

Title	視覚・言語情報を統合処理するニューラルネットワークに関する研究
Sub Title	Study on neural network for integrated processing of visual and linguistic information
Author	萩原, 将文(Hagiwara, Masafumi)
Publisher	
Publication year	2020
Jtitle	科学研究費補助金研究成果報告書 (2019.)
JaLC DOI	
Abstract	視覚情報と言語情報の統合的な処理機構の開発と、入力画像からの常識の自動獲得に関しては、画像を感性豊かに説明するキャプション自動生成システムの構築を行った。本システムは、物体の感性語を推定する際、画像特徴だけでなく物体の名詞情報も活用している。また表現力の高い感性語を生成するための感性語変換機構も有している。 自然言語を生成できるニューラルネットワークの開発に関しては、文脈を考慮した会話文の自動生成が可能な対話システムの構築を行った。さらにこれを、共感と助言に着目した自動相談システムへ発展させた。 Regarding the development of integrated processing of visual information and linguistic information and the automatic acquisition of common sense from the input image, an automatic caption generation system is constructed. It can explain an image richly. The system utilizes not only the image features but also the noun information of the object when estimating the affective words of the object. It also has an affective word conversion mechanism to generate expressive words. Concerning the development of neural networks that can generate natural language, we have constructed an automatic conversation system that can automatically generate conversational sentences considering the context. Furthermore, this was developed into an automatic consultation system focusing on empathy and advice.
Notes	研究種目：基盤研究 (C) (一般) 研究期間：2016～2019 課題番号：16K00338 研究分野：ニューラルネットワーク
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KAKEN_16K00338seika

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

令和 2 年 5 月 26 日現在

機関番号：32612

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K00338

研究課題名(和文) 視覚・言語情報を統合処理するニューラルネットワークに関する研究

研究課題名(英文) Study on neural network for integrated processing of visual and linguistic information

研究代表者

萩原 将文 (Hagiwara, Masafumi)

慶應義塾大学・理工学部(矢上)・教授

研究者番号：80198655

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：視覚情報と言語情報の統合的な処理機構の開発と、入力画像からの常識の自動獲得に関しては、画像を感性豊かに説明するキャプション自動生成システムの構築を行った。本システムは、物体の感性語を推定する際、画像特徴だけでなく物体の名詞情報も活用している。また表現力の高い感性語を生成するための感性語変換機構も有している。

自然言語を生成できるニューラルネットワークの開発に関しては、文脈を考慮した会話文の自動生成が可能な対話システムの構築を行った。さらにこれを、共感と助言に着目した自動相談システムへ発展させた。

研究成果の学術的意義や社会的意義

従来より学術的には、画像のような視覚情報と言語情報は別々に扱われることが多かった。しかし実生活やTV、書籍、Webなど、ほとんど場合、これらは分離されことなく総合的に扱われている。本研究の目的は、画像と言語を統合的に扱う方法を考案することにある。これは社会からの需要も高く、また学術的にも画像のようなパターン情報と言語のようなシンボル情報を統合処理するという意味で高い意義を有している。

研究成果の概要(英文)：Regarding the development of integrated processing of visual information and linguistic information and the automatic acquisition of common sense from the input image, an automatic caption generation system is constructed. It can explain an image richly. The system utilizes not only the image features but also the noun information of the object when estimating the affective words of the object. It also has an affective word conversion mechanism to generate expressive words.

Concerning the development of neural networks that can generate natural language, we have constructed an automatic conversation system that can automatically generate conversational sentences considering the context. Furthermore, this was developed into an automatic consultation system focusing on empathy and advice.

研究分野：ニューラルネットワーク

キーワード：視覚情報処理 言語情報処理 ニューラルネットワーク 感性情報

様式 C-19、F-19-1、Z-19（共通）

1. 研究開始当初の背景

高齢化による労働力の減少や介護の問題、産業分野における各種サービスや危険作業などのために、ロボット導入への需要は大きく高まりつつある。しかしながら、柔軟な情報処理を行う頭脳部の研究開発は遅れているのが実情である。近年、人工知能（AI）が大きく注目されているが、基本的には大容量データに統計・確率的な処理を施すものである。人間のような柔軟な思考やインタフェースをめざす場合には、知識や経験の扱いが不可欠である。したがって脳を考慮したアプローチが自然であり、将来的な可能性も高いと研究代表者は考えている。

人間の脳では、視覚、言語（聴覚情報を含む）、そして人間性に関連する感情や感性の情報処理が重要と考えられる。しかしながら従来のニューラルネットワークはいわばシングルタスク用で、これらの異種情報、すなわち視覚、言語、感性情報を統合して扱うモデルはなかった。本研究は特に、視覚と言語の異種情報を統合的に扱うニューラルネットワーク構築を目的としている。

2. 研究の目的

優れた特性を示す深層学習など、脳の情報処理の工学的実現をめざすニューラルネットワークが注目されている。本研究は、従来は別個に扱われていた画像処理機能と自然言語処理機能、すなわち視覚・言語情報を統合的に扱うことのできるニューラルネットワークの構築を行ない、以下の課題の解決を目指すものである。

- ・視覚情報と言語情報の統合的処理機構の開発
- ・統合的処理による入力画像からの物体に関する常識の自動獲得
- ・自然言語を生成できるニューラルネットワークの開発

3. 研究の方法

情報工学の研究であり、ニューラルネットワークを計算機に実装するアプローチを用いた。

4. 研究成果

年度毎に研究成果の一部を説明する。

(1) 2016 年度

平成 28（2016）年度における課題は「視覚情報と言語情報の統合的処理機構の開発」であり、以下の項目が具体的な課題である。まず第一に CNN（Convolutional Neural Network: 畳み込みニューラルネットワーク）をベースとした視覚情報処理ニューラルネットワークの構築である。人とのコミュニケーションを考えると、ユーザの表情や感情の推定は重要である。そこで本年度は、Double-Column Convolutional Neural Network（DCNN）と Spatial Pyramid Pooling（SPP）を融合させた新しい表情認識システムの構築を行った。

第二の課題は、知識を扱うことのできるニューラルネットワークの構築である。知識表現の代表例として意味ネットワークがある。これは事象の連想関係を、ローカル表現を用いたネットワークにより表現する方法であり、概念の類似性を扱うことができない。そこで、単語の分散表現と Restricted Boltzmann Machine（RBM）を用いて新しい意味ネットワークの構築を行った。具体的には、以下の 2 つの研究を行った。まず連想メモリとしての改良である。既存の RBM 連想メモリでは自己連想記憶にしか対応していない。そこでネットワークに新たにミドル層を追加することで、双方向の連想記憶、及び 1 対多の想起が可能な連想メモリとした。二つ目は、分散表現された単語を扱えるようにする研究である。

提案ネットワークを図 1 に示す。単語の分散表現を用いて単語同志の関係性の学習を行う。これにより単語をベクトル情報で保持するため、従来の知識表現では対応できなかった未知語への対応が可能になった。(①)

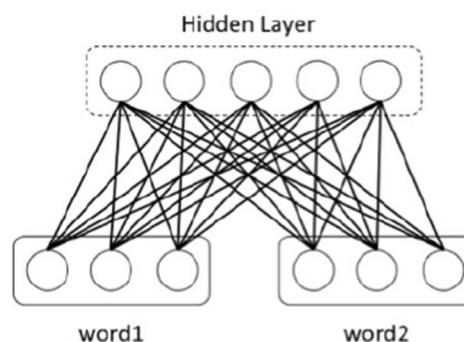


図 1 RBM アナログ連想メモリ

(2) 2017 年度

2017 年度における課題は「統合的処理による入力画像からの物体に関する常識の自動獲得」である。本年度は、自然言語処理研究と画像処理研究を組み合わせ、画像を感性豊かに説明するキャプション生成に取り組んだ。具体的には単語の分散表現と画像特徴を活用したシステムを構築した。

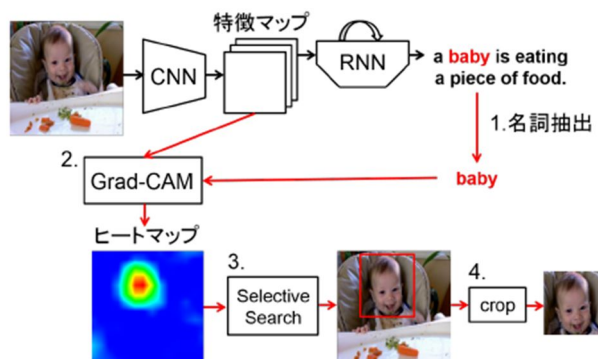


図 2 感性豊かに説明するキャプション自動生成システム



a **cuddly**(抱きしめたいくなるような) baby is eating a piece of food.



a **small cheeky**(小さく生意気な) bird sitting on a branch.



a **pathetic**(哀愁を誘う) dog is sitting on a bench in a room.

図3 提案手法で生成された画像キャプションの例

提案システムの構成を図2に示す。これには、以下の2点の特長がある。第一に、物体の感性語を推定する際、画像特徴だけでなく物体の名詞情報も活用した点である。既存研究では画像特徴のみを用いて感性語の推定を行っていた。提案システムでは、CNN (Convolutional Neural Network) を用いて抽出した画像特徴と word2vec で分散表現に変換した名詞情報を入力とし、感性語を推定するニューラルネットワークを構築することでより正確に感性語を推定する。第二に、表現力の高い感性語を生成するための感性語変換機構を構築した点である。表現力の高い感性語は複数の感性語のニュアンスの組み合わせで決定するという仮定のもと、word2vec を活用し、複数の感性語をより表現力の高い感性語に変換する。

図3に入力画像に対して生成されたキャプションの例を示す。評価実験では、3種類の実験を行った。まず、名詞情報を活用した感性語推定部の評価を行った。結果として、画像特徴のみを利用したシステムと比較して高い正解率で感性語が推定できることが確認された。次に、生成したキャプションの定量的評価実験と主観評価実験を行った。定量的評価実験では、3つの指標を用いてキャプションを評価し、既存の感性語を付与したキャプション生成システムを上回る結果になった。主観評価実験では、提案システムによって、多様な感性語を付与した表現力の高い画像キャプションが生成可能であることが示唆された。(②,③)

(3) 2018 年度

2018年度の課題は、「視覚・言語情報を統合処理するニューラルネットワークの完成」である。より細かい目標設定としては、「自然言語を生成できるニューラルネットワークの開発」と「ネットワーク全体での調整」であった。本年度は、特に前者の自然言語処理研究に力を入れた。具体的には、転移学習を用いた階層型潜在変数付きエンコーダ・デコーダによる自動相談システムの構築を行い、文脈を考慮した会話文の自動生成を可能とした。これは図4に示す Latent Variable Hierarchical Recurrent Encoder-Decoder (VHRED) というモデルを利用し、これに転移学習を用いて会話の流れの考慮を可能とした自動相談システムとなっている。

機械学習ベースの対話システムの作成においては、Sequence to Sequence(Seq2Seq)モデルが注目を集めている。しかし、Seq2Seq には1ペアの対話を学

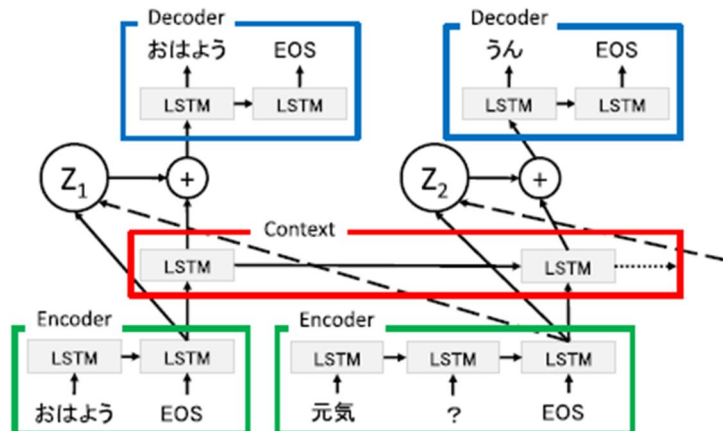


図4 会話生成を行うVHREDモデル図

発話者	発言
ユーザ	今好きな人がいます 彼が私を恋愛対象で見ているかどうか 教えてください
システム	こんばんはありがとうございます よろしくお願い致します
ユーザ	よろしくお願いします! 恋愛対象ですか?
システム	彼との相性は大吉ですね
ユーザ	なら頑張ってみます
システム	意見が合わないところがありそうですね
ユーザ	なるほど
システム	すごいですね ただ何をしても後悔しているようですね
ユーザ	ありがとうございます
システム	よい流れとなりますよう ありがとうございました

図5 提案手法により生成された対話例

発話者	発言
システム	お悩みをなんでも相談してくださいね
ユーザ	今付き合っている彼氏と将来うまくいくでしょうか
システム	彼氏の方とは仲良くしていけそうですが
ユーザ	やった
システム	よかったですね 彼氏の方はアピールすることが大切ですから よいご縁を運んでいるようなので今まで通り信頼できる お二人だという思いもあるけれどちょっと楽天的すぎる傾向にあります
ユーザ	なるほど！
システム	彼氏についてはちょっと子供っぽくて面倒をみた方がいいですね
ユーザ	わかりました結婚はできますかね
システム	結婚はいいですよ 結婚についてはまだ未定という感じですね
ユーザ	気長に考えることにしますね
システム	そうですね彼の方はまだ遊んでいたいという気持ちもありますね
ユーザ	ありがとうございました
システム	よい流れとなりますよう、ありがとうございました

図 6 提案手法により生成された共感のある対話例

習しているために会話の流れを考慮できない。したがって、同じ入力からは同じ出力しか得られないという欠点があった。この問題点を低減するために提案されたのが VHRED である。しかしこれを学習させるために必要な対話ターン数の長い日本語コーパスは、少量しか存在しないという問題があった。そこで VHRED に転移学習を適用した手法を提案した。具体的には、VHRED の Encoder 層と Decoder 層を Twitter から得た大量の対話ペアを用いて学習する。そして、その他の層を会話の流れを反映するための対話ターン数の長いコーパスで学習する。この学習方法により、少量のコーパスでも VHRED を学習させることが可能となった。

図 5 に提案手法により生成された対話例を示す。評価実験では、転移学習を行わない VHRED と比較して主観評価を行った。結果として、話の流れの自然さや応答文の多様性の項目において提案手法の有効性が示された (④)。

(4) 2019 年度

2018 年度に開発した自動会話システムを、共感と助言に着目した自動相談システムへ発展させた。相談においては特に相手の感情に対して共感する発話と相手に対して情報を与える助言の発話が重要である。提案システムでは相談者の入力文から抽出した感性語や話題語をもとにこれらの使い分けを行うことができる。図 6 に提案手法により生成された共感のある対話例を示す。

<引用文献>

- ① Yuichiro Tsutsui and Masafumi Hagiwara: “Analog Value Associative Memory Using Restricted Boltzmann Machine”, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.23, No.1, pp.60-66 (2019). doi: 10.20965/jaciii.2019.p0060.
- ② 瀬口太郎, 萩原将文: “画像特徴と単語の分散表現を活用した感性語を付与した画像キャプションの自動生成,” 電子情報通信学会ニューロコンピューティング研究会, 信学技報, NC2017-96, pp.169-174, 2018-03.
- ③ 三由裕也, 萩原将文 “感情推定を利用した感性的な画像説明文自動生成システム”, 日本感性工学研究論文集, Vol.18, No.2, pp.151-158, 2019.
<https://doi.org/10.5057/jjske.TJSKE-D-18-00071>
- ④ 和田翔, 萩原将文: “転移学習を用いた階層型潜在変数付きエンコーダ・デコーダによる自動相談システム,” 日本感性工学会論文誌, Vol.18, No.4, pp.307-317, 2019.

5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 12件／うち国際共著 0件／うちオープンアクセス 10件）

1. 著者名 武内達哉, 萩原将文	4. 巻 18
2. 論文標題 単語の持つ感情推定法の提案と単語感情辞書の構築	5. 発行年 2019年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 273-278
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.5057/jjske.TJSKE-D-18-00104	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 —

1. 著者名 和田翔, 萩原将文	4. 巻 18
2. 論文標題 転移学習を用いた階層型潜在変数付きエンコーダ・デコーダによる自動相談システム	5. 発行年 2019年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 307-317
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.5057/jjske.TJSKE-D-19-00008	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 —

1. 著者名 三由裕也, 萩原将文	4. 巻 18
2. 論文標題 感情推定を利用した感性的な画像説明文自動生成システム	5. 発行年 2019年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 151-158
掲載論文のDOI (デジタルオブジェクト識別子) 10.5057/jjske.TJSKE-D-18-00071	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 —

1. 著者名 佐川友里香, 萩原将文	4. 巻 18
2. 論文標題 感性語を考慮した敵対的生成ネットワークを用いたロゴマーク生成	5. 発行年 2019年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 215-222
掲載論文のDOI (デジタルオブジェクト識別子) 10.5057/jjske.TJSKE-D-18-00091	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 —

1. 著者名 武内達哉, 萩原将文	4. 巻 18
2. 論文標題 単語の持つ感情推定法の提案と単語感情辞書の構築	5. 発行年 2019年
3. 雑誌名 日本感性工学会論文誌	6. 最初と最後の頁 273-278
掲載論文のDOI (デジタルオブジェクト識別子) 10.5057/jjske.TJSKE-D-18-00104	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 和田翔, 萩原将文	4. 巻 18
2. 論文標題 転移学習を用いた階層型潜在変数付きエンコーダ・デコーダによる自動相談システム	5. 発行年 2019年
3. 雑誌名 日本感性工学会論文誌	6. 最初と最後の頁 307-317
掲載論文のDOI (デジタルオブジェクト識別子) 10.5057/jjske.TJSKE-D-19-00008	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 三由裕也, 萩原将文	4. 巻 18
2. 論文標題 感情推定を利用した感性的な画像説明文自動生成システム	5. 発行年 2019年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 151-158
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.5057/jjske.TJSKE-D-18-00071	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 佐々木駿也, 萩原将文	4. 巻 J102-D
2. 論文標題 動的多層化ニューラルネットワーク	5. 発行年 2018年
3. 雑誌名 電子情報通信学会論文誌(D)	6. 最初と最後の頁 226-234
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transinfj.2018JDP7045	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 —

1. 著者名 Yuichiro Tsutsui and Masafumi Hagiwara	4. 巻 23
2. 論文標題 Analog Value Associative Memory Using Restricted Boltzmann Machine	5. 発行年 2019年
3. 雑誌名 Journal of Advanced Computational Intelligence and Intelligent Informatics	6. 最初と最後の頁 60-66
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.20965/jaciii.2019.p0060	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 三上佳孝, 萩原将文	4. 巻 17
2. 論文標題 対話システムにおけるランダム性を考慮した話題展開手法	5. 発行年 2018年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 365-373
掲載論文のDOI (デジタルオブジェクト識別子) TJSKE-D-17-00084	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 三上佳孝, 萩原将文	4. 巻 17
2. 論文標題 多様な応答のための分散表現ベース対話システム	5. 発行年 2018年
3. 雑誌名 日本感性工学研究論文集	6. 最初と最後の頁 275-283
掲載論文のDOI (デジタルオブジェクト識別子) TJSKE-D-18-00006	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 —

1. 著者名 三上佳孝, 萩原将文	4. 巻 29
2. 論文標題 対話における文脈を考慮した分散表現生成ニューラルネットワーク	5. 発行年 2017年
3. 雑誌名 日本知能情報ファジィ学会誌	6. 最初と最後の頁 595-604
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 —

〔学会発表〕 計12件（うち招待講演 3件／うち国際学会 3件）

1. 発表者名 Masafumi Hagiwara
2. 発表標題 Toward next AI considering emotion and will
3. 学会等名 International Symposium on Affective Science and Engineering 2018（招待講演）（国際学会）
4. 発表年 2018年

1. 発表者名 筒井佑一郎, 萩原将文
2. 発表標題 RBMを用いたアナログ連想メモリ
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2017年

1. 発表者名 萩原将文
2. 発表標題 知情意の統合をめざして --- 人工 {知情意} 能 ---
3. 学会等名 日本知能情報ファジィ学会関西支部第114回例会（招待講演）
4. 発表年 2017年

1. 発表者名 三上佳孝, 萩原将文
2. 発表標題 対話システムにおける多様な話題の展開手法
3. 学会等名 日本感性工学会 第47回あいまいと感性研究部会研究発表講演会
4. 発表年 2017年

1. 発表者名 三由裕也, 萩原将文
2. 発表標題 感情推定を利用した感性的な画像説明文自動生成システム
3. 学会等名 日本感性工学会 第49回あいまいと感性研究部会研究発表講演会
4. 発表年 2018年

1. 発表者名 Kazuki Isoshima and Masafumi Hagiwara
2. 発表標題 A Non-Task-Oriented Dialogue System Controlling the Utterance Length
3. 学会等名 International Conference on Soft Computing and Intelligent Systems & ISIS (国際学会)
4. 発表年 2018年

1. 発表者名 Yuichiro Tsutsui, Masafumi Hagiwara
2. 発表標題 Continuous-Value Associative Memory Using Restricted Boltzmann Machine
3. 学会等名 International Conference on Soft Computing and Intelligent Systems (ISIS 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 萩原将文
2. 発表標題 心を考慮した人工知能へ ～知情意の考慮～
3. 学会等名 第50回日本人間工学会中国・四国支部大会 (招待講演)
4. 発表年 2017年

1. 発表者名 瀬口太郎, 萩原将文
2. 発表標題 画像特徴と単語の分散表現を活用した感性語を付与した画像キャプションの自動生成
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2018年

1. 発表者名 佐々木駿也, 萩原将文
2. 発表標題 段階的に多層化するニューラルネットワーク
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2018年

1. 発表者名 椎名健之, 萩原将文
2. 発表標題 分散表現を用いたRBM連想メモリによる知識処理
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2017年

1. 発表者名 山本真大, 萩原将文
2. 発表標題 連想層と抑制層を導入した言語処理ニューラルネットワーク
3. 学会等名 第32回ファジィシステムシンポジウム
4. 発表年 2016年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

—

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----