

活動内容

東京理科大学経営学部ビジネスエコンミクス学科3年 桐島雅也

目次

1. 画像の差分を自動で検出するAndroidアプリのチーム開発
 - 1.1. 経験内容
 - 1.2. 期間
 - 1.3. 使用技術
2. 音声合成を用いた合成音声エンジンのチーム開発
 - 2.1. 経験内容
 - 2.2. 期間
 - 2.3. 使用技術
3. Signate主催のStudentCup2021秋にチームで参加し、銀メダル獲得
 - 3.1. 経験内容
 - 3.2. 期間
 - 3.3. 使用技術
 - 3.4. 関連リンク

1. 画像の差分を自動で検出するAndroidアプリのチーム開発

1.1. 経験内容

私は画像の差分を自動で検出するAndroidアプリのチーム開発を行った経験があります。具体的には、インフラ系の企業からの依頼で開発したアプリで、カメラで撮影した類似している2枚の画像を比較した際に、それらの画像の差分を自動で検出するものとなっています。つまり、間違い探しの答えを自動で探し出してくれるというようなものです。これは、クライアント企業の作業員の方々の業務を効率化することを目的として、開発に至りました。作業員の方々は作業前に、操作盤などのボタンの操作を行います。そして、作業後にこれらのボタンを作業前の状態に戻さなくてはなりません。これまでは、この作業を目視で確認して行っていたため、画像を用いて自動化することで、作業の効率化を図りました。

開発は2つのチームで行われ、私のチームは3人のメンバーで構成されていました。チームでの私の役割は、画像解析の技術を用いることで、画像から自動で差分を検出してくれるアルゴリズムの開発、改良、そしてアプリへの実装を担当しました。アルゴリズムの完成に至るまでは、技術を調査することで仮説を立て、過去の画像データを用いた検証による差分検出の精度と解析時間、メモリの測定、そしてその結果から技術を採用するかどうかの検討といったPDCAサイクルを回してきました。

この開発プロジェクトでは、実際に作業員の方々が使用しているAndroid端末のスペックの関係もあり、メモリや差分検出における精度と解析時間という面での制限が非常に厳しかったため、使用する技術の選定はかなり難しかったです。特に、仮説を立てる際は、精度と解析時間の向上はもちろん、メモリの制限を超えないかどうかにも着眼して仮説を立てました。採用した技術の一例としては、画像の局所ごとの差分計算とノイズ処理といったものがあります。これは差分を検出する際の細かな領域ごとの処理を行うために用いました。ノイズ処理では、オープニング処理や膨張処理、二値化といったような処理を用いることで、ノイズによる不要な差分検出の対策をしました。また、採用に至らなかった技術も多くあり、ARマーカを用いた射影変換もそのひとつです。射影変換は、操作盤やディスプレイなどの対象物の2枚の画像をマッチングさせ、対象物を正面に向かせるために必要な技術でした。そこで、ARマーカを用いて射影変換をすることを検討しました。理由としては、ARマーカを用いることで、射影変換を行う解析時間とメモリの大きな削減につながると考えたためです。しかし、ORBというアルゴリズムを用いて特徴量を抽出することで射影変換を行なった場合の方が、解析時間やメモリの面では多少劣っていても、射影変換の精度という面で優れていたため、こちらの技術を採用することにしました。このような制限下でのアルゴリズムの開発の過程は、技術的な知識やコーディング力はもちろん、新しい技術への探究心やその取り組み方といった面でも大きく成長につながりました。

また、この開発プロジェクトでは、約10ヶ月間チームでの活動をしてきました。そのため、チームでのコミュニケーションの取り方の重要性も実感しました。特に、自分では苦手であったり、わからなかったりするようなことは、無理して自分でやろうとせず、チームメンバーに頼ることも非常に重要であると学びました。私は他人に頼らず、ひとりで取り組んでしまおうとする点が弱みとしてあるため、それによってチーム全体での進捗を停滞させてしまったこともありました。このような経験から、なんでもひとりで取り組もうとするのではなく、時にはチームメンバーに頼ることも、チーム全体での成果を上げるうえでは重要であると学ぶことができました。

1.2. 期間

2021年6月～2022年3月

1.3. 使用技術

- OpenCV
- Python
- Kotlin

2. 音声合成を用いた合成音声エンジンのチーム開発

2.1. 経験内容

私は音声合成技術を用いた合成音声エンジンのチーム開発に取り組んでいます。これは、人の独特な言葉選びや話し方を再現することを目的として取り組んでいます。この合成音声エンジンは、ある文章を入力すると、その文章をその人がどのような文章に変換して読むのかを予測し、文章を変換します。そして、変換された文章に対してその人の音声合成することで、話し方を再現して出力するというものです。

開発は3つのチームで行っており、私のチームは4人のメンバーで構成されています。私のチームでは、文章に音声合成するといった音声合成の分野を担当しています。開発を行うにあたって、論文などの文献を参考に音声合成モデルについて調査を行い、実用性の高いモデルをいくつか選定します。そして、選定したいくつかのモデルを実際に音声データを用いて学習させ、パラメータチューニングなどを行うことで、評価・比較し、優れたモデルの開発を行っています。

現在もプロジェクトは進行中で、今後はマルチスピーカーモデルを用いて、テンションを分けて話せるようなモデルの開発なども検討しています。そのための準備として、音声データのテンションを分類する取り組みもしています。そこで、音声データから抽出した多次元の特徴量データを主成分分析することで次元削減し、K-means法を用いたクラスタリングをしています。このように、合成音声エンジンを開発していく過程で、音声合成の技術のみならず、データ分析に用いるさまざまな技術も活用しており、幅広い知識を学ぶことができています。

2.2. 期間

2021年12月～2023年3月(現在進行中)

2.3. 使用技術

- Python
- Pytorch

3. Signate主催のStudentCup2021秋にチームで参加し、銀メダル獲得

3.1. 経験内容

私はSignateが開催するコンペティションであるStudentCup2021秋にチームで参加し、銀メダルを獲得した経験があります。友人と自分達の実力を試し、成長していきたいと考え、このコンペティションに参加しました。このコンペティションでは、シェアサイクルサービスの各ステーションの自転車数を予測することで、各ステーションの自転車数に過不足が生じないように、夜間に自転車の再配備を行うという対策を講じることを目的として分析を行いました。分析の際は、モデルの学習データとして、全70ステーションについて、2013/09/01～2015/08/31の730日分、それぞれの日について1時間おきのデータや、日時ごとの天候のデータを用いました。

モデルの精度を向上していくために特に工夫した点は、実際のシェアサイクルユーザーの視点に立って考えることで、必要なデータの選定や加工を行ったことです。一例としては、雨が降っていると自転車に乗ることは躊躇うと考えたため、自転車数の予測において降水量は必要なデータであると考えたことです。それだけでなく、さらに工夫した点として、直感的に、0.2、0.3mmくらいの雨であれば、0みたいなものではないかとも考え、降水量を四捨五入するという思い切った工夫もしました。結果的には、この工夫が功を奏し、最終的な精度向上につながり、銀メダルを獲得することができました。この工夫で精度が向上した理由ははっきりと理解できていませんが、モデルの学習における複雑化を避けれたことや、人間の直感的な判断が影響していたためであると考察しています。そして、この降水量を四捨五入するといった工夫は、論理的だけでなく直感的な視点も持って考えることができたからこそ

の発想であると考えています。そのため、この経験を通して、論理的でなく直感的でクリエイティブな視点でも、データを活用していくことが重要であると考えようになりました。

3.2. 期間

2021年9月～2021年10月

3.3. 使用技術

- Python

3.4. 関連リンク

- <https://signate.jp/users/68784>