過去卒論のまとめ（永林さん）

# アブストラクト

　テレビcmカテゴリのクラス分けシステムの開発について述べる。今日人々は日常的にテレビを見ている。だからテレビcmを調査することは社会分析に役に立つ。なぜならテレビcmは人々や文化に影響を与える体。それで、CNNを使ったテレビcm分類システムの開発を決めた。話は変わるが、CNNを構成する活性化関数について紹介したいと思う。開発において、システムを様々な活性化関数を用いて評価する。この研究において、ReLUとその派生を活性化関数として用いたとき、良い正答率が得られた。

# イントロ

近年、ほとんどの人々がテレビを所有しCMを見ている。テレビcmは様々な商品やサービスに関する情報を私たちに提供する。そして視聴者の購入意欲に影響する。テレビcmを調査することは社会分析に役立つ。なぜならテレビcmはその当時の社会の状況やトレンドを反映するものだからだ。テレビcmを調査するために、テレビcmをカテゴリーを用いて分類することは大切なことだ。

　自動的にテレビcmを分類するシステムを開発することを決めた。そのシステムはCNNを使用されている。開発を決めた理由は、人間がテレビcmを分類するのは非効率だからである。CNNは近頃頻繁に使用されている。多くの場合、画像が一つ一つCNNに入力される。この研究では入力として、ビデオをいくつかの連続した画像としてCNNを応用する。OOO。また、より良いシステムを開発するために、認識精度の観点からCNNに用いられる様々な活性化関数を評価する。さまざまな活性化関数はこれまでCNNでの使用を提案されてきた。しかしながら、どの活性化関数もそれぞれいくつか問題がある。それゆえ、CNN

モデルを開発するときのために最適な活性化関数を調査することは重要なことだ。

# ニューラルネットワーク

　ニューラルネットワークは人間の脳と神経システムを模したコンピューティング知るテムで、文字認識や言語認識などのパターン認識の分野においてしばしば用いられる。ニューラルネットワークは、入力層と隠れ層と出力層によって成り立っている。複数の隠れ層を持つニューラルネットワークを持つ機械学習はとりわけ、深層学習と呼ばれる。ニューラルネットワークの一例はFig.1で示されている。このFig.1のような環はニューロンと呼ばれる。それは重みとバイアスをしようして計算され、その値が次のニューロンに渡される。パラメータの調節は最終的なニューラルネットワークの出力結果を開発することを可能にする。

# CNN

CNNは深層学習のニューラルネットワークの一つだ。それは畳み込み層とプーリング層によって成る。ニューラルネットワークはしばしば画像認識の分野において用いられる。畳み込み層においてCNNは入力データを畳み込みによって計算し、部分部分の特徴を得る。この方法によって計算されたデータはフィーチャーマップと呼ばれる。プーリング層は基本的に畳み込み層の後で、入力情報を圧縮しより扱いやすい形式に変換する。プーリング層は“平均プーリング”や“最大プーリング”などといった種類がある。この研究においては“最大プーリング”を使用した。“最大プーリング”は入力の最大値を出力する。

# OpenCV TensorFlow

OpenCVは動画や画像を要約するオープンソースライブラリの一つである。CやC＋＋、PythonやJavaなど様々な言語においてサポートされている。この研究において、動画や画像の処理に使用された。

　TensorFlowもオープンソースライブラリの一つである。このライブラリは機械学習の開発にしばしば使用される。このライブラリはGoogleによって開発されたものである。このライブラリもCやC++、またはPythonなど様々な言語に対してサポートがある。この研究において、ニューラルネットワークの開発に使用された。

# 活性化関数

活性化関数はニューロンの出力を決定づける関数である。例えば、Fig.1のようなニューラルネットワークの出力層のニューロンの計算は、Fig.2によって示される。ここでx1, x2 などは前の層の出力で、w1,w2などは重み、bはバイアスである。ニューロンによる出力は活性化関数の適用によって派出される。活性化関数をh()とすると、ニューロンによる出力はh(y)として与えられる。様々な活性化関数がこれまで提案されてきた。この研究では７つの活性化関数を使用してみた。

# シグモイド関数とハイパボリックタンジェント関数

　シグモイド関数は、[0, 1]の地域をとる。シグモイド関数は下のように表現される。活性化関数は原点を通る方が良いとされているが、それを改善したのがFig.4のようなハイパボリックタンジェント関数である。ハイパボリックタンジェント関数は、[-1, 1]の地域をとる。ハイパボリックタンジェント関数は下のように表される。これまで、シグモイド関数とハイパボリック関数はしばしば使用されてきた。しかしながら最近ではReLUという関数がよく使用されるようになった。その理由は、ReLUはバニっシンググラディエント問題に善処したからである。

# ReLU

　ReLUの値域は[0, ♾)である。入力が負の値の場合０を出力する。入力が正の値の場合、その値を出力する。多くの研究において、ReLUを実装したものは最も良い認識精度を記録した。しかしながら、入力が負の値の場合いつも０を出力することは問題である。いくつかの活性化関数はこの問題に対して改善策を提案してきた。

# ReLUの派生系

　ReLUの問題を解決するために、負の引数をとるいくつかの関数が提案されてきた。それらは下にあるような、ELU, SELU, Leaky ReLU そしてReLU6で、この研究において使用されている。

　ReLUと比較してSELUは負の値をとる。SELU関数は以下のように表現される。λとαの値の例として、λ＝1.0607とα＝1.6732のように代入した。

　ELUは、SELUのとりわけλとαがそれぞれどちらも1の場合の時のことをいう。ELU関数は以下のように表現される。

ReLUと比較して、Leaky ReLUも負の値をとる。Leaky ReLUは下のように表現される。ここで、α＝0.5である。RELU6は最大値が６でこれ以上は増えない関数である。ReLU6は下のように表現される

# 実験

## 前準備

　CNNモデルを訓練するために、245ものの様々かつ複数のテレビcmを収集した。またそれらは実時間１５秒のもので、それらのみがこの研究において使用される。それらのCMを”food”, “car”, “cosmetic”, “other”によって分別した。とりわけ”other”は、他３種類に分別できないもののことである。またその内訳は、foodと分別されたものが103種類、carと分別されたものが49種類、cosmeticと分別されたものが41 種類、otherと分別されたものが52 種類である。

　同様にCNNモデルの認識能力を測るテストのためのデータとして40種類のCMを収集した。その内訳はそれぞれ１０種類ずつとなっている。

## CNNの構造

　任意のCMを４種類（”food”, “car”, cosmetic”, “other”）に分別する意図がある機能を持ったCNNを含むモデルを開発した。それは大きく4つ、入力層、３D畳み込み層、全結合層、出力層から成る。入力層において、１つのCMから取り出した３０枚一組の画像を使用した（cm動画１つは１５秒なので、およそ0.5秒毎にサンプリングした）。すべてのCMのオリジナルサイズは1920x1080である（それは放送メディアに依る）。それらを研究環境の都合上、80x45にサイズを変換した。３D畳み込み層は４つの層から成り、それぞれにプーリング層が適応される。プーリング層として“最大プーリング”を適応した。この研究においては、３D畳み込み層の活性化関数を変えることによって活性化関数を評価する。出力層は最終的に４つ中のどのカテゴリかを決定する。出力層の活性化関数としてはソフトマックス関数を使用した。学習率は0.00001にしてある。学習は学習中の識別率が98%になったときに終了する。

# 結果と考察

この実験によって、CNNをもつモデルを先に挙げた７つの活性化関数によって学習させた。まずCNNモデルは、CMを４種類に分別できるように開発され、学習した。その結果は下のテーブル１に示される。結果から、テストにおける識別率のものはおおよそ46.8%で、訓練時のものと比較して半分以下であった。

この低い識別率の原因を調査するために、種類数に関して４以外の数字でも分類してみる必要があると判断した。それを実現するため、”food”, “car”の二種類のみに分別するCNNモデルを開発した。その後CMを分類し識別率を計測した。テーブル２はそれぞれの活性化関数に関するテストデータに対する識別率と、平均値である。シグモイド関数の結果に関していえば、平均値に対してその識別率は15%から20%低かった。それはつまり、シグモイド関数は７つの活性化関数の中ではるかに識別率が低いことを示した。シグモイド関数以外の活性化関数の結果に関して言えば、ほぼ同じ値であった。

学習では識別率は９８％まで上昇した。しかしながら、テストの際の識別率は全体的に低いものとなった。例えば、４種類の分別に関しての平均値は46.8%である。これは半分よりも低い結果である。その理由は、モデルの過学習にあるのだと考えている。そのため、トレーニングの際の最大識別率を変更し、過学習が起きるかどうか確かめた。その結果をテーブル３とFig.12で示した。小さなズレはあるが、テストでの識別率と学習時での識別率では正の相関がみられる。つまり、過学習は起きていないことがわかる。

# 結論と未来

　この研究ではCMを分類し、様々な活性化関数を評価した。その結果は、シグモイド関数が他の活性化関数と比較して、とりわけ識別率が低いことを示した。また、他の活性化関数に関して言えば大体同じであることもわかった。ReLUの派生系の関数は、ReLUの識別率とほとんど同じであった。どの活性化関数を使うのがベストかは、状況によるものである。しかしながら現状、ReLUまたはその派生系のどれかを使用するのがベストだと言える。

　将来的な研究に関して言えば、トレーニングデータを増やすことでより良いシステムを作るべきだと思う。またデータの情報量を増やし、学習率を上げ、隠れ層のパラメータを増やし、ニューラルネットワーク全体の層を増やすことも大切だ。さらに転移学習を応用することで、入力データをより効率的に使用することも念頭におきたい。