

情報と心理

第4回

分散表現・ニューラルネット

担当: 森田純哉

j-morita@inf.shizuoka.ac.jp

ファイル置き場

- <https://moritalab.inf.shizuoka.ac.jp/sites/inf-and-mind2019/>

ブックマーク推奨

今日のスケジュール

- 記号表現と分散表現(20分)
- 先週の実習の解説・深層学習(40分)
- 実験(10分)
- 人間の視知覚(20分)
- 授業後アンケート(宿題)

これまでの要点と今後

1 (10/2)	【イントロダクション】シラバスの読み合わせ、情報システム・情報技術と人間の関係、講師紹介
2 (10/9)	【方法論と歴史】行動主義、チューリング、ダートマス会議、人工知能・認知心理学・認知科学の関係
3 (10/23)	【方法論と歴史】物理記号システム仮説、記号表現と分散表現
4 (10/30)	【分散表現】ニューラルネットによる知覚のモデル、誤差からの学習
5 (11/6)	【分散表現】深層学習の到達点、視覚的注意とニューラルネットの限界
6 (11/13)	【議論】概念駆動 v.s. データ駆動、トップダウン v.s. ボトムアップ、分散表現と記号表現のインタフェース
7 (11/20)	【記号表現】プロダクションシステムと積み木の問題 (IF-THENルールの理解)
8 (11/27)	【記号表現】プロダクションシステムと滑車の問題 (パターンマッチの理解)
9 (12/4)	【記号表現】プロダクションシステムと記憶のモデル (意味ネット、フレーム、チャンク、系列位置曲線)
10 (12/11)	【記号表現】類推のモデルと情報処理的人間観 (述語論理式、グラフマッチング)
11 (12/18)	【認知アーキテクチャ】認知の統合理論、ユーザビリティ評価のモデル、ACT-Rのデモ
12 (1/8)	【認知アーキテクチャと身体】ACT-Rと脳機能の対応、生理と認知の統合による感情の説明、情報技術への応用
13 (1/15)	【認知アーキテクチャと社会】コミュニケーションシステムの形成実験
14 (1/22)	【認知アーキテクチャと社会】コミュニケーションシステムの形成シミュレーション
15 (1/29)	【まとめ】情報処理的人間観の確立
16 (1/29)	試験

- 「情報と心理」の歴史→
科学としての人工知能の追求
(認知科学の一部)

人間の認知のモデル化

人間の意識・思考・概念がコンピュータによってどう**表現**されるか

2つの認知のモデル

歴史的には対立

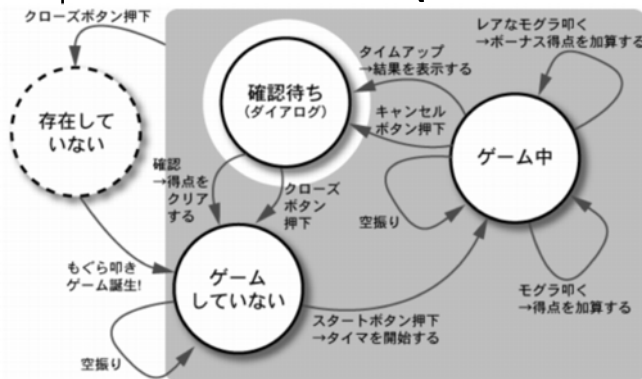
記号表現は時代遅れ ⇔ 分散表現じゃ何もわからない

記号表現

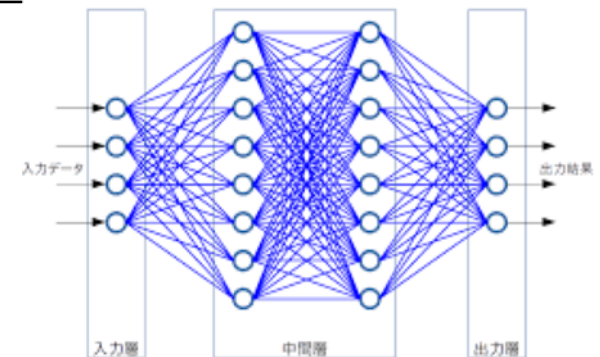
- 意識の流れ(記号の系列)
- 主観的(意識的)な認知プロセス
- 思考や推論のモデル
- 第一次人工知能ブーム(1950年代)、第二次人工知能ブーム(1980年代)

分散表現

- 意識の発生(記号の発生)
- 意識されない認知プロセス
- 知覚のモデル
- 第一次ニューロブーム(1950年代)、第二次ニューロブーム(1980年代)、**第三次人工知能ブーム(2010年代)**



2つの名称は文脈によって様々
トップダウン vs. ボトムアップ,
記号処理 vs. 並列分散処理...



分散表現

過去

マカロック・ピッツ
による人工ニューロンモデル 1943

ローゼンブラッドによる
パーセプトロン 1958

ラメルハート、マクレランド、ヒントンらによる
誤差逆伝播法 1986

チューリング
マシン 1936

ダートマス会議
1956

ニューエルとサイモンによる
一般問題解決器
1957

ミンスキーによる
パーセプトロン批判
1969

エキスパートシステムの
商用化 1980年代

ヒントンによる深層学習
の技術革新 2006

深層学習による
画像認識精度の飛躍的
向上 2012

現在

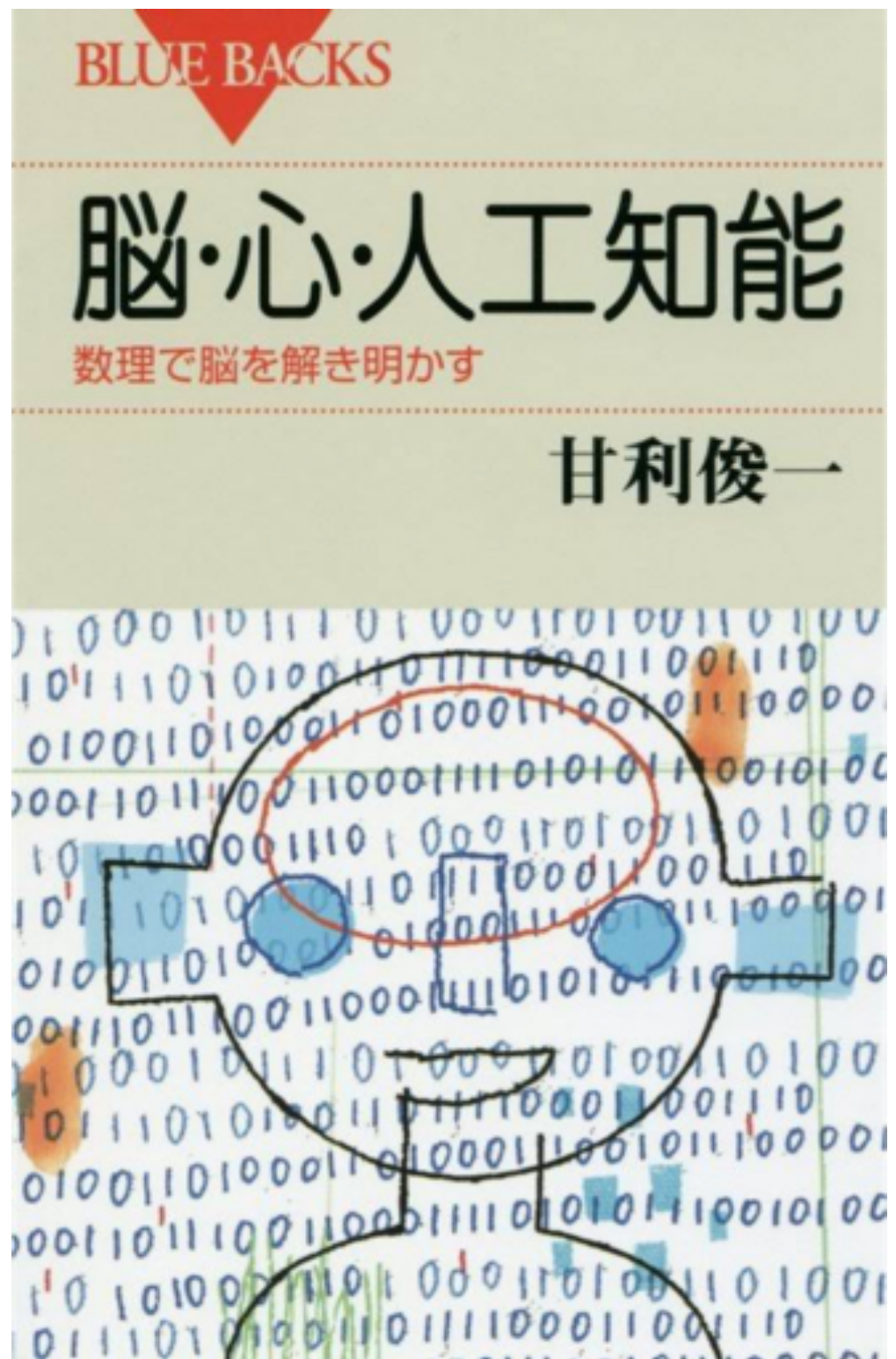
未来

記号表現

- 記号表現(AI)のブーム
- 分散表現(ニューロ)のブーム

より詳しくは

日本の計算神経科学の
開拓者
分散表現サイドからみた
「情報と心理」の歴史



近年のCognitive Science Societyでも

Keynote: Nature versus Nurture in Artificial Intelligence

Matthew M. Botvinick

27th July 2018 9:00 – 10:00 CDT

<https://cogsci.tv/category/cogsci2018/>

Password: cogscitv

CogSci 2018 - Keynote Matthew Botvinick

Watch later Share

学習するもの
ボトムアップ
= 分散表現

Nature and Nurture in AI

与えられるもの
トップダウン
= 記号表現

Matthew Botvinick
DeepMind, London UK
Gatsby Computational Neuroscience Unit, UCL

MONOHA TERRACE

認知科学・人工知能における永遠のテーマ

今年の認知科学会でも



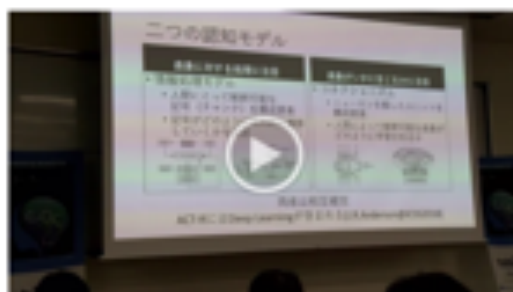
2019年9月4日（金）

プログラム委員会企画シンポジウム

深層学習時代に認知科学の歴史と価値を見つめなおす

主観を扱う科学としての認知モデリングの未来

企画：重塚博幸（北海道大学）、小野哲雄（北海道大学）、高橋英之（大阪大学）、森田純哉（静岡大学）



概要説明

森田純哉（静岡大学）

本シンポジウムは、深層学習が様々なタスクにおいて人間を凌駕するパフォーマンスをたたき出す現代において、あらためて認知科学ができることは何か、今後、どのような方向を目指すべきかを議論する。深層学習・ニューラルネットワークを用いつつ人間性の本質に迫ろうとする立場、神経科学や認知科学の伝統を受けつつ主観の科学を追求する立場などが、それぞれの考えを交流させつつ、新たな知の統合科学の姿と課題を語り出す。

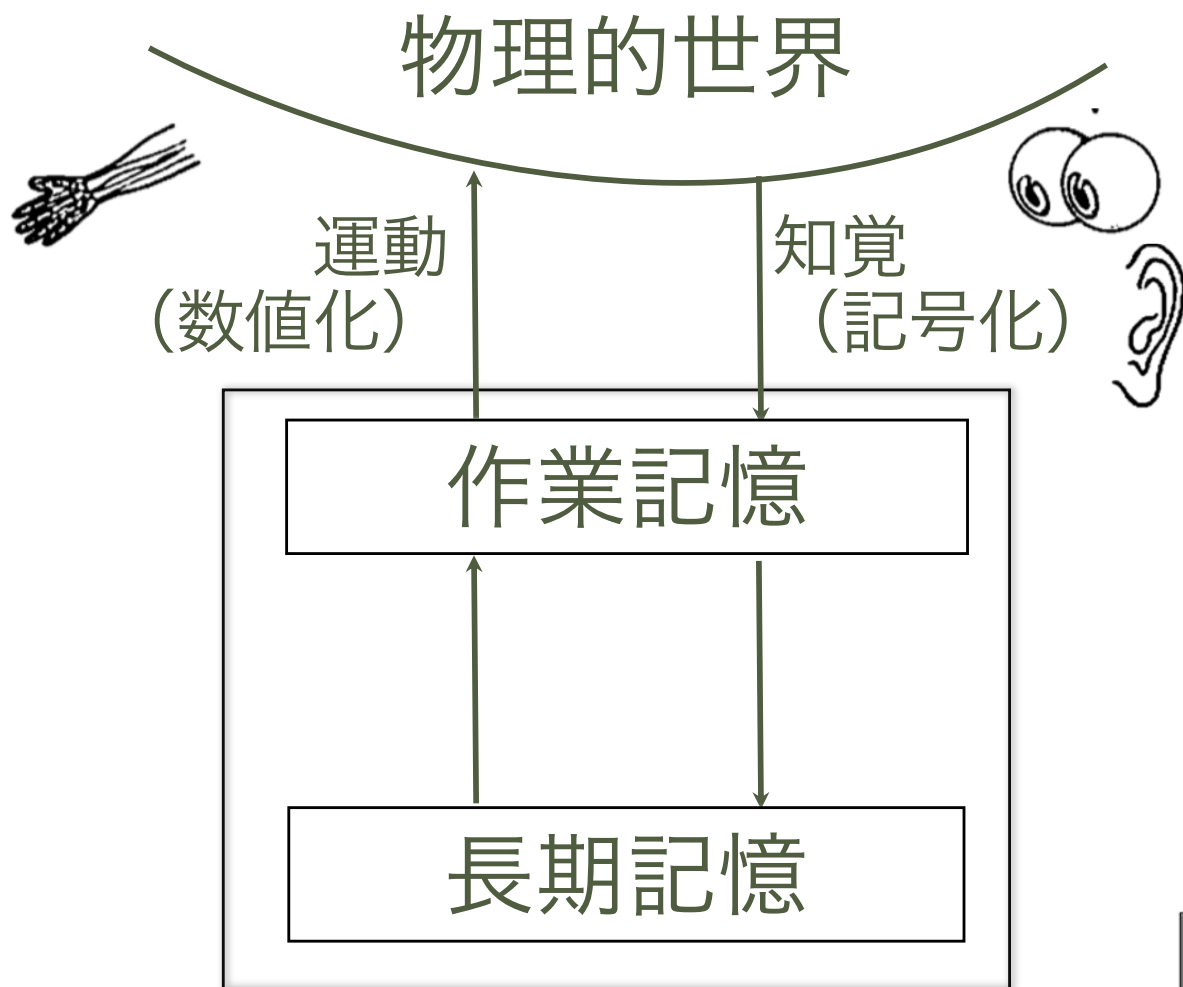


認知科学分野のDynamic Capability

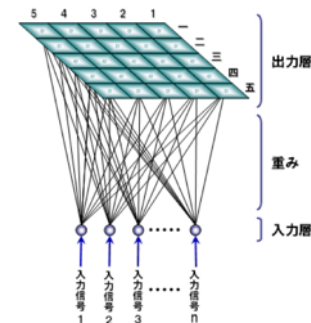
山川 宏（全脳アーキテクチャ・イニシアティブ）

今後の授業

以下の図(アーキテクチャ)の具体化



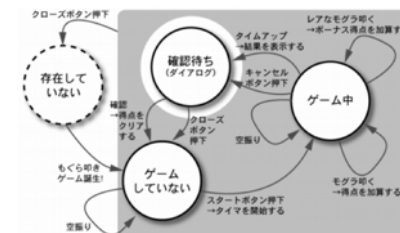
分散表現モデル
ニューラルネット



記号表現モデル
状態遷移
記号論理

記号と分散のハイブリッドで
人間の認知を説明

```
- (Cell object1)  
- (Pulley object2)  
- (Pulley object4)  
- (Weight object6)  
- (Weight object7)  
- (Rope object3)  
- (Rope object5)  
- (Rope object8)  
- (Rope object9)  
- (Rope object10)  
- (Value object1 Emp)  
- (Value object6 100)  
- (Value object7 Emp)  
- (Value object3 Emp)  
- (Value object5 Emp)  
- (Value object8 Emp)  
- (Value object9 Emp)  
- (Value object10 Emp)
```



入力装置としての知覚

物理的世界から記号的世界へ

分散表現を知るための実習

1. 以下のサイトにアクセスする

– <https://moritalab.inf.shizuoka.ac.jp/sites/inf-and-mind2019/nndemo/#3>

• オリジナル: http://jsdo.it/z_kro/u84J

2. 適当に数字を選び、Recognizeボタンを押してみる

A rectangular button with a light blue border and the text "recognize" in a dark blue font.

3. パラメータを変えつつtrainしてみる

A rounded rectangular button with a light gray border and the text "train" in a dark gray font.

– 以下のフォームの問いに答える

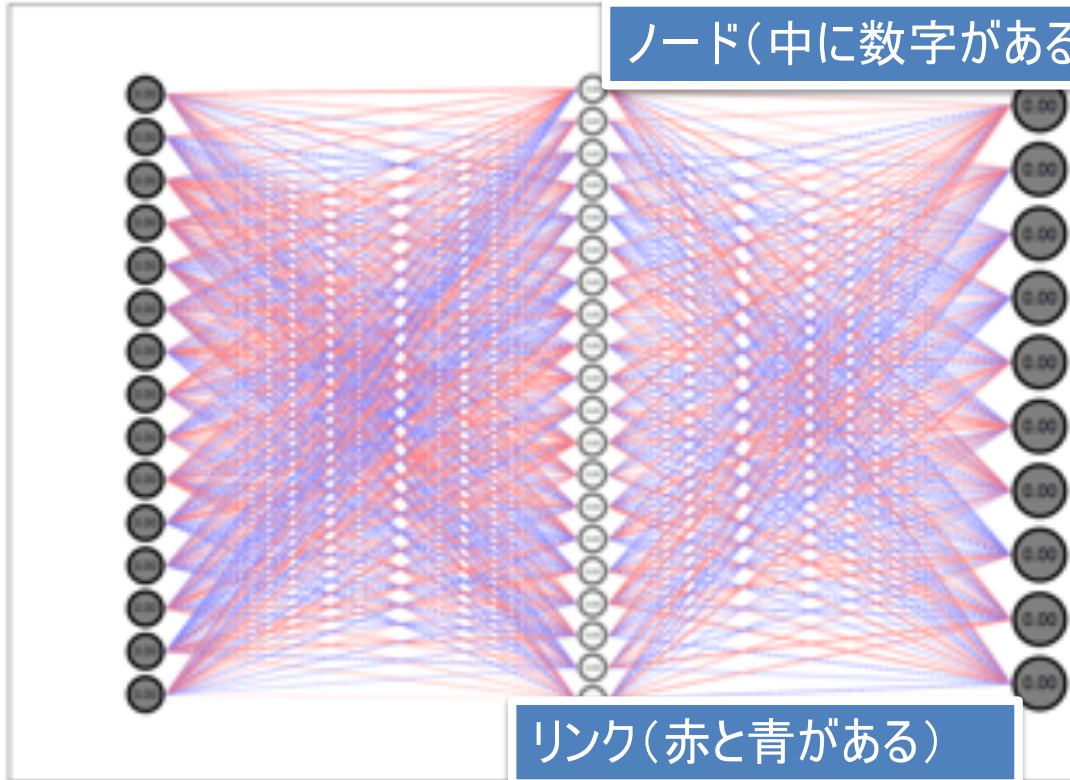
• <https://goo.gl/forms/krdvIFxRJT99Gu672>

上記とは別に授業後アンケート(学情経由)もあり

JavaScript Neural Network Demos

Demos **Pattern Recognition** | Numbers of nodes of hidden layers (comma separated) 20

数字を入力できる



ノード(中に数字がある)

リンク(赤と青がある)

Learning rate 0.5

End conditions

by error 0.05

by steps step

By errorにすると勝手にとまる

initialize train

step: 0
max error: 0

step 学習した回数
max error 誤差の大きさ(認識精度)

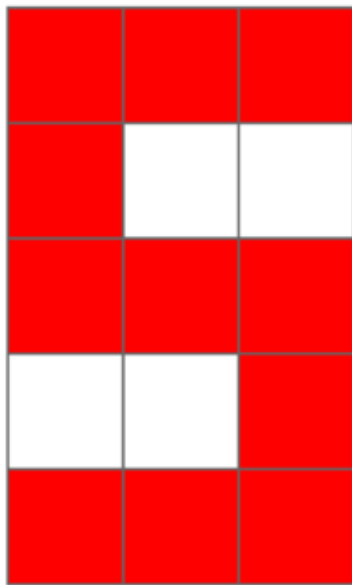
0 1 2 3 4 5 6 7 8 9 CLEAR

recognize



人間の何をモデリングしているのか = 数字の認識 (パターン認識)

曖昧なパターンも認識

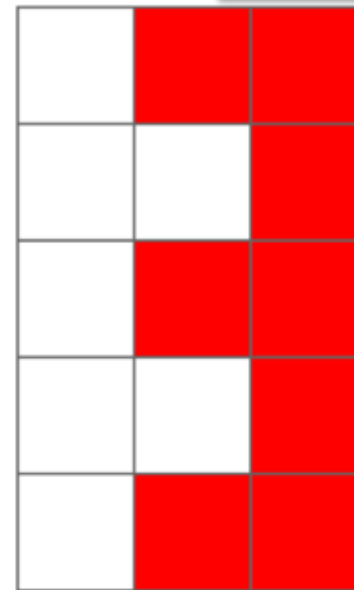


recognize

網膜に対応



認識された概念



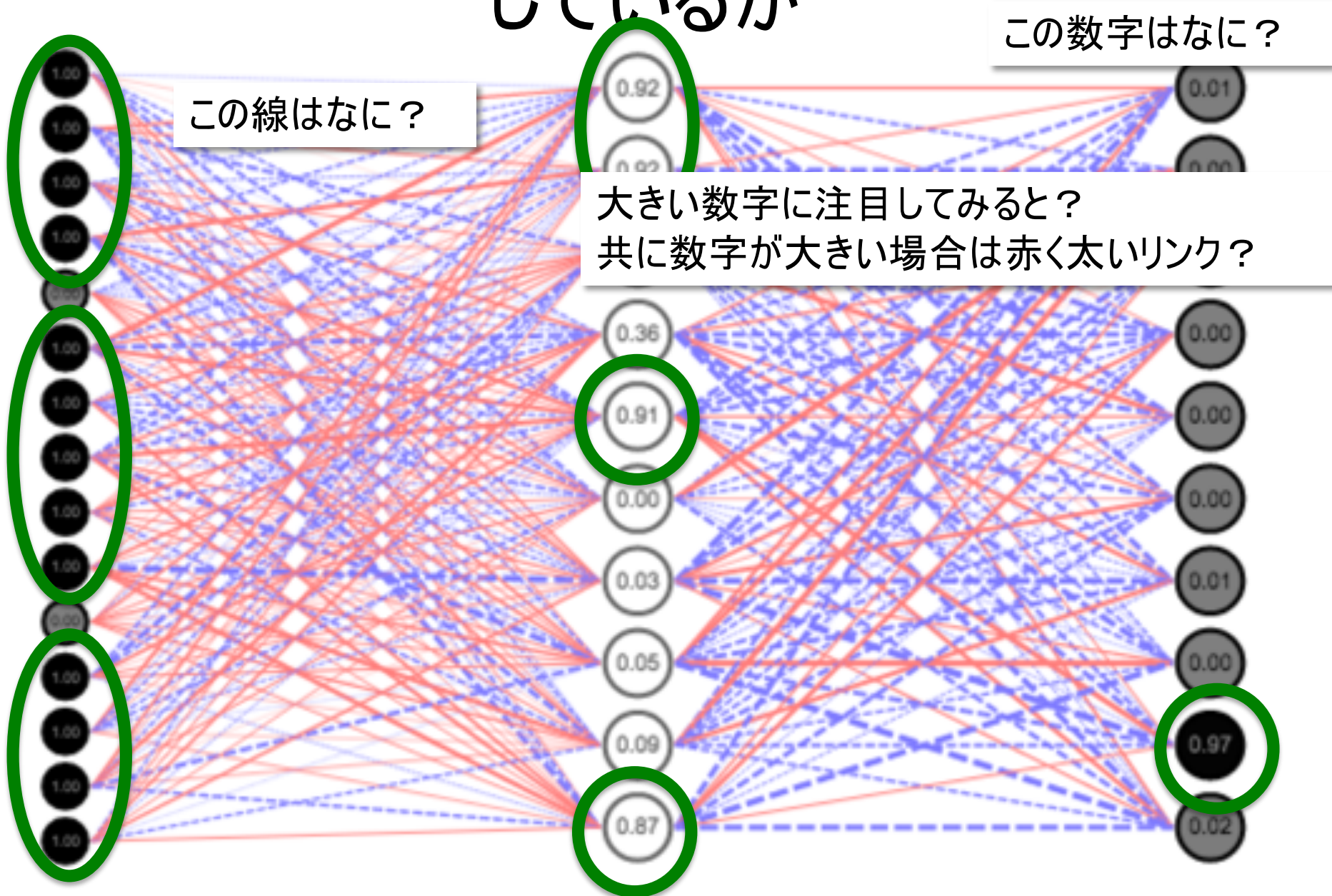
recognize

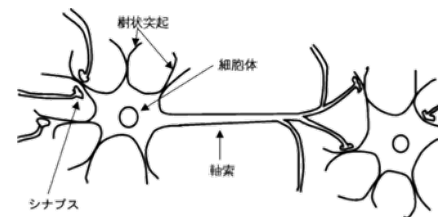
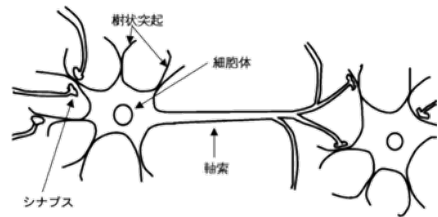
網膜に対応



認識された概念

リンクとノード中の数字がどのように関係しているか





入力層

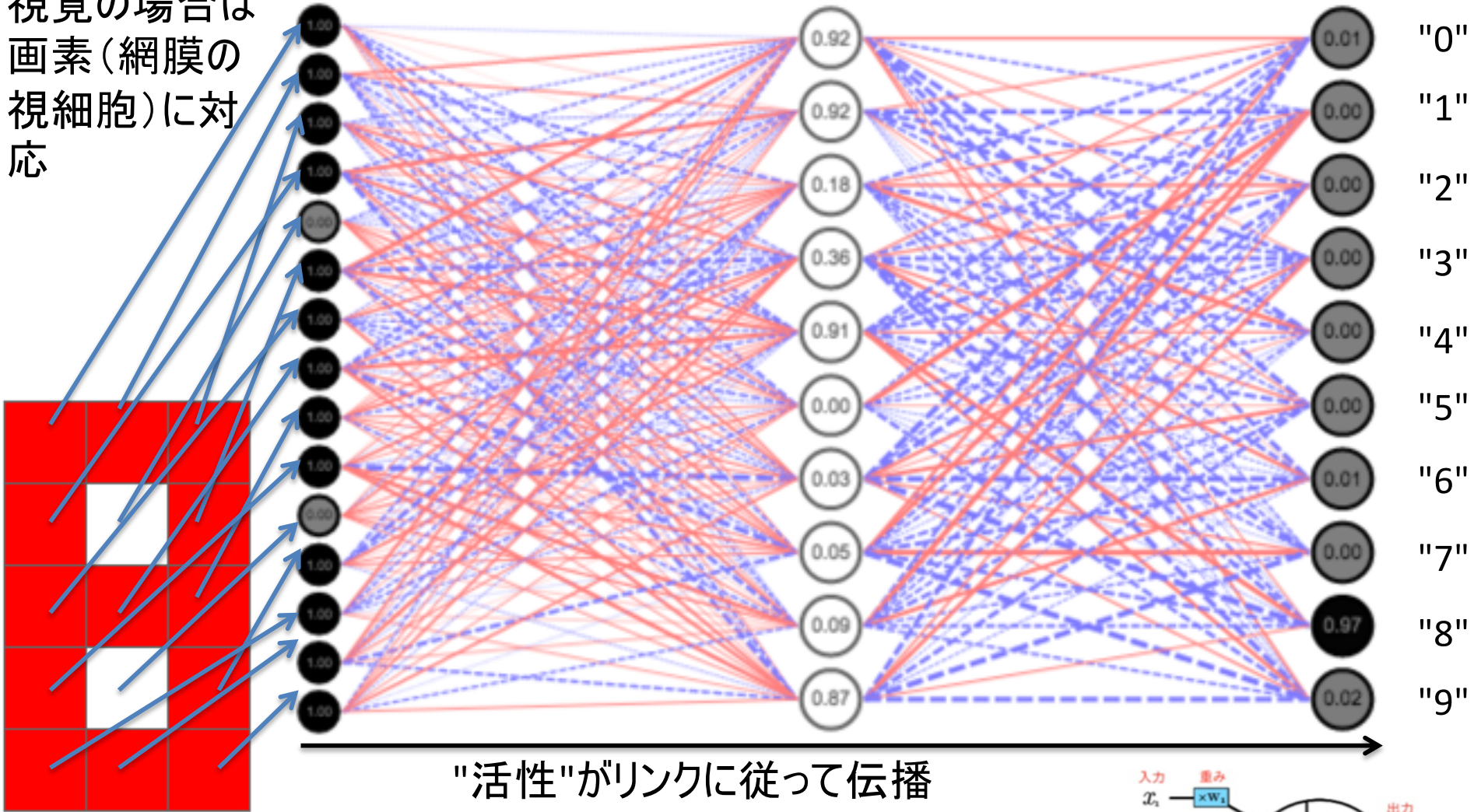
視覚の場合は
画素(網膜の
視細胞)に対
応

中間層

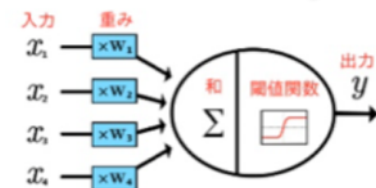
任意のノード数

出力層

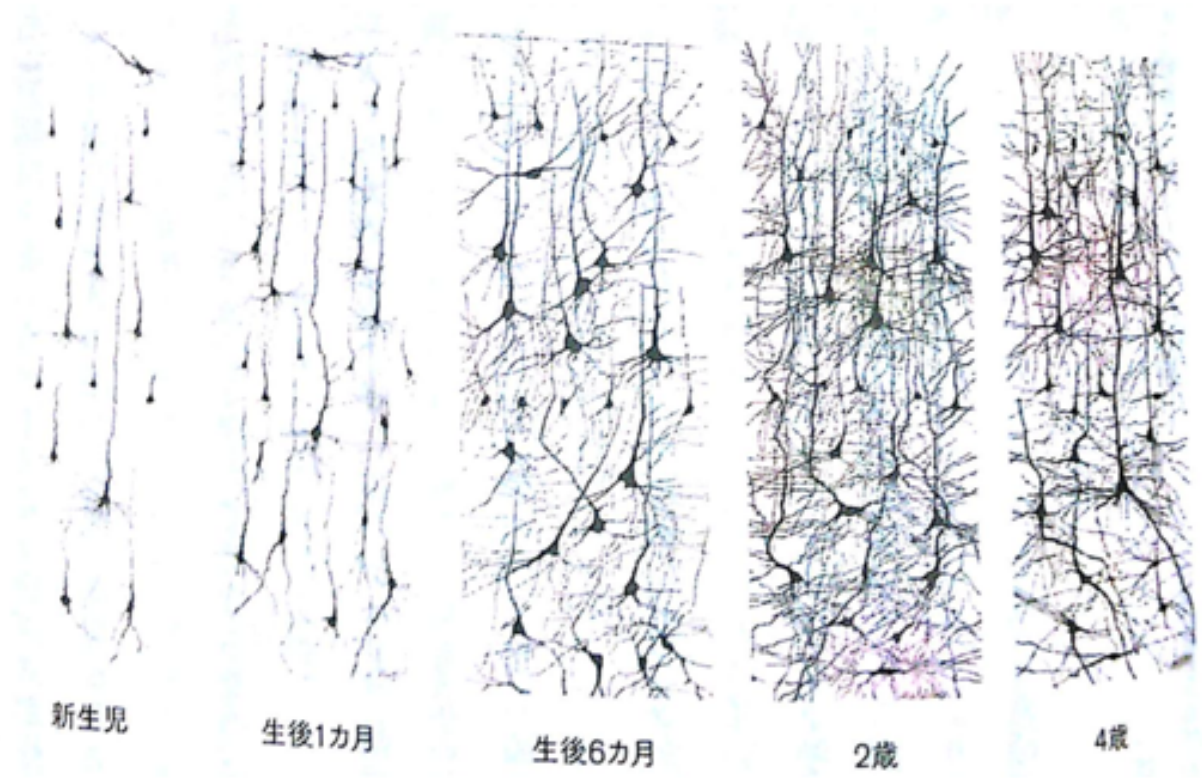
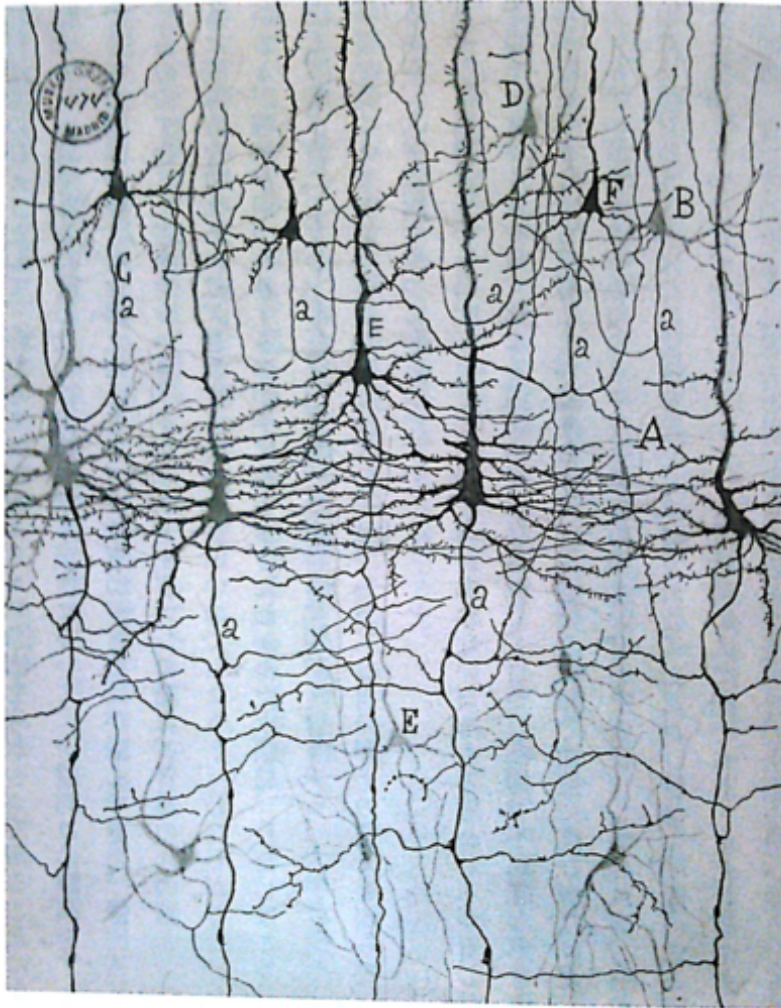
"記号"に対応



"活性"がリンクに従って伝播
 赤く太いリンクはプラスの影響
 青いリンクはマイナスの影響



脳内の神経回路

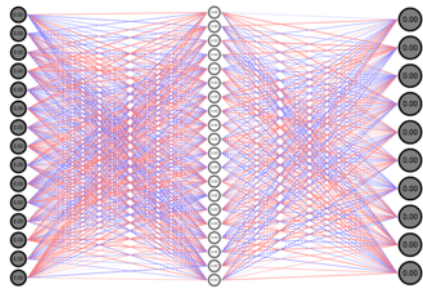


大脳皮質のニューロンも人工ニューラルネットと同じく多層構造

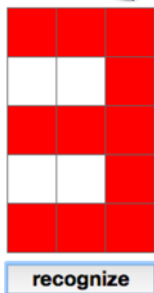
セバスチャン・スン (2015).コネクトーム:脳の配線はどのように「わたし」をつくり出すのか

重みの調整（誤差逆伝播法）

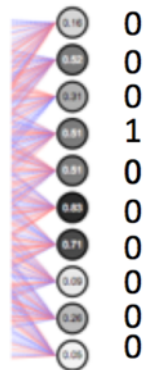
はじめはランダムに重み設定



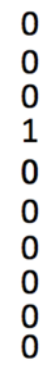
パターンを入力
活性化値が最大の出力
ノードを選択



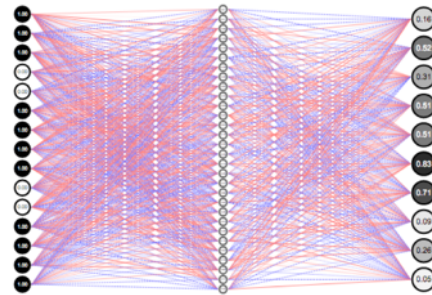
出力と
正解の誤差（差分）
をフィードバック



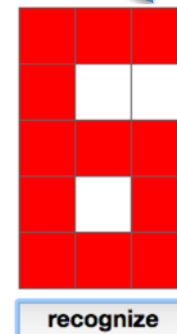
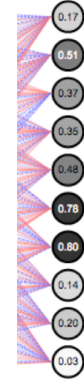
正解
パターン



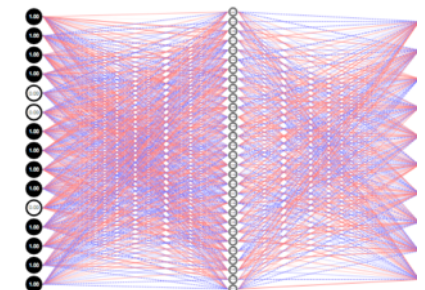
誤差を減らすよう全体の重み調整



パターンを入力
活性化値が最大の出力
ノードを選択



誤差を減らすよう全体

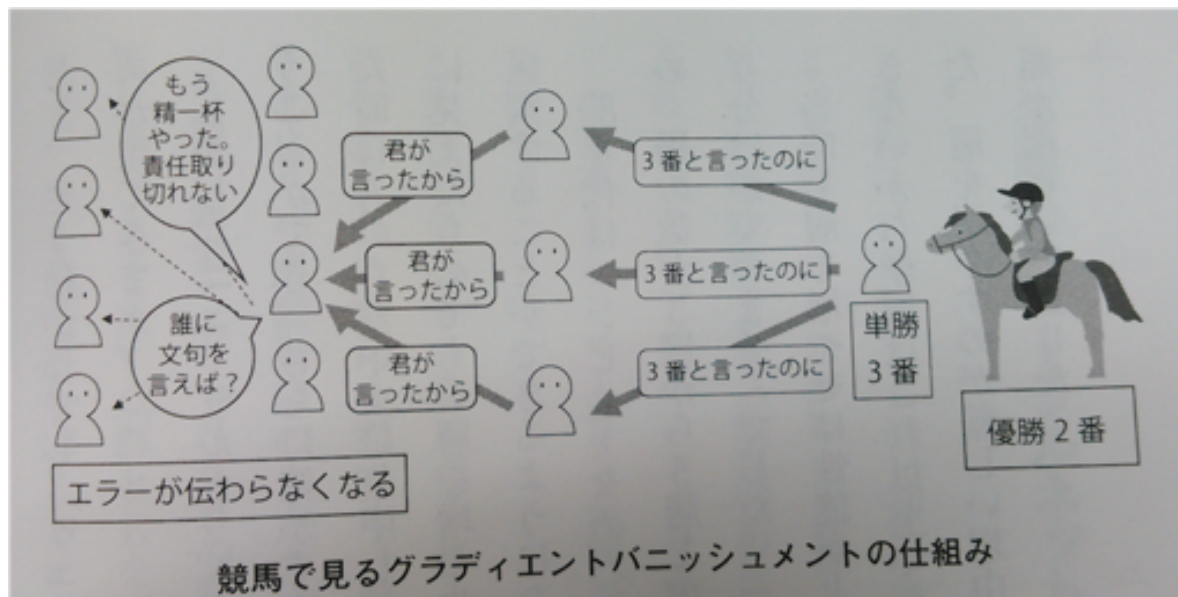
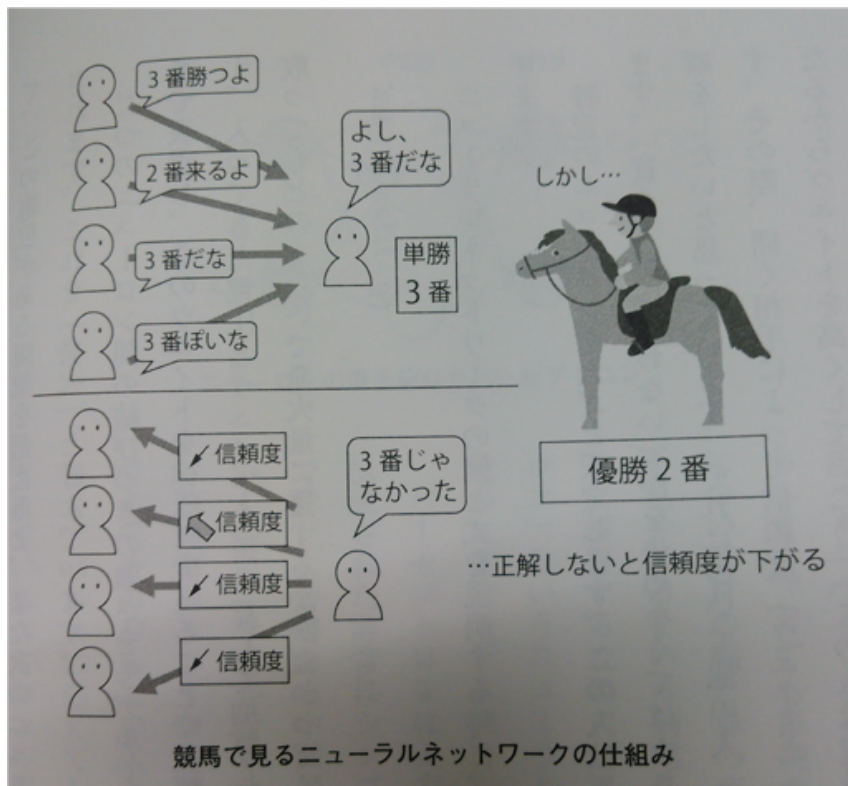


正解
パターン

誤差が一定以下になるまで学習ステップの繰り返し

馬券での説明

多層ニューラルネットの場合



初心者におすすめ→



深層学習（深層ニューラルネット）

深層学習におけるニューラルネットはただ階層が深いだけでなく、多数の特徴を扱うため各層の幅が広く、**パラメータ数が数億から数百億**と大きいのが特徴である。これだけ大きなモデルを学習するためには学習データも大きくなければならない。**数億個の学習データを利用**するといったような、これまでと比べ、桁違いに大きな学習が実現されている。データやモデルを大規模にすることで精度は著しく向上し、これまで実用化の壁を超えられなかった画像認識や音声認識といった技術が次々と陽の目を見ている。**量が質を変えた**と言える。人の脳が約数千億のニューロンから構成され、それぞれが1000個程度のニューロンとシナプス結合していることを考えると、さらに大きなニューラルネットワークが必要である知能学会監修『深層学習』2015



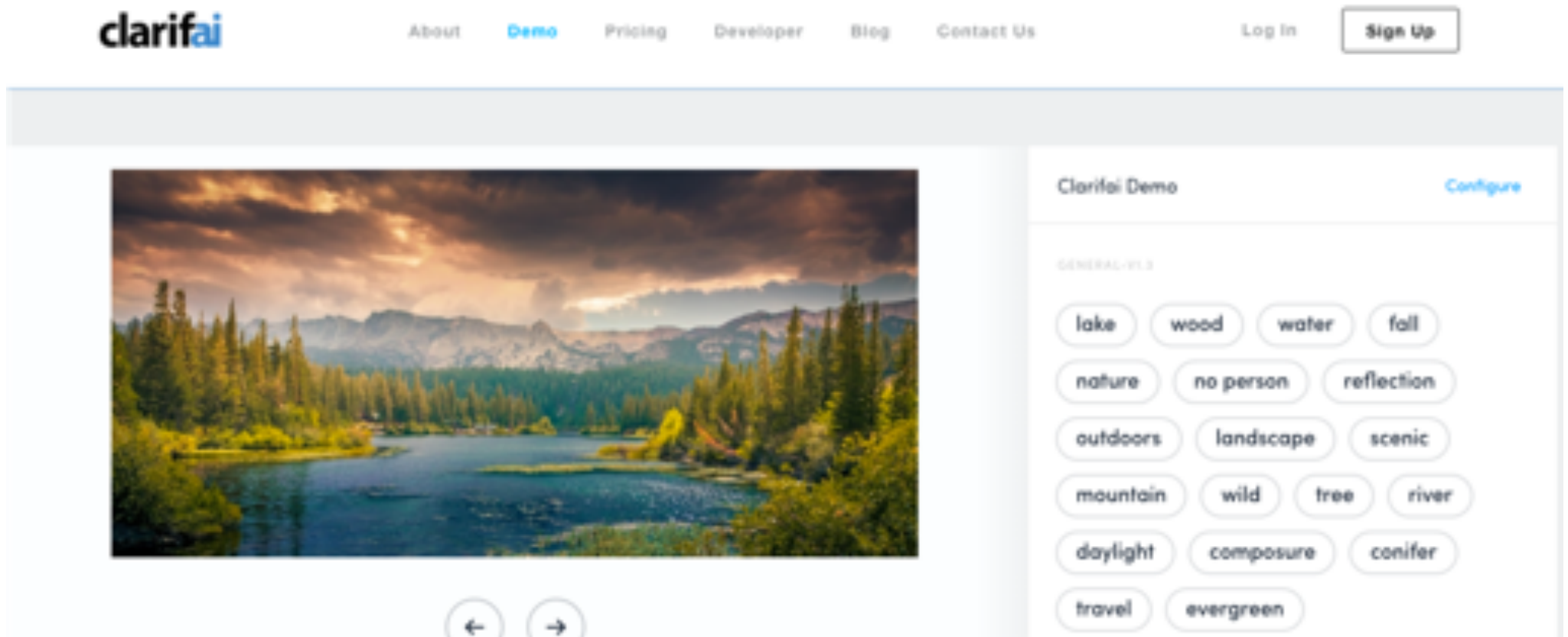
Google LeNet

深層ニューラルネットで実現されていること

- 手軽な体験

- <https://www.clarifai.com/>

自分の写真もドラッグ可能



The screenshot displays the Clarifai website interface. At the top left is the Clarifai logo. The navigation menu includes links for About, Demo, Pricing, Developer, Blog, and Contact Us. On the right side, there are links for Log In and a Sign Up button. The main content area features a large landscape photograph of a lake surrounded by trees and mountains. Below the image are two navigation arrows. To the right of the image is a panel titled 'Clarifai Demo' with a 'Configure' link. Underneath, a section labeled 'GENERAL TAGS' contains a grid of tags: lake, wood, water, fall, nature, no person, reflection, outdoors, landscape, scenic, mountain, wild, tree, river, daylight, composure, conifer, travel, and evergreen.

その他のデモ (Amazon Rekognition)

Amazon Rekognition

Facial analysis

Get a complete analysis of facial attributes, including confidence scores. (Your images aren't stored.)

Done with the demo?
[Download SDKs](#)

▼ Results

looks like a face	99.9%
appears to be female	80.5%
age range	28 - 43 years old
smiling	93%
appears to be happy	84.6%
wearing glasses	97.8%

[Show more](#)

Choose a sample image

Use your own image

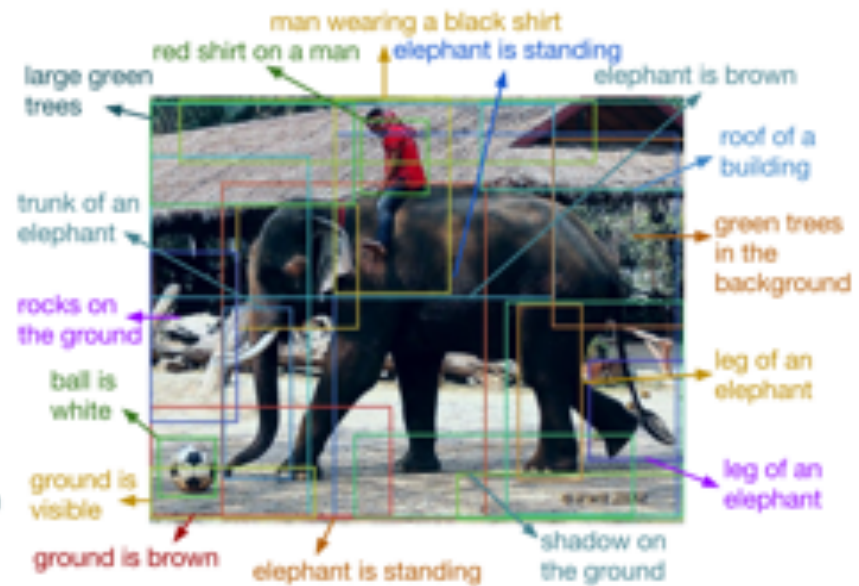
[Upload](#) or drag and drop

Use image URL [Go](#)

人間を超える知覚？

- DenseCap もはや古いが。。。
 - Johnson, Karoathy, Fei-Fei. CVPR 2016

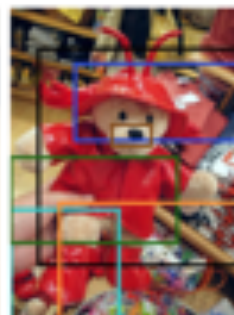
<https://www.youtube.com/watch?v=2wRnmRSrgCo>



Our Model: plane is flying, tail of the plane, red and white plane, plane is white, engine on the plane, windows on the plane, nose of the plane.



woman wearing a black shirt, baby bear is brown, chair is black, glass of wine, table is brown, woman with brown hair, paper on the table.



teddy bear is wearing a red shirt, red and white teddy bear, bear is wearing a red hat, red and white shirt, bear is brown, black nose of a bear.



train on the tracks, trees are green, front of the train is yellow, grass is green, green trees in the background, photo taken during the day, red train car.

ここまでのまとめ

- 物理的世界の記号化はすでにかなりの精度
 - 物理的な世界(画像や音声)から人間にとって理解のできる概念を出力
- 問い
 - 知覚の研究はこれで終了か？
 - 人間の知覚のモデルとして欠けているものがあるとすれば何か？

人間の視知覚

実験

指示にしたがって
ください

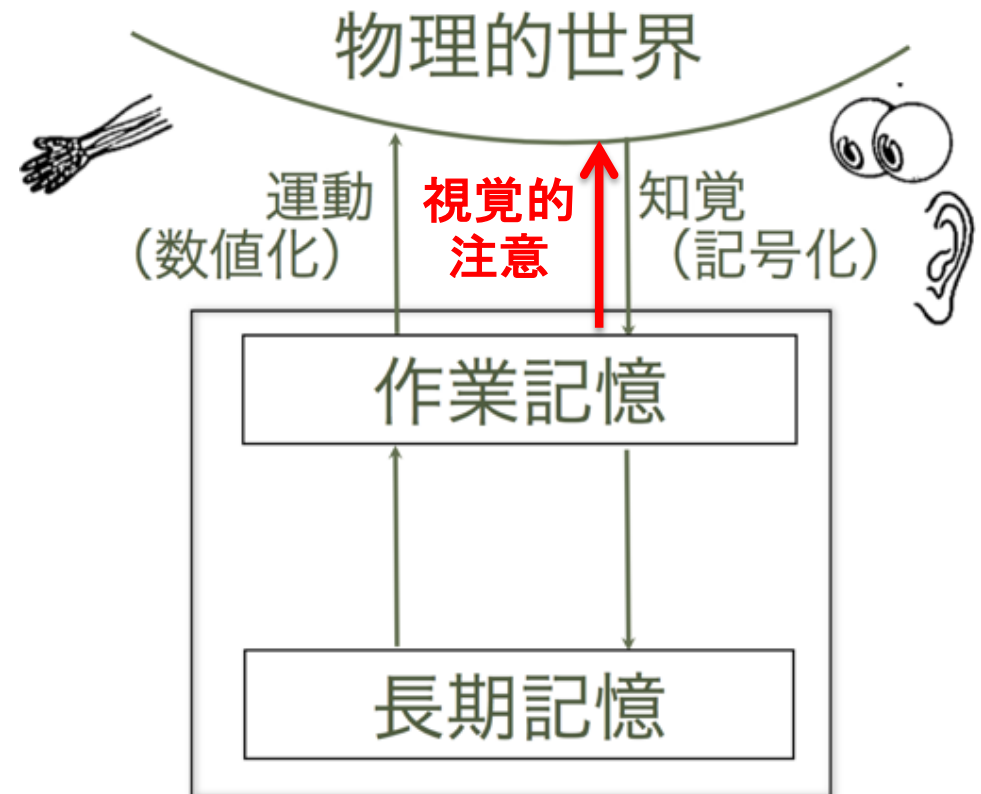
実験材料へのリンクは授業後

授業内アンケート

- 学情システムから届いたアンケートに回答してください。
 - 周りの人と意見交換してもOK
 - 時間は5分程度

人間の視知覚の特徴

- 目に映る全てを認識（記号化）しない
- 頭の中（作業記憶）のゴールに従って取捨選択する



概念駆動型知覚

トップダウン的知覚

←→ボトムアップ的知覚

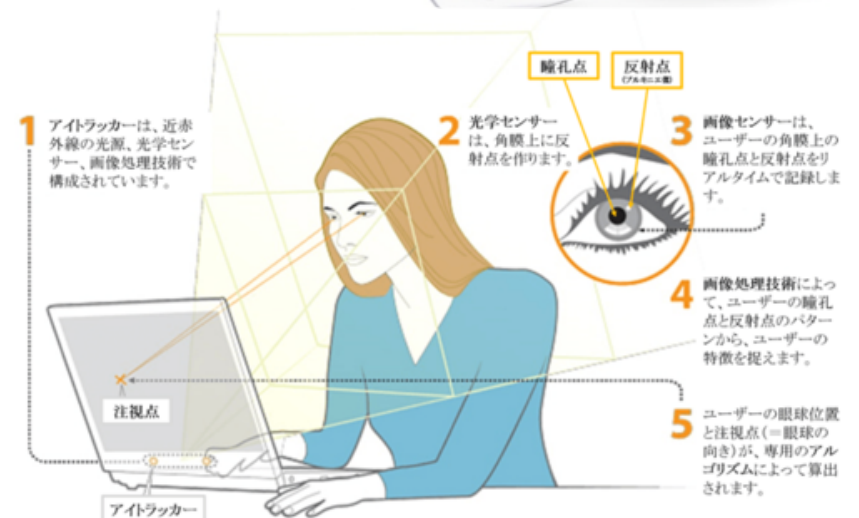
データ駆動型知覚

視覚的注意を計測する装置

- 視線計測装置＝アイトラッカー

<http://www.tobii.com/ja/product-listing/tobii-pro-glasses-2/>

こちらは100万オーバー。。。



https://commerce-star.com/okaimono/blog/tobii_eyetracking

THEEYETRIBE PRODUCTS DEVS PRESS BLOG ABOUT CONTACT US ORDER

AFFORDABLE EYE TRACKING JUST GOT BETTER!

EYE TRIBE TRACKER PRO

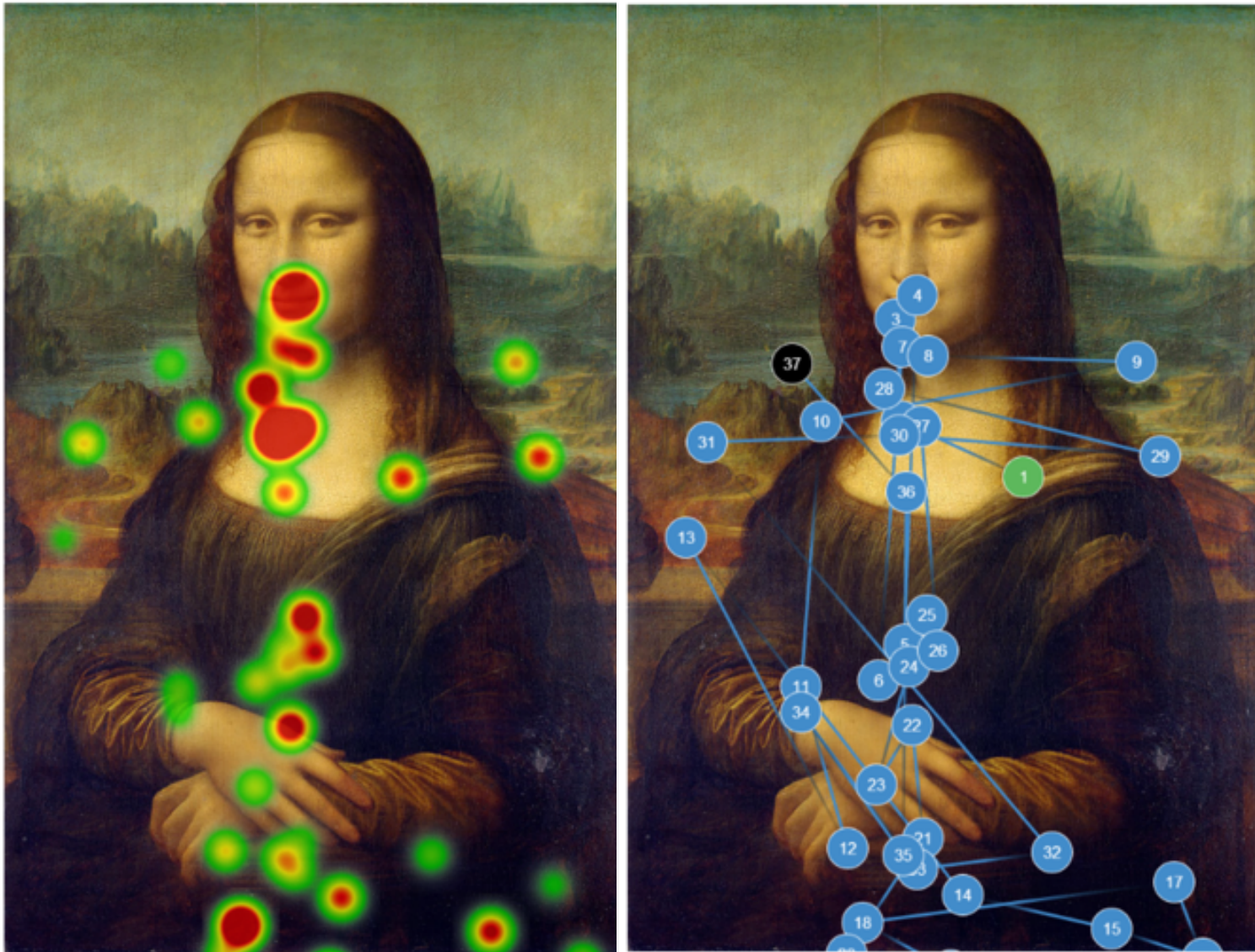
\$199 PROFESSIONAL EYE TRACKING

お値段なんと2万円!!
SDKもフリー

PRE-ORDER NOW

<https://theeyetribe.com>

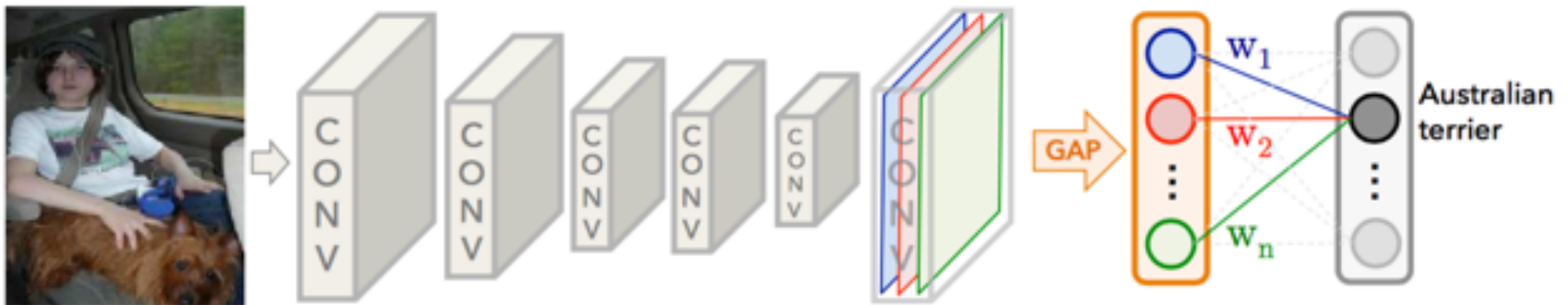
視線計測装置の出力



- <http://d.hatena.ne.jp/naraba/20140625/p1> より抜粋
- <https://www.youtube.com/watch?v=shsezhQtKf8> (動画デモ)

ディープラーニングで視覚的注意を 予測できるか

- Learning Deep Features for Discriminative Localization
 - Zhou, Khosla, Lapedriza, Oliva, Torralba, CVPR2016
 - <http://cnnlocalization.csail.mit.edu>



- 概念(記号)ありきの認識をする
- この写真がオーストラリアテリアなら入力層のどこが活性化している？

デモ

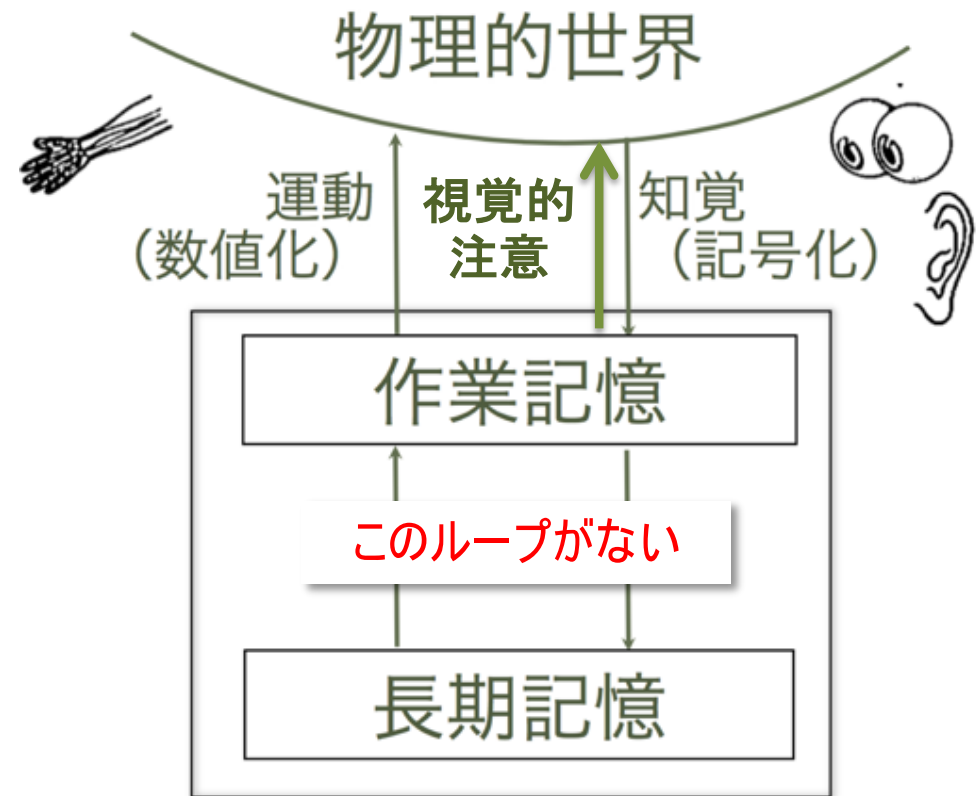
abaya



- 画像から概念を抽出し、抽出した概念(出力層)と強い重みでつながっている領域をハイライトしているだけ？
- 結局、歩いているのは人間
- 概念と概念がどう結びつくか、概念がどこから来たのかは説明されていない

ニューラルネットの限界

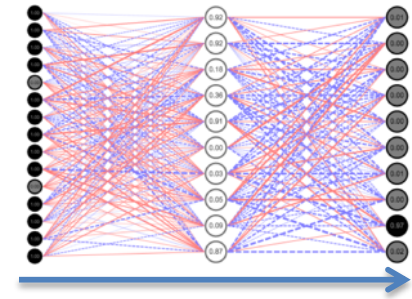
- 記号操作を組み入れられない
- 時系列的な処理が苦手
- できないこともないが、わかりづらいモデルになる(コスパが悪い)



頭の中のループ、補足は本日の資料の付録に

ニューラルネットデモ 補足スライド

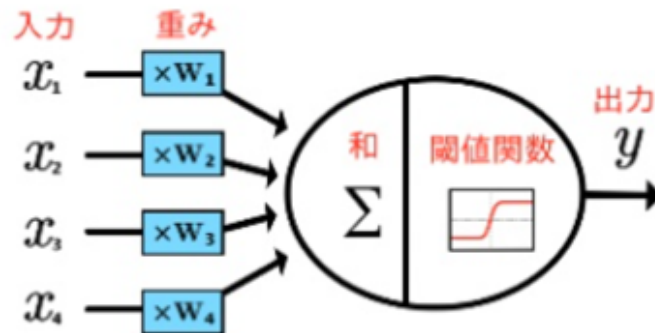
認識のプロセス



http://www.slideshare.net/z_kro/html5-36655060

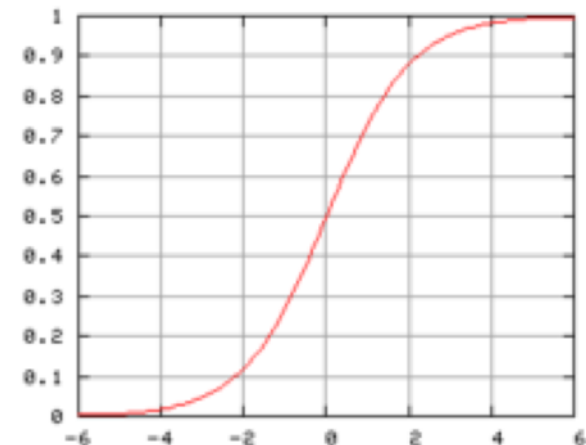
ニューロン

- ニューラルネットを構成する素子 → いわゆる「神経細胞」
- 多入力1出力
- 入力にはそれぞれ「重み」係数が掛けられる
- それらを足し合わせた後、閾値関数を通して出力される



φ の例 → シグモイド関数

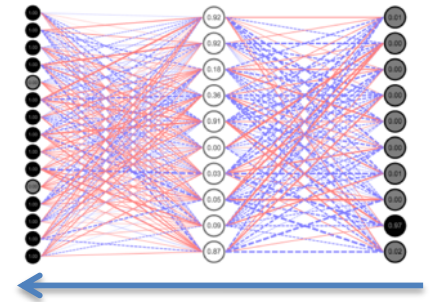
$$\varphi(x) = \frac{1}{1 + \exp(-x)}$$



直感的な説明

- 入力×重みの和が小さければOFF(0)を出力
- 入力×重みの和が大きければON(1)を出力

学習のプロセス



更新後のリンクの重み

現在のリンクの重み

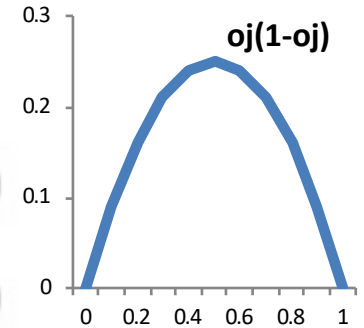
$$w'_{ij} = w_{ij} - \rho \epsilon_j o_i$$

変化量
 ρ は学習係数
 o_i は下位ノード i の出力

誤差 = 正解 (b_j) と実際の出力 (o_j) の差分

シグモイド関数の微分

$$\epsilon_j = \begin{cases} (o_j - b_j) o_j (1 - o_j) & \text{(出力層)} \\ (\sum_k \epsilon_k w_{jk}) o_j (1 - o_j) & \text{(中間層)} \end{cases}$$

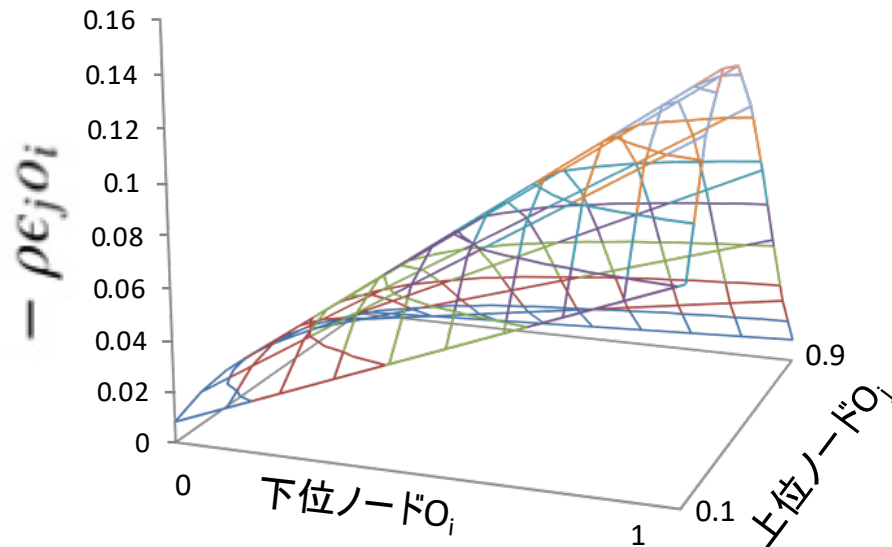


上位層の誤差の重みつき加算

重み (w_{ij}) の変化量

ただし $\rho = 1$ として計算

正解が1(活性)の場合



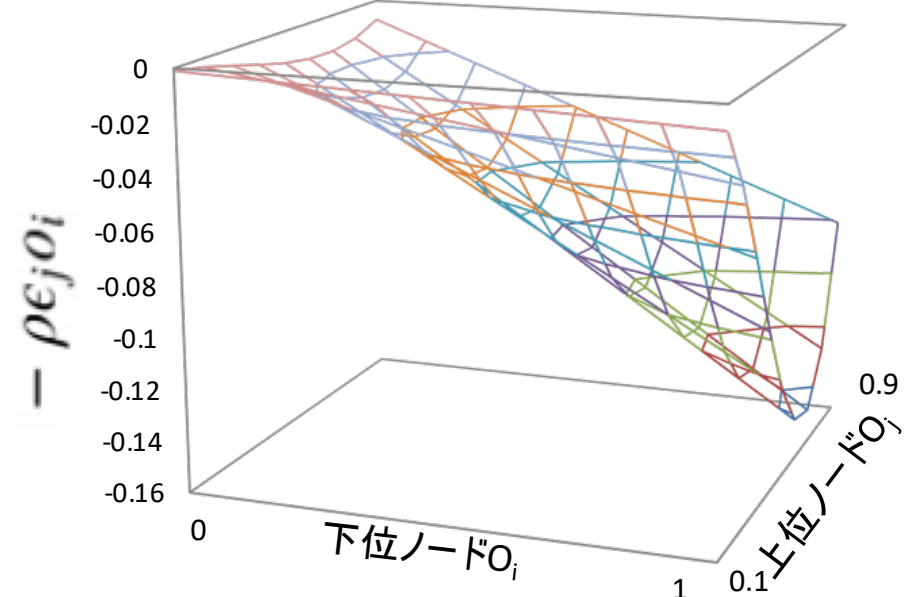
変化しない

- 下位ノード i が活性化していない(入力がない)
- 上位ノード j が1(正解した)

変化する

- 下位ノードが発火かつ不正解
- 変化量のピークは $o_j = 0.3$ 付近

正解が0(不活性)の場合



変化しない

- 下位ノード i が活性化していない(入力がない)
- 上位ノード j が0(正解した)

変化する

- 下位ノードが発火かつ不正解
- 変化量のピークは $o_j = 0.7$ 付近

end conditionをby errorとしたとき、trainの数を減らすには

模範回答

中間層のノード数を適当に調整する
(ただし、いくつかが最適なのかはわからない)

The image shows two side-by-side training configuration panels. Both panels have a 'Change' button next to a text input field for the number of nodes. The left panel has '30' in the input field, and the right panel has '30,30'. Both panels have a 'Learning rate' slider set to 0.5, 'End conditions' set to 'by error' with a slider at 0.05, and 'Speed' set to 'fast'. At the bottom of each panel are 'initialize' and 'retrain' buttons. A red arrow points from the 'step: 11832' text in the left panel to the 'step: 8736' text in the right panel, indicating a reduction in training steps.

Configuration	Hidden Layer Nodes	Learning Rate	End Conditions	Speed	Step	Max Error
Left	30	0.5	by error (0.05)	fast	11832	0.04996618894471749
Right	30,30	0.5	by error (0.05)	fast	8736	0.049974558213469744