

Over the AI ― AIの向こう側に(22):

官能の人工知能 ～深層学習を最も分かりやすく説明するパラダイム

<http://eetimes.jp/ee/articles/1806/04/news030.html>

深層学習を正しく理解するのは困難を極めます。ですが、あるパラダイムで考えると、大変に分かりやすくなるのです。そのパラダイムとは、これまであらゆるテクノロジーの進化と発展をけん引してきたと言っても過言ではない、最も偉大なるコンテンツ―そう、「エロ」です。

2018年06月04日 11時30分 更新

[江端智一, EE Times Japan]



今、ちまたをにぎわせているAI(人工知能)。しかしAIは、特に新しい話題ではなく、何十年も前から隆盛と衰退を繰り返してきたテーマなのです。にもかかわらず、その実態は曖昧なまま……。本連載では、AIの栄枯盛衰を見てきた著者が、AIについてたっぴりと検証していきます。果たして”AIの彼方(かなた)”には、中堅主任研究員が夢見るような”知能”があるのでしょうか―。[⇒連載バックナンバー](#)

「深層学習」を理解するための2つの前段

【その1】

成人向けの雑誌の中でも特に性的な娯楽要素を扱う分野の書籍および雑誌が、いまだにコンビニなどで販売されていることについて、私は、ビジネスモデルの面(×倫理面、×教育面)から、不思議に思ってきました。

インターネットを使えば、ほぼ無限と言っても過言ではないほどのエロコンテンツにアクセスできる現代にあって、数々の障壁(女性店員(特に若くて好みのタイプ)がレジにいる場合は差し控える、わざわざ少し離れたコンビニにまで出向く、表紙を裏向きに向けてレジに出す、など)を乗り越えてまで、購入を試みるモチベーションが分からないのです。

ところが、最近、町内会の委員を担当して分かってきたことがあります。「わが国の、ITリテラシーは、私が想像するよりも、はるかに低いレベルに滞っている」ということです。

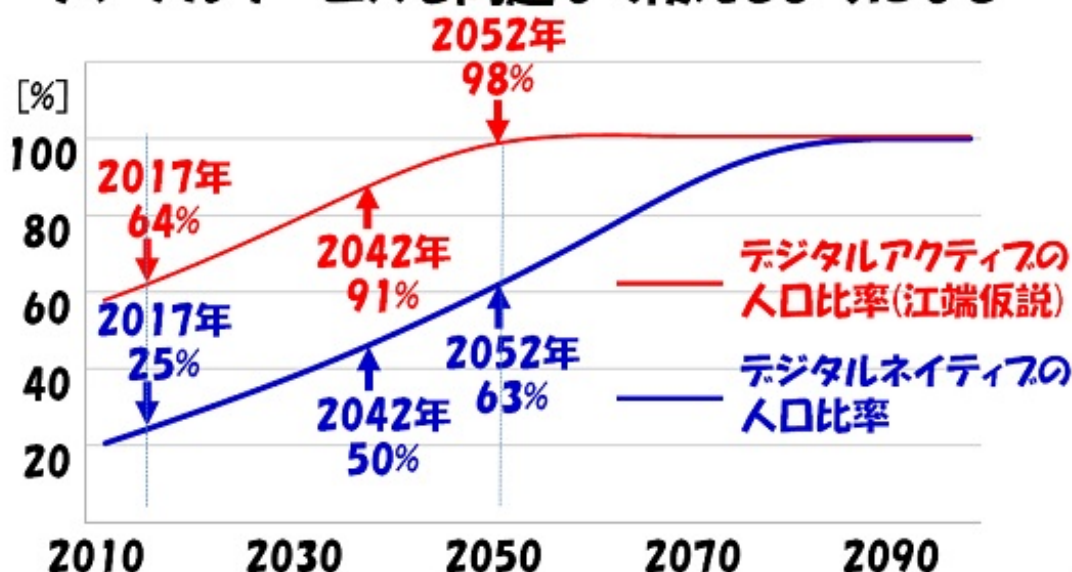
「PCやスマートフォン(スマホ)が使えない」「Webアクセスができない」「メールの文面が作

成できない」というレベルではなく、そもそも、メールアドレスを知らない、持っていないという人が、委員の中に、無視できない人数いることが分かったのです —— とりわけ多いのがリタイヤ後のシニア世代です。

調べているうちに、これを裏づける事実が山ほど分かってきました。「購買層の中心は高齢者」「販売されている半分以上が、人妻・熟女系エロ本」といった具合です。

再掲：デジタル利用者の比率

**2052年(35年後)には、人口の98%が
デジタルサービスを問題なく使えるようになる**



35年後にはアナログサービス消滅する・・・か？

「[力任せの人工知能 ～ パソコンの中に作る、私だけの「ワンダーランド」](#)」の図版を再掲

日本国内に、メール、パソコン、スマホが使えない人は、一定数存在しており(江端のシミュレーションでは、日本の人口の36%と推測)、そこから「エロ本は、もうしばらく生き残ることができる」という仮説は成立しそうです。

つまり、ITリテラシーが十分でない人にとって、コンビニで入手できるエロ本は、文字通りコンビニエンス(便利)なものなのです。

前段その2

しかし、私は、もう一つの仮説として、以下のようなことを考えています。

個人的なエロの嗜好性のマッチングに始まり、コンビニでの購入、その保管(例:ベッドの下)、そして廃棄(例:廃品回収日に、玄関の前にエロ本を積み上げる勇気(?))に至るまで、そこには、「エロ本の『ライフサイクル』としての様式美」があるのではないかと。

ネットのない時代においては、エロ本には、その本の価値だけでなく、その所有者に対するリスクがありました（エロ本を友人に貸せる人間は「勇者」で、借りる人間は「ヘタレ」）。

また、エロ本には、当時、画像としての制約（いわゆる、モザイクなどの画像処理）が加えられており、そこに一種の「制約としての官能」があったのです。

——それは、あたかも、俳句が5・7・5の語句の中に、一つの世界観をパッケージするのと同様の、情報圧縮技術です。

【その2】

少し前、嫁さんと長女（19歳）に質問をしました。

質問：ボーイズラブ（BL）は好きか？ 今、好きでないなら、今後、後発的に好きになれる可能性はあると思うか？

2人の解答は以下の通りでした。

嫁さん：「その可能性は絶無に近いと思う。もし、世の中で、BLが唯一の愛の形というのであれば、可能性はあるかもしれないが、私の世界はそうなってはいないから」

長女（19歳）：「私は愛に対してこれといった信念はないので、世の中のマジョリティーがBLの愛を受けいれるようになれば、私もそこに「流される」という自信はある」

江端：「ほう。ちなみにマジョリティーって、どれくらい？」

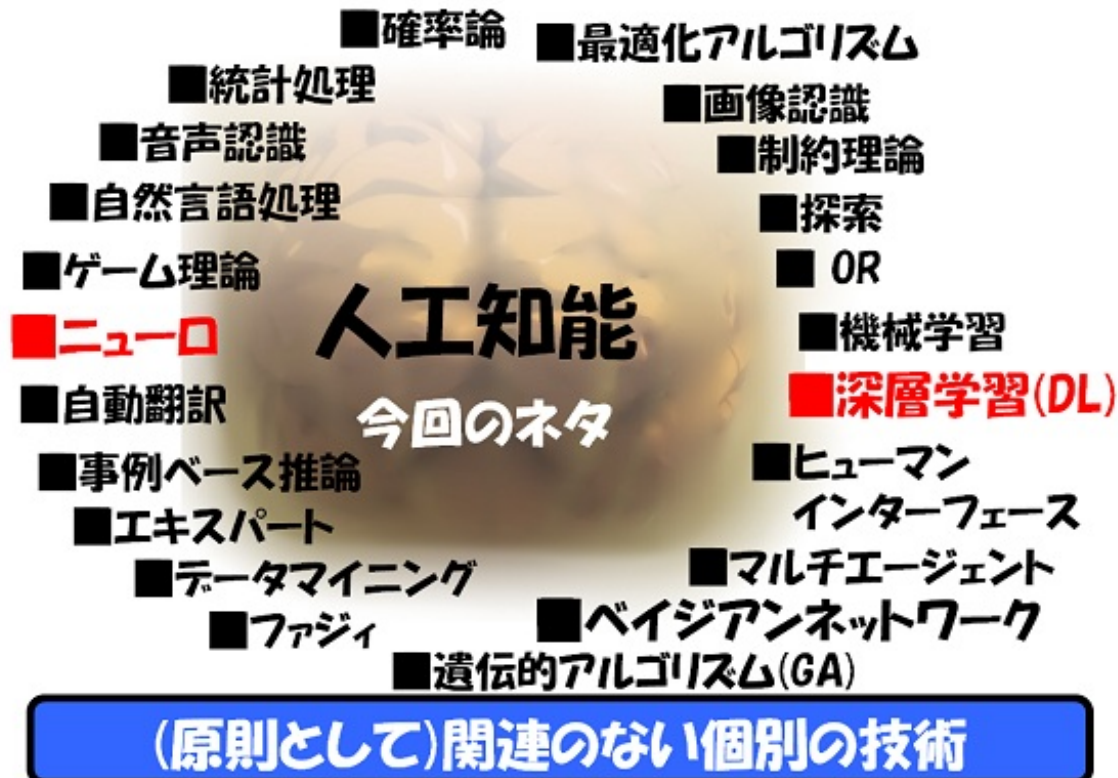
長女：「6割くらいかな」

□

ここまで読んできて、「一体何の話……？ ここってテクノロジーの記事が読めるサイトだよな？」と不安になってきた方、ご安心ください。大丈夫です。

これは、EE Times Japanが提供する、連載「Over the AI ——AIの向こう側に」の第22回で、今回のテーマはニューラルネットワークの、特に「深層学習（後半）」に関するものです。

“人工知能”という技術は存在しない



冒頭の2つの話は、今回の深層学習を理解する上で、非常に重要な示唆になっています。

なぜなら、今回私は、恐らくAI技術に関する解説としては、人類始の試みとなる、

—— 深層学習を、「エロ本」や「BL」のパラダイムから解説する

に挑戦してみたいと思っているからです。

江端を混迷の縁に追いやった、2つの写真

こんにちは、江端智一です。今回は前回に引き続き、現在進行中の第3次AIブームの最大のけん引技術である「ニューラルネットワーク」の、「深層学習(ディープラーニング)」についてお話したいと思います。

以前にもお話しましたが、私は、第2次AIブームの真只中であって、大学でAI研究のチームを立ち上げ(させられ)、ゼロからプログラムを書いて、ニューラルネットワークをテーマとして論文を執筆して大学を卒業した身の上です。

ですので、他の人より、かなり楽勝で、ニューラルネットワークの新しい学習方式である「深層学習」を理解できるだろう —— と、気楽に考えていました。

しかし、私は、この「深層学習」を"学習"するために、かなりの時間を費やしたのですが、こ

れが、どうにも理解できないのです。

私を混乱の縁に追いやったのが、どこでも頻出してくる、以下の2つの写真です。

本気で理解させようとしているのか？(1)

分からせたいのか？ 分っていることを誇示したいのか？

(ケース1) 猫



(ケース2) セレブ



どの解説もこの画像を「引用」しているだけ

ニューラルネットワークは、データを取り扱うことはできます。しかし、それは信号としての数値のみです。ニューラルネットワークそのものに、画像データを、画像データの形式のまま、各層に転送する能力などありません。

画像の上の写真は、まるでニューロの各層が画像を映像として転送しているかのような錯覚を与えますが — そんなこと、人間の脳だってやっていません（人間の脳のシナプスも、単なるニューロンの発火信号をリレーしているだけです）

ところが、一部の学術系の論文やプレゼンテーションの資料を除き、明らかに、その事実を放置して、誤った記載をしています。

もっとも、理解していなくても、結果的にそれが正しい内容になっていれば問題はありません。しかし、これらの記事はニューラルネットワークの事実を歪めている内容となっています。

さらに問題なのが、以下の写真です。

本気で理解させようとしているのか？(2)

本当は、「分からせる気がない」のではないか？



First Layer Representation Second Layer Representation Third Layer Representation

「中間層に、この特徴が現われる」という言い方は
ぶっちゃけ「詐欺」じゃないか？」

「どうせ説明しても分からないんだろう？」
と思っているのか？

この写真が掲載されている記事を読んでいると、「ニューラルネットワークが、中間層にこのような特徴が現われる画像を、自動的に抽出している」というように読めます。

うそです。

これは、畳み込みネットワーク (Convolutional Neural Network : CNN) が、このような特徴を出るような画像を取り出す前処理の説明に使っているものです。CNNでは、小さい画像ピクセルの集合から、そのピクセルの集合の纏めとなる情報を纏めて、さらに次の層に渡しています(後述します)。

そういう処理をしているのですから、最初の層で「部分」の画像が出てきて、最後の層で「全体」の画像がでてくるのは当たり前です。そういう風に作っているのですから。ニューラルネットワークが、自動的に部分から全体の画像特徴を抽出している訳ではありません。

にもかかわらず、これを「ニューラルネットが自動的に部分から全体への自動抽出を実行している」と記載している記事が散見されます。ただ、仕方ないともいえるかもしれません。なにしろ私自身も、つい先日までそう信じていたくらいですから。

私は、自分でCNNのサンプルプログラムをプログラミングするまで、このうそを見破ることができませんでした。

上の写真は、特定の画像を与えた時に、最終層でニューロンが発火した時に、それを逆方向から探して、どの「前処理をした画像」に反応をしていたのかを調べて、該当するたくさんの画像を、単に並べただけのものです(後述します)。

深層学習を知らなくても、地球は回るけど

プレスの記者が、聞きかじった話だけで分かったような気になり、「うそ」をまき散らしていることについては、私は、遠慮なく批判します。「お前たちの勉強不足が、深層学習についてのデタラメを世界にまき散らしている」と私は、声高に叫びます*)。

*) 昔、私は、「日本の経済」のとある分野についての新聞を出版している会社の記者にインタビューを受けたことがあるのですが、やつら、終始、態度がタカビで、不愉快な言動で、さらに私の提出した資料を、完全に「デタラメでうその内容」に書き換えて、新聞に発表しました。私、死ぬまで(死んでも)あの記者を許しません。

一方、研究開発やエンジニアの人の解説資料や、プレゼンテーション資料については、かなりの同情の余地があると思っています。彼らは、「誤差の逆伝搬学習を知らない人」までも想定してプレゼン資料を作成することが、不可能に近いことを知っているからです。

なぜ、私が、そう思うことができたのか ―― それは、今回の、深層学習のコラムで、私自身が「どう説明したら良いのか、さっぱり分からん」と、混迷の蟻地獄に陥ったからです。

こここのところ、毎日、夢の中で七転八倒しているうちに、

――別に、深層学習なんて知らなくたって、それでも地球は回るよね

という、ちょっと投げやりな気持ちになってきました*)。

*) 私の嫁さんの名言に、「リトマス試験紙が、青になろうが赤になろうが、それが一体何だというの?」があります(関連記事:「[6万人のメイドが“合体”!? EtherCATの通信方式](#)」)。

とはいえ、それでは、この連載の目的(数値や数式を使わないAI技術の解説)をひっくり返すことになります(そして、担当のMさんを大慌てさせることになります)ので、なんとか今回もがんばっていきたいと思います。

「教師あり学習」「教師なし学習」って、結局何なのか

ニューラルネットワークに限らず、どのAI技術においても、「教師あり学習」と「教師なし学習」は、いろいろなところで登場し、かつ、混乱して使われています。そこで、再度整理してみたいと思います。

いろいろなニューラルネットワーク

「教師あり」と「教師なし」の2つがある、と覚えていけば良い

タイプ	名称	意味	期待される動き
教師あり	全結合	入力と出力がペアになったらデータが大量にある	別の新しいデータを入力したら、 それらしい出力データ を吐き出す
	CNN		
	RNN		
教師なし	ボルツマン	入力データ だけ が大量にある	別の新しいデータが、 入力データに似ているか否か を吐き出す
	AE		

「教師なし」は「学習データがない」ではない

教師あり学習とは、どんな猫の写真を入力しようとも"1000"を出力するように、どんな犬の写真を入力しようとも"0100"を出力するように、どんな猿の写真を入力しても"0010"を出力するように、どんな雉(きじ)の写真しようとも"0001"を出力するように、力づくでニューラルネットワークを強制的に矯正するものです(前回ご説明した通り)。

学習データは「猫1、1000」「犬1、0100」「猿1、0010」「雉1、0001」「猫2、1000」「犬、0100」「猿2、0010」「雉2、0001」「猫3、1000」「犬3、0100」「猿3、0010」「雉3、0001」…のように、4種類の動物と出力値のペアとなります。

比して教師なし学習とは「どんな猫の写真」を入力しようとも、出力が"1"となるように、力づくでニューラルネットワークを強制的に矯正するものです。

学習データは、「猫1」「猫2」「猫3」「猫4」「猫5」「猫6」「猫7」「猫8」「猫9」「猫A」「猫B」「猫C」「猫D」「猫E」「猫F」「猫10」……と、ひたすら猫の写真だけです(出力値はいらない)。

「教師なし学習」は、「データの必要がない」ということではありません。データは山ほど必要です。ただ、『もし、このデータが与えられたら、〇〇をしる』の「〇〇をしる」の部分がないだけなのです。

手作り学習データでCNNを学んでみる

今回は、この教師あり学習としてCNN(Convolutional Neural Network: 畳み込みニューラルネットワーク)を、教師なし学習としてAE(Auto Encoder: 自己符号化)をご紹介します。

今回は、CNNとAEだけに限定する

この2つ(の違い)が分かっているれば、
それ以外は、知らなくても「なんとかなる」(多分)

通称	名称	江端なりの理解
CNN	Convolutional Neural Network	■ニューラルネットワークの学習が面倒なく運ぶように、画像を前処理するもの →学習の成功率を劇的に上げる
	畳み込みニューラルネットワーク	
AE	Auto Encoder	■ある種類の絵を見せたら、興奮するように強制的に学習させるもの →同じ種類の絵で興奮し、別の種類の絵で無反応するようにする
	自己符号化器	

“研究者”にとっては、失神しそうなくらい凄い技術だが・・・それ以外の人にはどうかな？

ニューラルネットワークには、いろいろありますが、前回の(1)階層型の層間全結合のニューラルネットワークと、この(2)CNNと(3)AEの3つさえ分かっているれば、おおむねニューラルネットワークは理解しているといっても良いでしょう。

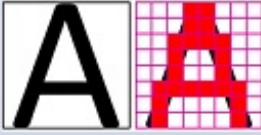
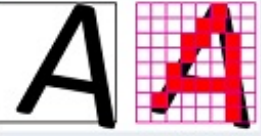
では、CNNから解説します。

CNNについては、いろいろな文章を読みまくったのですが、どうにもよく分からなかったのので、書籍やネットに記載されているプログラムをパクリ……もとい、参照させて頂きながら、自分でプログラムを書いてみました(プログラムは[こちら](#))。

学習データとしては、A,B,C,Dの4種類の文字を8つ用意しました。いくつかは文字の角度や大きさや場所を変えてみました。

プログラムで理解するCNN

文章を何度読んでも分からなかったので、プログラムを書いて、学習データも手作りしてみました。

"A"の学習データ	"B"の学習データ	"C"の学習データ	"D"の学習データ
			
			

単純な層間全結合型のニューラルネットワークの場合、大抵、学習に失敗する

前回、私は、XORという、線形では分離できない、簡単な4つの信号の学習を試みて、かなり頻繁な確立で学習に失敗することを示しました。

今回の入力データは、 $8 \times 8 = 64$ ビットのデータを8つも学習させることになりますので、前回のような全結合型のニューラルネットワークを使って、逆伝搬学習(バックプロパゲーション)を試みれば、ほぼ確実に失敗することは分かっています。

文字認識は、ニューラルネットワークの代表アプリケーションと言えるもので、世の中の多くの論文で成功例の報告がありますが、大抵の場合、マグレで学習に成功したケースです(いわゆる、チャンピオンデータ、というやつです)。

まあ、マグレであれ何であれ、一度学習に成功してしまえば、そのニューラルネットワークはどこでも使うことができるようになりますので、それはそれで良いのです。——が、結局のところ、マグレや設計者のセンスに依存するような技術は、結局使いものにならないのです。第2次AIブームのニューラルネットワークは廃れていったのは、こうした理由からでした。

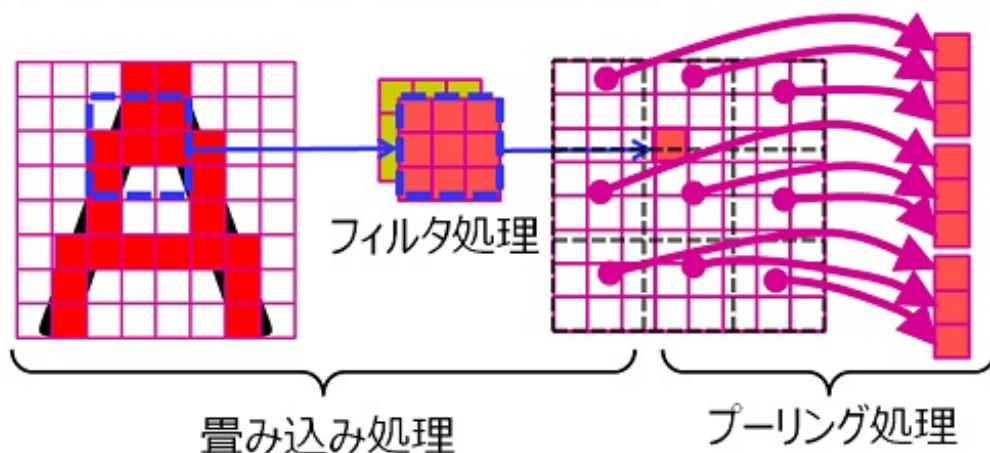
なんかちょっとずるい!? 「CNN」は「学習しない」がキモ

さて、ここからは、実際のプログラムの内容に合わせて説明します。

CNNとは、Convolutional Neural Network、畳み込みネットワークのことです。

CNNとは、「学習しない」がキモ

その時、江端は、「え～、それってズルいじゃん!」と叫んだ



**別にこの部分は学習しているのではなく、
前処理をしているだけ**

「畳み込み処理」とは、文字通り「布団を畳む」と同じ意味です。要するに、畳んで小さくする、ということです。しかし、小さくても布団は布団です。布団の性質が損なわれる訳ではありません。

前述の私のプログラムでは、画像の3x3(9ビット)の部分を1x1(1ビット)に畳んでいる訳です(9畳の部屋を1畳に圧縮しているイメージ)。ただ、単に畳み込むのではなく、畳み込む前に、フィルターにかけます。

フィルターとは、簡単に言えば、画像の一部をあえて、見えなくしてしまうことです。このフィルターには、タイプの違う複数の種類を準備しておきます。フィルターを通した画像は、当然に不完全な画像になりますが、その特徴が際立つことになります。

その不完全な画像を、頭の中で(妄想で)組み立てることができれば——完全なエロ本の写真……ではなく、原画像として理解できるはずです(後述します)。

次に「プーリング処理」ですが、畳み込み処理で作られたデータの塊の中の代表値に置き換えるものです(私のプログラムでは3x3の9ビットのデータの最大値を使っています)。そして、これも基本的には「エロ画像のモザイク」と同じ理解で良いのです(後述します)。

今回、私が驚いたのは、この「畳み込み処理」「フィルター処理」「プーリング処理」が、画像処理の前処理にすぎない、ということでした。形の上では、ニューラルネットワークの層の処理をしているように見えますが、要するに画像の繰り返し処理を、for ~ nextのように行っているだけなのです——つまり、この部分に関しては、私が、大学時代に地獄を見てきたバックプロパゲーション(逆伝搬学習)は、ここには全く登場してこないのです。

私は、「深層学習」という名前を聞いた時から、

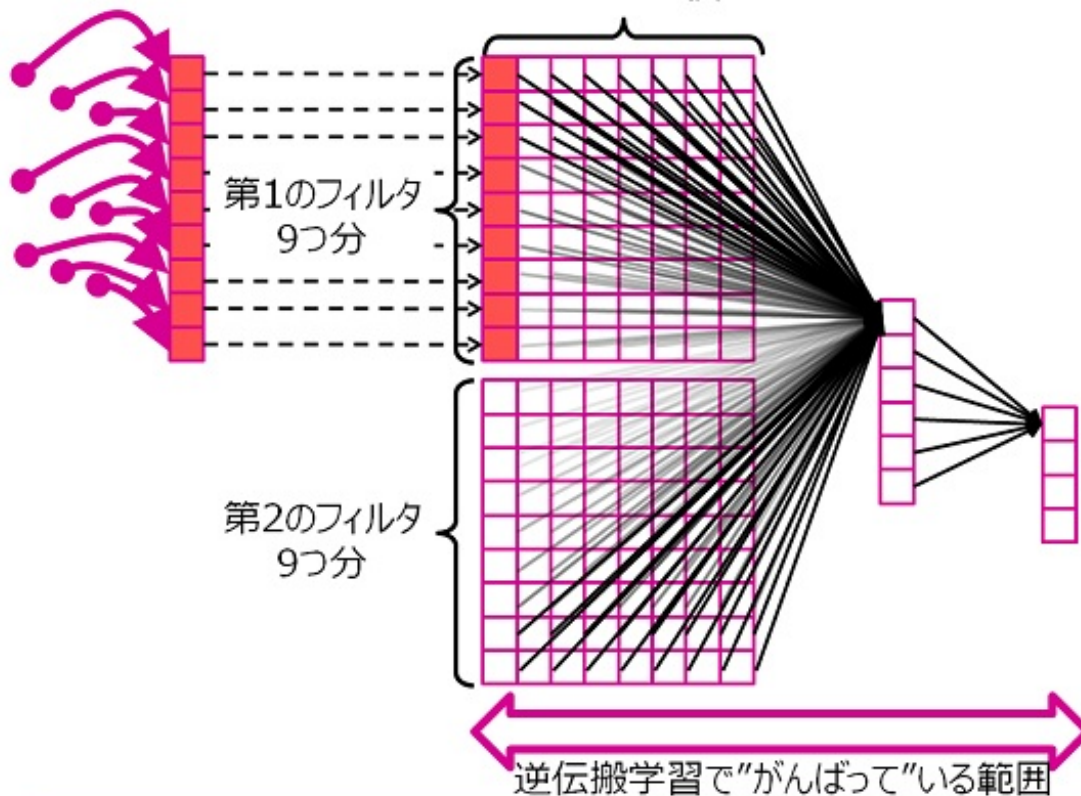
—— 奈落の底（深層）にあるニューロンの出力層から、地表の入力層に向けて逆走する、ダイナミックなバックプロパゲーションの驀進が見られると思ってワクワクしていた

のです。

ですから、このコードを理解した時 —— 返して！ この私の純粋なエンジニア心の期待を返して!! と叫びそうになりました（通勤電車の中だったので、我慢しましたが）。

CNNは最後の2層だけ学習する

江端:「これを「深層 **学習**」と言われると、なんかムカつく」
学習データ8個分



前処理で、学習収束しやすいデータとなるように加工を完了させてしまっている

とはいえ、バックプロパゲーションのニューラルネットワークが全く登場しないという訳ではありません。最後の3層のみは、バックプロパゲーションが必要となります。

いかに「モザイク」の画像処理でデータ数の削減を図ろうとも、たかだか8x8ビットの画像（

文字)のデータ8パターンでも144ビット(9×8×2)の加工済みデータが出てくるのです。これを最終的に、4ビット(1000, 0100, 0010, 0001)にまで圧縮するには、全結合型のニューラルネットワークに登場してもらわざるを得ません。

全ての試みで学習に成功

私のプログラムには、入力層144ノード、中間層6ノード、最終層4ノードという、私の人生では試したことがない巨大なニューラルネットワークが登場します(私は、これまで、3桁もの数のノードを持つニューラルネットワークを試したことがありません。なぜなら、1桁でも失敗の日々だったからです)。

学習の結果ですが ―― 腰が抜けるほど驚きました。

100回以上、初期乱数のシードを変えて、シナプス結合値を変化させて試してみたのですが、学習成功率100% ―― 全ての試みで学習に成功しました。

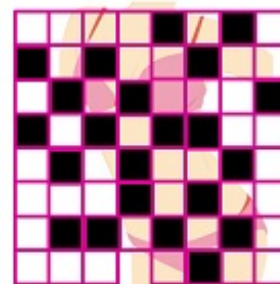
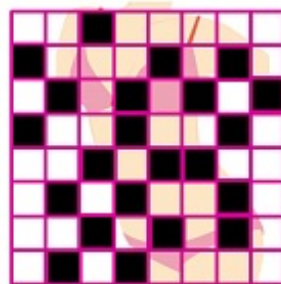
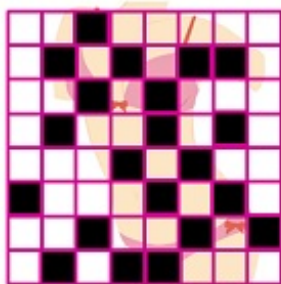
『何これ? 入力層のノード144個だよ? そんなこと、本当にあり得る?』 ―― 信じられませんでした。そして、「なるほど、世界が大騒ぎをするはずだ」と納得したのです。

これはニューラルネットワークというよりは、画像の前処理技術(画像データを、ニューラルネットワークに食べやすいデータに加工した技術)の勝利と言えましょう。私は今、CNNの開発者に対して、心からの最大級の賛辞を申し上げたいです。

なぜ、CNNがこのような、劇的な成功に至ることできたのか ―― 私は、ここに「エロ写真のモザイク加工処理のコンセプト」を主張したいのです(開発者はきっと怒るでしょうが)。

CNNのフィルタとは「モザイク」である

「興奮」…ではなく「特徴」が強化されるという解釈でいい?



「完全な情報」に、敢えてフィルタリングをかけるという、逆転の発想

本来、画像フィルタとは、「縦方向」とか「横方向」にスリット状にするなど、恣意的に作るも

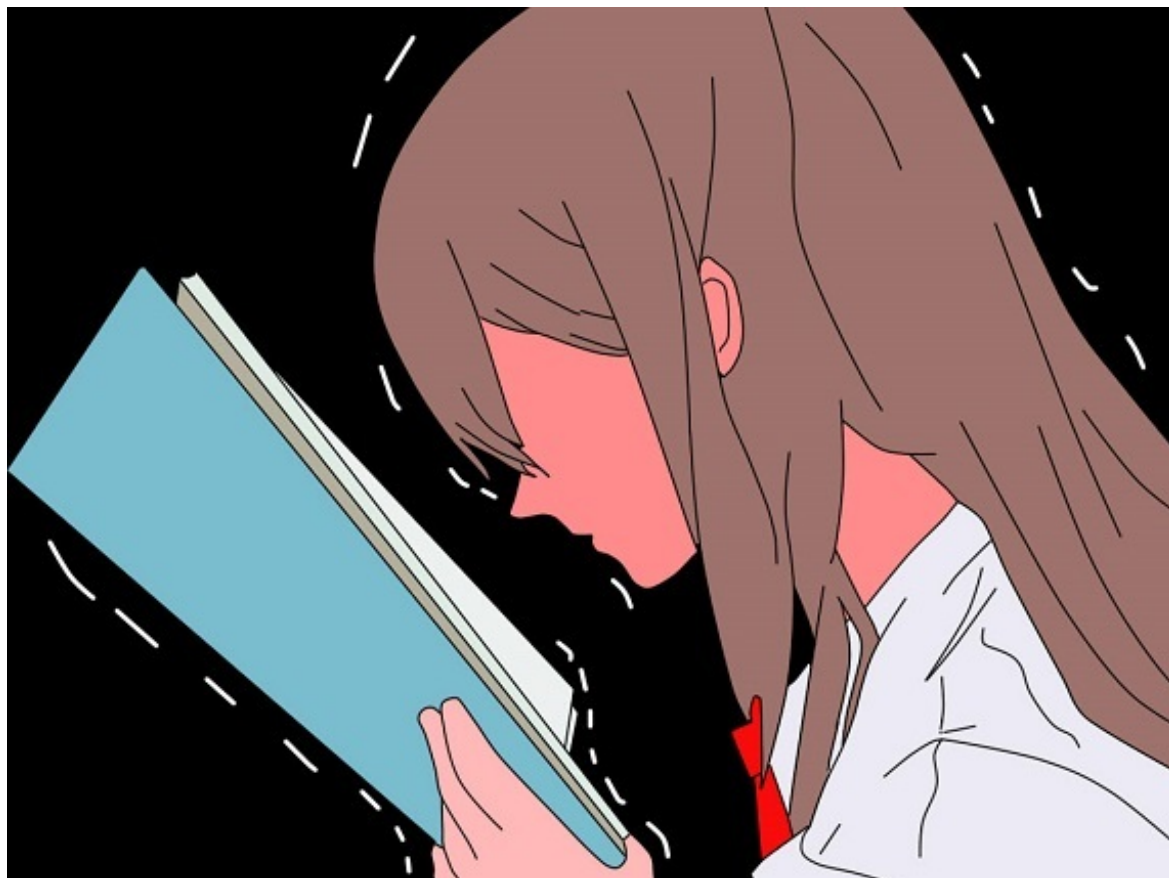
のですが、私のプログラムではフィルターは乱数で適当に作られていました（当初、私は、この理由が分かりませんでした（ソースコードをコピー&ペーストしていたからです））。

しかし、「これは、エロ写真のモザイクだ」と気がついた瞬間、CNNの各種のコードの意味が、私の頭の中で一気に氷解したのです*）。

*)「畳み込み」はちょっと説明できませんが、「フィルタリング」と「プーリング」については、十分に解説可能です（このコラムが発禁になるので、具体的な解説は差し控えます）。

要するに、細部の描写を曖昧にすることで――まさに、映倫（映画倫理機構）の狙い通りに――画像を「ぼかす」のです。そして、「私たち」……ではなく、「CNN」が、「細部の情報に心を奪われないように」……ではなくて、「細部の情報を想像（妄想）で補完できるように」データの加工を行う訳です。

「全体としての画像の概要を開示しつつ、部分の画像に対して、あえて『ぼかす』『隠す』『（見えそうで）見えなくする』という処理を施す」――これが、人類の、特にティーンエイジャの脳の「官能をつかさどる」ニューロンの発生とシナプスの活性化に、どれほど資してきたか、今さら皆さんに説明する必要もないでしょう。



「ぼかして補う」。官能の世界とCNNの世界は似ている？

ただ、勘違いされては困るのですが、「CNNとは、ニューラルネットワークに、画像に対する官能感知機能が具備したもの」ではありません。間違っても、そのような宣伝をしないようにしてください（一部のプレスが、また変な記事を書きかねないので、念のため）。

私は、「CNNとは、ニューラルネットワークと好結合する、優れた画像のデータ加工の前処理技術を具備している」と申し上げたいのです。

猫を見たら興奮するニューラルネットワークを作る

では次に、学習なしの深層学習である、AEについて解説をしたいと思います。

AEについては、一般論にすると分かりにくくなるので、今回は「猫の写真」に寄せて解説します。

AEの目的を一言で申し上げます。

—— 猫を見たら興奮するニューラルネットワークを作る、です。

「猫に興奮する変質者を作るのか？」と誤解されては困ります（編集担当のMさんを不用意に激怒させるだけ*）ですし、そもそも、そういうことではないからです）。

＊）編集注：担当Mは猫好きです。犬も好きですが、やはり面白いのは猫だと思います。大して要らない情報ですが、飼い猫がこの5月で20歳を迎えました……！ パチパチパチ……！

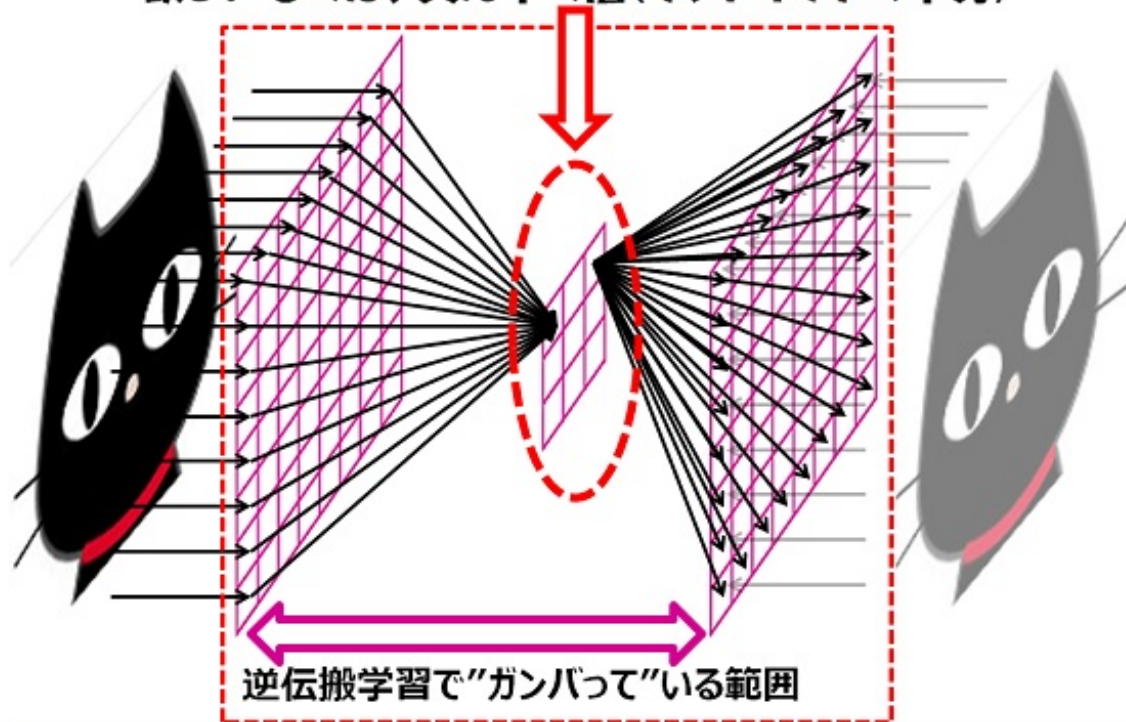
つまり、このニューラルネットワークには「猫」しか見せず、世界には「猫」以外には何も存在しないと考えるニューラルネットワークを具現化するものなのです。

さらに「猫だけに反応しろ」と教えることすらもしません。ただ、ひたすら猫の写真を、—— どんな写真でもいいですが、とにかく「猫」の写真を見せ続けて、ニューラルネットワークの学習を続けます。この学習方法には、やはりバックプロパゲーションを使います。

AEでは、ちょっと変な学習の仕方をさせます。入力データと全く同じ出力を出すニューラルネットワークを、何個も作るのです。

ANとは、鏡に映った自分を学習する

欲しいものは、真ん中の層(サンドイッチの中身)



ネコの画像情報がサンドイッチの中身に凝縮

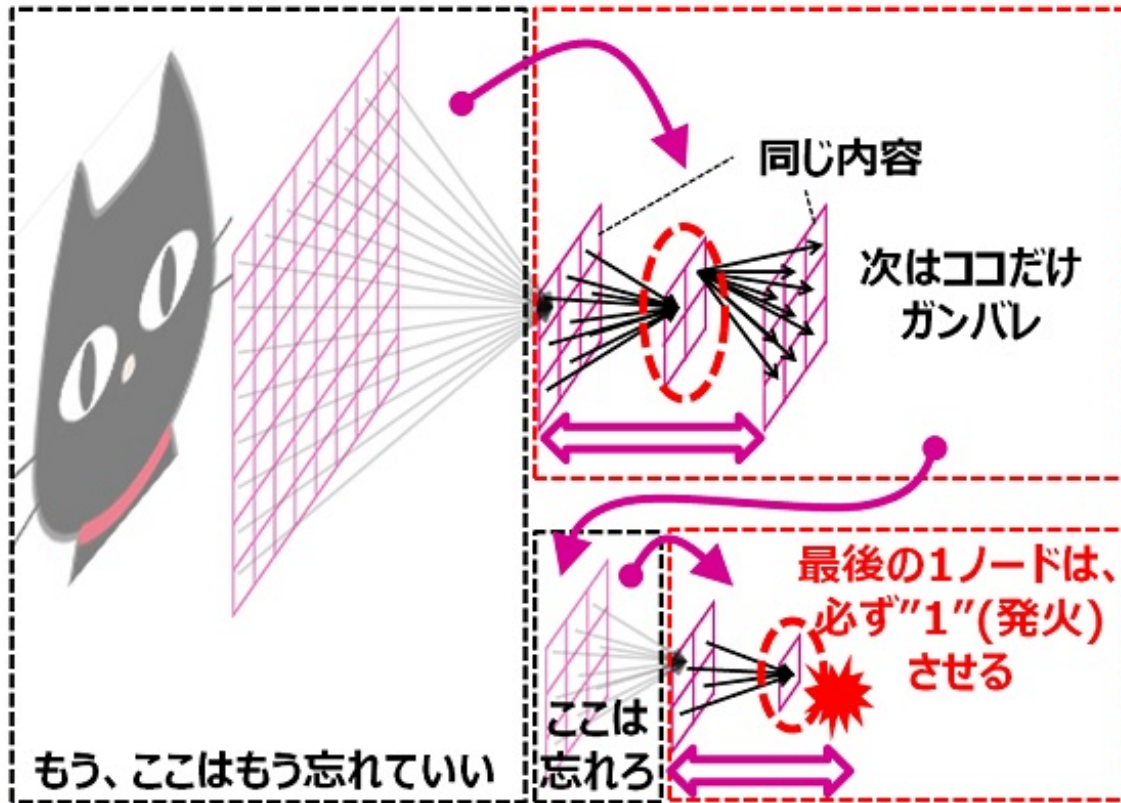
ポイントは、「入力層と出力層のノード数もデータも同じになるようにしますが、中間層のノードはそれより小さく設計する」という点にあります。

この中間層は、ちょうどサンドイッチの中身のように見えますので、ここからは、このパラダイムで説明します。つまり、サンドイッチの両側の食パンのサイズよりも小さい「具」を作ることが目的です。

このサンドイッチの具は、出力層で、入力層のデータを復活することができるものなのです。ということは、このサンドイッチの具は、入力層のデータを圧縮したものになっていることになります。

学習したもので、学習をする

サンドイッチの中身をパンにして、新しい中身を作る



欲しいのは、「どんどん小さくなっていく中身」

さて、今度は、このサンドイッチの具をパンに見立てて、新しい、そして、さらに小さいサンドイッチの具を、バックプロパゲーションを使って作り出します。これを何回も繰り返し、最後、サンドイッチの具は最終的に1ノード(これ以上圧縮できない情報)になります。

そして、この1ノードは、必ず、値が"1"となるようにバックプロパゲーションで学習させます。こうして、学習用データである猫の写真で、かならず「発火」……「興奮」するように"作る"のです。

この学習を、全ての猫の写真で実施します(Googleは1000万枚の猫の画像を使ったそうです)。

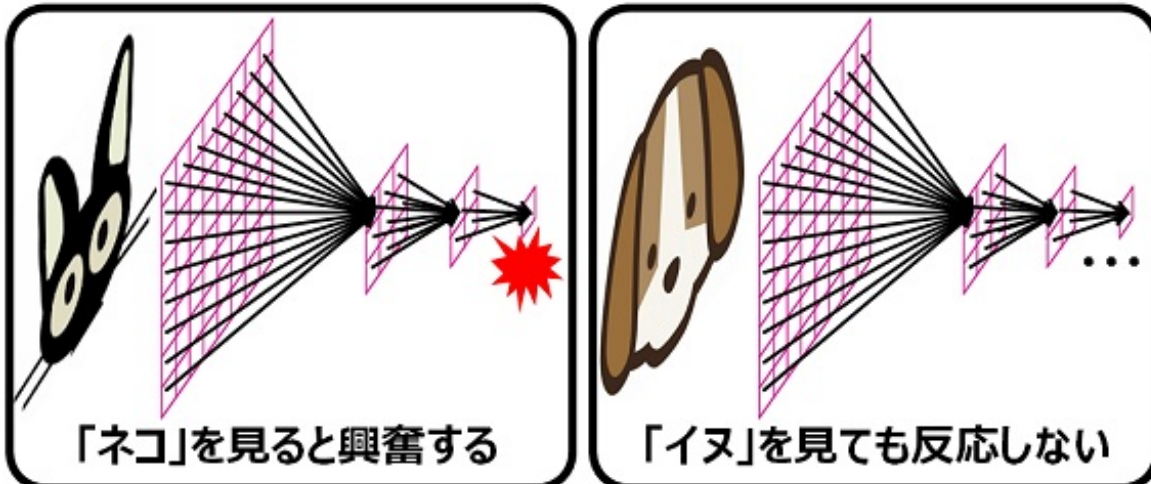
こうやって、3層に限定した対象型のニューラルネットワーク(サンドイッチの形をしたネットワーク)を、順次バラバラに学習させることで、各層のサンドイッチの具が出来上がります。多層ネットワークではなく、3層に限定して学習するので、学習の成功率は格段に向上することになります。

こうして、全ての猫の写真によって学習を続けて、サンドイッチの具を重ね合わせた、N段のケーキ型のサンドイッチのネットワークは、「猫を見たときだけ、最後の1ノードが発火する」("1"に

なる)、つまり「猫だけに興奮するニューラルネットワーク」として、完成することになります。

特定の“モノ”に反応する装置の完成

サンドイッチの中身(だけ)を積み上げた“N段のケーキ”



ネコにのみ反応するネットワークの完成!!
.....で？

このネットワークには、猫の写真しか見せておりません。そして、とにかく最後の1ノードが発火するように（興奮するように）、各層を丹念に学習さえ続ければ良いのです。まさに、「教師なし学習」の面目躍如と言えましょう。

AI技術のメカニズムを人間に反映させて考えてみる

さて、この「後発的な学習によって作られた興奮」が、冒頭の話、「ボーイズラブ(BL)は好きか？ 今、好きでないなら、今後、後発的に好きになれる可能性はあると思うか？」につながる訳です。

私は、嫁さんと長女に対して『私たちは、今まで存在しなかった対象に対するものに対して、学習によって後発的に「興奮」を作り出すことができるか』という、思考実験を仕掛けてみたのです。

嫁さんは「愛という概念のない世界に限定すれば、後発的な人工的な愛（興奮）の発生の可能性はある」といい、長女は「マジョリティーが是認する世界であれば、後発的な愛（興奮）が受入可能なプラットフォームは作れる」と言いました。

似たような話で、「結婚相手の多くは、(オフィス内などの)半径3m以内の異性である」という説があります(参考:[他社のサイトに移行します](#))。

これは、頻繁に視野に入る異性が、結婚の対象になる、という点では、AEに近いように思えるのですが、しかし、「猫」と同じ範ちゅうで考えるのであれば、別段、近くにしようとも遠くにしようとも、あまり関係がないとも言えそうです。

まあ、これは単なる思考実験ですので、あまり真剣に考える必要はないのですが、AI技術のメカニズムを、逆に人間の側に反映させて考える試みは、これは、これで、なかなか楽しいものなのです。

□

さて、ここから真面目な話ですが、CNNにしてもAEにしても、これが、すごい発想に基づく、素晴らしい効果を発揮するAI技術であることに疑いはありません。私は、これらの技術と、その開発者に跪く(ひざまずく)ことにやぶさかではありません。

しかし――です。

これって、一体どこに、どういう風に使えば良いのか、と考えると、首をかしげてしまうのです。

CNNとAEの応用例

CNNもAEも『凄い技術』であることは間違いない。だが・・・

通称	凄さ	江端の感想	江端の疑問
CNN	多層の同時/一斉学習をスッパリ止めた	画像認識に強い ――― というか、 画像処理に特化 しすぎている ぶっちゃけ、それ以外の「 凄い 」成功事例はない	凄いのは、学習ではなく、 画像の前処理の発想 だよな？
AE		とにかく、何も教えなくても、特定の“モノ” だけ に反応する ――― ぶっちゃけ、「猫」と「セレブ」くらいの成功例しかない	これって、「猫」「セレブ」以外に、 どこに、どうやって、どういう用途 に使えるの？

「これで、何でもできるようになる」
という感じはしない

CNNもAEも、既存のニューラルネットワークの技術を援用している点には変わりはありません

せん。

しかし、第2次AIブームでつまづいた、多層ネットワーク全体での学習をキッパリと諦め、逆に少層(3層)ネットワークにぶった切ってバラバラに学習するようにして、加えて、学習前の前処理に膨大な工夫を施すという、思いきったパラダイムシフトを行ったという点において、「すごい」の一言に尽きます。

特に、CNNは画像にはやたらめったら強いですが(そもそも、そういう目的に作られているからです)、それ以外にうまく援用できるのかというと、私が調べた限り、他分野での「すごい」成功事例はありません。

また、AEに至っても、「猫」と「セレブの顔」以外の顕著な応用例の報告が見当たりません。

私は、CNNもAEも、このような対象物の認識や判断ができるだけでも、十分に成功である、と思いますし、研究員としてまたエンジニアとして、私はその立場を支持します。

――しかし、「世界」はそう思ってくれるでしょうか？

- 誤ったAI技術の知識をまき散らし、それに対する修正も自己批判も一切行わず、
- アニメの美少女アンドロイドを、AI技術の最終形のように考え、
- AIに経営判断を全部丸投げして楽ができるようになると、夢のようなことを真面目信じている

そういう「世界」は、CNNやAEのようなすごい技術を着想し、開発を成し遂げた、研究者やエンジニアのことを、きちんと評価してくれているのでしょうか？

今回のまとめ

それでは、今回のコラムの内容をまとめてみたいと思います。

【1】今回は、ニューラルネットワークの後半として、深層学習について説明しました。特に、前代未聞の挑戦として、深層学習を、「エロ本」や「BL」のパラダイムから解説することを試みてみました。

【2】深層学習に関するプレスの説明は、表層的であるばかりではなく、誤った情報を世間に散布しているという事実と、研究者による説明であっても深層学習を完璧に説明するものにはなっていないという事実を明らかにしました。同時に、深層学習を解説することが、恐ろしく難しいということを、私の体験談から述べました。

【3】深層学習を説明するに際して、教師ありタイプである「畳み込みニューラルネットワーク(CNN)」と、教師なしタイプである「自己符号化器(AE)」の2つを紹介し、「教師あり／なし」と「学習データのある／なし」が混同されている事実を指摘し、そもそも、学習データがないニューラルネットワークなど存在しないことを説明しました。

【4】CNNが、「超多層のニューラルネットワークの逆伝搬学習を行っているものではなく」、最終層を除けば、画像データの加工処理を行っているだけであることを解説しました。しかし、それで

も入力層100ノードを超えるニューラルネットワークの学習成功率100%という、驚異の性能を発揮したことをご紹介しました。

さらに、CNNを、エロ本の写真の画像処理のパラダイムで考えると、極めて理解が進むことを、江端の体験から説明しました。

【5】AEは、とにかく特定の種類のデータ(「猫」「猫」「猫」「猫」「猫」「猫」「猫」「猫」……)を見せて、「猫」を見せるだけで興奮するネットワークを作ることであることを説明し、同じように、人間も後発的な学習によって「興奮」するメカニズムを作れるか否かについて、簡単な思考実験を試みました。

【6】CNNもAEも、ニューラルネットワーク技術の弱点である逆伝搬学習を逆手にとって、その周辺技術を磨き上げることで、第3次AIブームを牽引するに足る素晴らしい技術となったことを説明しました。しかし、それでもCNNもAEも今後の産業応用には、不安が残ることを述べました。

以上です。

□

以前、「[ご主人様とメイドは、通信の設計図を数秒で書く](#)」にも執筆しましたが、人類のライフスタイルを決定的に変えたイノベーション(活版印刷(かっぱんいんさつ)、写真、インターネットなど)が、エロによってけん引されたという事実は、よく知られています。

実際、私は、インターネットが「エロ」コンテンツによって ―― いや、正確に言うと「エロ」コンテンツだけで発展してきた経緯を、目の当たりにしてきた当事者です。

今回、私は、深層学習の雄である「畳み込みニューラルネットワーク(CNN)」の中に、「エロ」の基盤技術である「モザイク」が含まれていることを発見することができたのですが ――

このパラダイムにたどりついていた最初の一人が、「この私」であるはずがない

と思っています。

絶対に、私より先に、この考えに至った先行者はいたはずです。しかし、その先行者の方々が、その事実を、学会や研究会や論文で明示的に発表することはなかったのでしょうか。みすみすキャリアをドブに捨てることもないでしょうから。

まあ、それはさておき。

この発想を、逆方向から考えてみると、「全てのAI技術は、人間の本質の一つである『エロ』を、運命的に含んでいる」と考えるのは、それほどムチャな仮説であるとは思えません。

これまで私は、この連載を22回に至って続けてきたわけですが、私の新仮説「AI技術には『エロ』が包含される」に言及したのは、今回が初めてで ―― そして、今、私は猛烈な悔悟の念

に襲われています。

これまで連載してきた全てのAI技術を、「エロ」の観点から再検討すれば、AI技術の飛躍的發展に寄与する新しいパラダイムに至れたのではないかと思います。——そして、その貴重なチャンスを、2年近く見逃してきたことに地団駄(じだんだ)を踏む思いです(しかし、担当Mはホッとしております……)。

ですが、寂しいことに、本連載もそろそろ最終回が近づいてまいりました。

「エロ」の観点からのAI技術の包括的な検討については、後人に託したいと思います。どうか私の思いを受けとり、この検討を行って欲しいと、強く希望しております。

□

では、次回より、最終章「Over the AI -- AIの向こう側の私たち」について、お話ししたいと思います。

本連載に、どうか最後までお付き合い頂けますよう、よろしくお願い申し上げます。

⇒「Over the AI ——AIの向こう側に」[⇒連載バックナンバー](#)



Profile

江端智一(えばたともち)

日本の大手総合電機メーカーの主任研究員。1991年に入社。「サンマとサバ」を2種類のセンサーだけで判別するという電子レンジの食品自動判別アルゴリズムの発明を皮切りに、エンジン制御からネットワーク監視、無線ネットワーク、屋内GPS、鉄道システムまで幅広い分野の研究開発に携わる。

意外な視点から繰り出される特許発明には定評が高く、特許権に関して強いこだわりを持つ。特に熾烈(しれつ)を極めた海外特許庁との戦いにおいて、審査官を交代させるまで戦い抜いて特許査定を奪取した話は、今なお伝説として「本人」が語り継いでいる。共同研究のために赴任した米国での2年間の生活では、会話の1割の単語だけを拾って残りの9割を推測し、相手の言っている内容を理解しないで会話を強行するという希少な能力を獲得し、凱旋帰国。

私生活においては、辛辣(しんらつ)な切り口で語られるエッセイをWebサイト「[こぼれネット](#)」で発表し続け、カルト的なファンから圧倒的な支持を得ている。また週末には、LANを敷設するために自宅の庭に穴を掘り、侵入検知センサーを設置し、24時間体制のホームセキュリティシステムを構

築することを趣味としている。このシステムは現在も拡張を続けており、その完成形態は「本人」も知らない。

本連載の内容は、個人の意見および見解であり、所属する組織を代表したものではありません。

関連記事



[儲からない人工知能 ～AIの費用対効果の“落とし穴”](#)

AI技術に対する期待や報道の過熱が増す中、「抜け落ちている議論」があります。それが、AIの費用対効果です。政府にも少しは真剣に考えてほしいのですよ。例えば「荷物を倉庫に入れておいて」と人間に頼むコスト。そして、AI技術を搭載したロボットが「荷物を倉庫に入れておく」という指示を理解し、完璧にやり遂げるまでに掛かるトータルのコスト。一度でも本気で議論したことがありますか？



[力任せの人工知能 ～パソコンの中に作る、私だけの「ワンダーランド」](#)

私はこれまで、人口問題や電力問題、人身事故などさまざまな社会問題を理解するためにシミュレーションを利用してきました。シミュレーションは、AI（人工知能）という概念を飛び越えて、「人間が創造した神」と呼べるかもしれません。今回は、シミュレーションに最適なAI道具の1つとして、「オブジェクト指向プログラミング」を解説します。これは、PCの中に“私だけのワンダーランド”を力任せに作る技術ともいえます。



[外交する人工知能 ～理想的な国境を、超空間の中に作る](#)

今回取り上げる人工知能技術は、「サポートベクターマシン（SVM）」です。サポートベクターマシンがどんな技術なのかは、国境問題を使って考えると実に分かりやすくなります。そこで、「江端がお隣の半島に亡命した場合、“北”と“南”のどちらの国民になるのか」という想定の下、サポートベクターマシンを解説してみます。



[光による量子コンピュータ、現行技術で実現へ](#)

北海道大学と京都大学の研究グループは、現在の技術レベルでも光による量子コンピュータを実現できる方法を開発した。量子ビットの誤り耐性を従来の100億倍に高めることで可能とした。



[富士通のAIプロセッサ、演算精度とμアーキに工夫](#)

富士通は、同社のプライベートイベント「富士通フォーラム2018 東京」（2018年5月17～18日、東京国際フォーラム）で、ディープラーニング処理に特化したプロセッサ「DLU（Deep Learning Unit）」を展示した。競合となる既存のアクセラレーターと比較して、10倍の電力性能比を実現するとしている。



[リーマンショックも影響？“蔵出しFPGA”の真相を探る](#)

前回に続き、2017年発売ながらチップに開発した年を意味する「2009」と刻まれていたIntel製FPGA「Cyclone 10 LP」を取り上げる。さらに多くのCycloneシリーズ製品のチップを観察し、2009年に開発されたチップであるという確証を探しつつ、なぜ2017年の発売に至ったのかをあらためて考察していく。

Copyright © ITmedia, Inc. All Rights Reserved.

