユーザーを減らすために、画像認識技術が利用されています。

2.2.1 カーセンサーにおける車種判別機能

カーセンサーにおける別の画像認識技術活用事例をご紹介します。

課題 道端で車を見かけて「カッコいい車だな」「おしゃれな車だな」などと思い、購入しようと思って検索しようとしたとします。皆さんならどのように検索しますか。車にとっても詳しい人であれば、「これは何というメーカーの何という車種だから・・・」とキーワードを入力することで検索することができるかもしれませんが、実際そこまで詳しい人はなかなかいません。「あの車が欲しいのに探し方がわからず購入ができない」という状況は大きな機会損失と言えるでしょう。

解決策 ではどのようにこの状況を改善することができるのでしょうか。私達は画像認識による車種判別を行うことによりこの問題を解決しています(図6)。道端で車を見つけたときに、カーセンサーのアプリを利用してカメラを起動し、その車を写真に撮ります。あるいは雑誌に載っている車を撮っても良いでしょう。そうすると、A3RTの車種判別システムが画像から車種を予測して予測結果を教えてください。それによって得られた車種情報から中古車一覧が検索され、閲覧できるようになります。



絞り込む

図 7: ゼクシにおける従来の絞込機能

技術 このシステムでは CNN を用いて画像と、その正解の車種名の組み合わせを学習しています。CNN が車のそれぞれの車種の画像が持つ特徴を自動で抽出してくれることで、高精度な車種名の予測が実現しています。

2.2.2 ゼクシにおけるウェディングフォトサーチ機能

結婚準備のウェディング情報を提供しているゼクシ net での事例をご紹介します。

課題 ゼクシでは結婚式場を掲載しており結婚式場を検索することができます。このサービスを利用しているユーザーに対して調査を行ってみると、自分の結婚式のイメージがまだ無かったり、頭の中で漠然としたイメージがあってもうまく言葉に表せないカスタマーが多く、そうした人々がうまく式場を見つけられず離脱しているということが分かりました。

従来のシステムでは図 7 のように「教会式」や「神前式」などの挙式スタイルや、「ホテル」や「レストラン」などの会場タイプ等を選ぶことができました。しかし、これは初めから「ホテルがいい」と明確に基準が決まっている人にとっては探しやすいのですが、「この式場はアットホームでいいけど、この式場は堅苦しくてなんだか好みじゃないな」などといった漠然とした状態ではなかなかうまく検索できない状態でした。

解決策 そこで私達は図 8 のような画像を選んでもらうインターフェースを用意しました。従来型の検索クエリや条件は入力することが大変ですが、自分の好みのイメージに従って画像を選ぶのはより簡単で、カスタマーにとって楽しい体験となります。画像を選ぶとシステムが画像にあったテーマ（診断キーワード）をサジェストしてくれます。システムが代わりにイメージを言語化してくれることで、カスタマーが「私って



図 8: ゼクシフォトサーチの機能

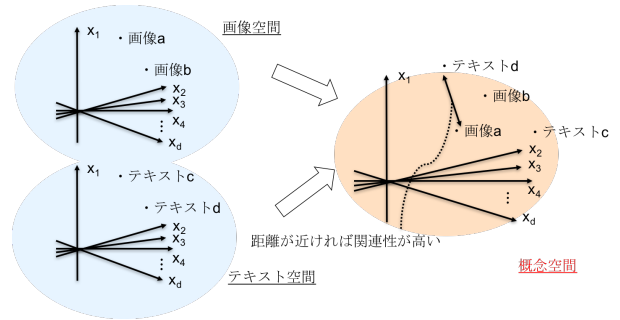


図 9: マルチモーダル学習による共通空間への射影

こんな式場を探していたのね」と納得感を得られるようになります。また、イメージにあったテーマが見つからない場合は類似する画像やシステムで提案される関連タグを選択しながら、自分にぴったりあったテーマに少しずつ近づいていけるようになっていきます。そして最終的に提案されたテーマを選ぶと、それにあった式場の一覧が出力され、式場のフェアを予約できるようになります。

技術 このシステムの実現にはマルチモーダル Deep Learning という技術を利用しています。画像とその画像に付与されたキャプションを大量にニューラルネットワークに学習させることで、画像とテキストの関係性をシステムが学習していきます。具体的には図 9 のように画像特徴空間とテキスト特徴空間から関連性が高いものが近くに来るような共通の埋め込み空間をシステムに学習させておき、画像から近いテキスト、テキストから近い画像をその空間上で検索することで画像とテキストの双方向的な検索を実現しています。

2.3 「探したけれど結果に納得感がない」を解決する仕組み

検索結果は正しくても、視覚的に入ってくる情報のマッチ度が低いと結果の納得感が低くなってしまいます。そうした結果の見せ方も画像認識技術を利用することで制御することができます。

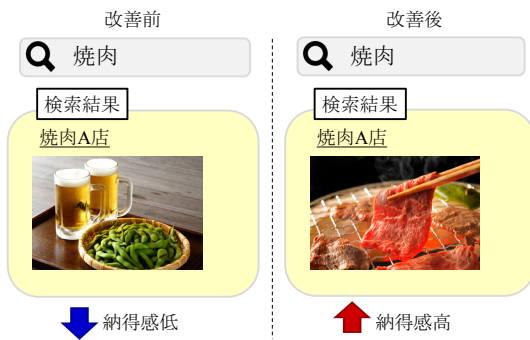


図 10: 料理画像タグ付け機能

2.3.1 ホットペッパーグルメにおける料理画像タグ付け機能

ホットペッパーグルメでの事例をご紹介します。ホットペッパーグルメは飲食店を検索したりネット予約することができる Web サイトです。

課題 カスタマーが焼肉が食べられる飲食店に行きたいと思って「焼肉」と検索した際に、店舗検索結果一覧に期待するのは美味しそうなお肉の画像でしょう。ですが別の画像が出てきてしまうと「検索された結果があまり合っていないのかな」と思ってしまい、店舗のページを見てもらえず離脱してしまいます。このように、検索結果は正しいにも関わらず、カスタマーにとって納得感がない状況になってしまっていました。

解決策 この状況を解決するために、検索されたキーワードに対して適切な画像が検索結果一覧に出るようにしました。このためには画像認識技術を利用して予め料理画像に料理名のタグを付けておき、検索した際に検索キーワードとタグを比較し、キーワードにマッチするような画像を優先的に表示するようにします。これにより期待はずれな画像が減り、食べたい料理の画像が出てくることで、より納得感のある検索結果となりました (図 10)。

技術 この機能を実現するために、まず画像に検索されやすいキーワードを調査しておき、画像認識で学習するタグにしました。料理画像に対して CNN によって画像とタグの組み合わせを学習しておき、定期的にバッチ処理を行うことで新たな料理画像に対してタグを付与するようにしました。そして、検索された際にそのタグが該当すれば、その画像を優先的に表示することにしました。



図 11: 飲食店かぶり画像検出

2.3.2 ホットペッパーグルメにおけるかぶり画像検出機能

ホットペッパーグルメにおけるもう一つの事例をご紹介します。

課題 飲食店を検索した際に検索結果の画像リストに複数の同じ画像が並んでしまうことによって様々なタイプの画像を見ることができなくなってしまうという課題がありました。ユーザーにとって画像は飲食店を選ぶ重要な基準となるので、ユーザーが見れる画像の種類が減ってしまうことは、飲食店を選ぶ際に判断材料が減ってしまうこととなります。

解決策 この課題に対して、類似画像検出技術により類似する画像を予め表示する画像リストから省いておくことで、似たような画像が並ぶ状況を排除しました。これによって、様々な画像を見ることができるようになり、飲食店の料理や雰囲気を検索結果のページからより理解できるようになりました (図 11)。

技術 aHash (Average Hash) というアルゴリズムを利用して画像をハッシュ値に変換した上で画像比較を行い、ある閾値を超えて類似しているものに対しては、かぶり画像として排除しました。

3 まとめ

これまで見ていただいたとおり、検索と一言で言っても様々な場面や、課題の改善方法があるということが分かっていただけたかと思います。画像認識技術はこのように検索における様々な課題を解決するために有益な手段であり、より人々の感覚に合った検索を実現する方法なのです。