

人間情報学会

Academy of Human Informatics



第37回 人間情報学会講演集

2020年12月10日

オンライン開催

人間情報学会 オーラルセッションプログラム

◆日時：2020年 12 月 10 日（木）11:00-13:30

◆場所：ウェブ会議システムにて

.....
発表 10 分、質疑 4 分、切り替え 1 分

[11:00 - 12:00] オーラルセッション 1（15 分×4 件）

1. 時空間敵対的生成ネットワークを用いた動画像における異常行動検知
橋本慧志, 工藤謙一, 高橋孝幸, 梅田和昇
2. セマンティックセグメンテーションに基づく線分の属性判定と画像をクエリとする位置
推定手法への応用
渡瀬渉太, 高橋淳二, 余永
3. ナッジ理論を応用した AI による省エネ Web サービスと人間活動への影響
岩崎哲, 阪井遼, 原祐太, 益川陽平
4. 文献の特徴語から抽出される看護倫理の特性の考察-テキストマイニングを用いて
平木早紀, 川原靖弘

[12:00 - 12:10]休憩

[12:10 - 12:55]オーラルセッション 2（15 分×3 件）

5. 複数人対話での共通基盤構築における非言語行動因子および社会的対話行動
因子の分析
山口留実, 渡辺 巧登, 桑原多瑛, 古谷優樹, 光田航, 石井亮, 東中竜一郎, 高汐一紀
6. 心拍数からの勉強中の集中, 眠気, 疲労状態検出手法の検証
川崎勇佑, 横窪安奈, ロペズギョーム
7. 感情制御を伴うテキスト対話システムの実現
片山晋, 米澤拓郎, 大越匡, 中澤仁, 河口信夫

[12:55 - 13:00]休憩

[13:00 - 13:30]企業セッション（15 分×1件）

1. アフターコロナのスマートオフィスを支える無線通信環境の仮想空間における検証技
術活用の可能性
コクヨ株式会社 坂口綾華
2. 赤ちゃん中心設計を目指した育児用品の開発
ニューウェルブランズ・ジャパン合同会社 日下理

時空間敵対的生成ネットワークを用いた 動画画における異常行動検知

橋本 慧志¹ 工藤 謙一² 高橋 孝幸³ 梅田 和昇⁴

¹ 中央大学大学院理工学研究科 〒112-8551 東京都文京区春日 1-13-27

² 中央大学研究開発機構 〒112-8551 東京都文京区春日 1-13-27

³ プリマハム株式会社開発本部 〒300-0841 茨城県土浦市中向原 635

⁴ 中央大学理工学部 〒112-8551 東京都文京区春日 1-13-27

概要

近年、日常生活や製造業の現場等における深層学習を用いた動画画の異常を捉える試みが検討されている。特に、敵対的生成ネットワーク (GAN: generative adversarial networks) の活用が盛んであるが、非効率的であることや、不安定性が課題となっている。本稿では、時空間敵対的生成ネットワークを用いた教師なし学習による新たな動画画異常検知手法を提案する。提案手法はフレーム予測型のモデルであり、U-Net Discriminator の出力を活用するため、効率的かつ高精度な異常検知が可能である。

キーワード: 教師なし学習 敵対的生成ネットワーク 異常検知

1. 序論

近年の深層学習の発展に伴い、日常生活や製造業の現場等における監視カメラ画像の異常をとらえる動画画異常検知に関する研究が盛んに行われている[1-7]。動画画異常検知では主に時空間ネットワーク (STN: spatio temporal networks) を用いた手法[1-4]と、appearance 特徴と motion 特徴に分離してモデル化する手法[5-7]の 2 つに大別される。さらに、GAN (Generative Adversarial Networks) の活用が盛んであり[5-7]、異常検知の精度向上に貢献している。しかし、これら既存手法の多くに共通する課題として以下 3 点が挙げられる。1 つ目は非効率性である。STN を用いた手法は推論時に直近のフレームのみで異常度を計算する場合が多い。また、GAN を用いた手法の多くは推論時に Discriminator を無視する[8]。2 つ目は不安定性である。敵対的学習は一般的に不安定であり、手法の再現性に課題がある[9]。3 つ目はノイズの問題である。motion 特徴として用いる optical flow にはノイズがしばしば混入するので、性能に悪影響が及ぶ。

本稿では、以上の背景を踏まえ、時空間敵対的ネットワークを用いた新たな動画画異常検知手法を提案する。本稿の貢献は次のとおりである。

- フレーム予測型の時空間敵対的生成ネットワークを構築し、効率的な手法を確立する。
- GAN を用いた手法の課題である不安定性について、安定化手法の実装により改善する。
- U-Net Discriminator の高レベル特徴の活用により、高精度な異常検知が可能となる。

以下では、最初に提案手法のフレームワークを示す。さらに、検証実験に関して示し、最後に結論と今後の展望を述べる。提案手法の概要を Fig. 1 に示す。

2. 提案手法

提案手法では、動画画を効率的にモデル化するために、フレーム予測型の時空間敵対的生成ネットワークを用いる。また、敵対的学習の不安定性を改善するために安定化手法を用いる。さらに、U-Net GAN[10]を参考に、Discriminator に U-Net を使い、その出力を融合することで、高精度な異常検知手法を確立する。

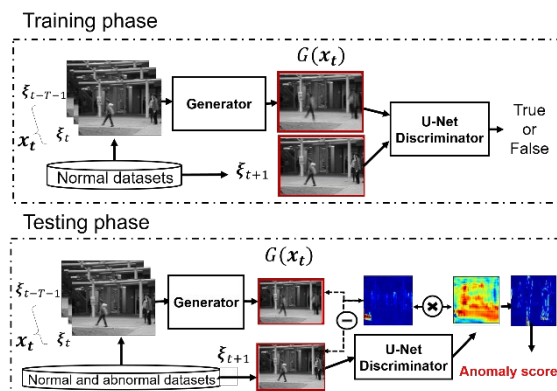


Fig. 1 提案手法の概要

2.1 時空間敵対的ネットワーク

モデルの詳細を Fig. 2 に示す。我々のモデルは、Generator, Discriminator の 2 つから構成されている。Generator は Encoder と Decoder からなる。Encoder は入力の動画画から畳み込み層と Convolutional-LSTM[11]を用いて特徴を抽出し、Decoder は Encoder により抽出された特徴を用いて逆畳み込み層により動画画の 1 フレーム先のフレーム画像を予測する。Generator の構造は、Lee ら[3]が採用したモデルを参考に、予測型のモデルへと拡張した。Discriminator には U-Net を採用し、真値か Decoder により予測されたフレーム画像かをピクセルレベルで識別する。

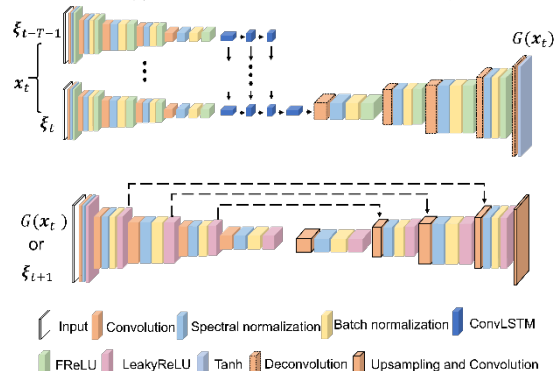


Fig. 2 Generator (上) と Discriminator (下) の詳細

2.2 敵対的学習

我々の時空間敵対的生成ネットワークは、以下の最小最大化問題を最適化する。

$$\min_G \max_D L_{D_{enc}} + L_{D_{dec}} + L_{consist} + L_{recon} \quad (1)$$

ここで、

$$L_{D_{enc}} = -\mathbb{E}_{\xi \sim p_\xi} [\log(D_{enc}(\xi_t))] - \mathbb{E}_{x \sim p_x} [\log(D_{enc}(1 - G(x_t)))] \quad (2)$$

$$L_{D_{dec}} = -\mathbb{E}_{\xi \sim p_\xi} \left[\sum_{i,j} \log [D_{dec}(\xi_t)]_{i,j} \right] - \mathbb{E}_{x \sim p_x} \left[\sum_{i,j} \log [D_{dec}(1 - G(x_t))]_{i,j} \right] \quad (3)$$

なお、 G はGenerator、 D, D_{enc}, D_{dec} はDiscriminatorとそのEncoder module, Decoder moduleである。 L_{recon} は真値と予測結果間の予測損失、 λ はそれに対する重みである。入力 x_t はある時刻 t における固定長 T を有する部分時系列 $\xi_{t-T-1}, \xi_{t-T-2}, \dots, \xi_t$ から構成され、 ξ は各フレーム画像である。 $[D_{dec}(\xi_t)]_{i,j}$ と $[D_{dec}(G(x_t))]_{i,j}$ は、ピクセル (i, j) におけるDiscriminatorの出力である。 $L_{consist}$ は[10]で導入されたCutmix[12]ベースのConsistency Regularizationである。さらに、提案手法では安定化手法として、TTUR[13]、Self-Attention[14]、Spectral normalization[15]を導入する。

2.3 異常検知

入力 x_t に対する異常度 $a(x_t)$ を以下に示す。

$$a(x_t) = ||\xi_{t+1} - G(x_t)||_{map} \quad (4)$$

$$map = (A - D(\xi_{t+1}))|F_2(\xi_{t+1}) - F_2(G(x_t))| \quad (5)$$

map はU-Net Discriminatorからの出力 $D(\xi_{t+1})$ と、成分がすべて1で $D(\xi_{t+1})$ と同じサイズの行列 A との差のマップ $A - D(\xi_{t+1})$ に、 D_{enc} の畳み込み第二層の中間表現 F_2 の差を融合したものである。我々がこのU-Net Discriminatorからの出力を融合するのは、異常な画像が入力されたときに、モデルがその領域を偽と判別することが期待できると考えるからである。異常算出に用いる最終的なスコア $S(x_t)$ は式(6)を用いて正規化することで求められる。

$$S(x_t) = \frac{a(x_t)}{\max(a(x_{1..m}))} \quad (6)$$

ここで、 m はテストデータの総数である。

3. 検証実験

本稿では、動画画像異常検知において一般的な公開データセットであるAvenue[16]を用いて実験を行った。Fig. 3にAvenueの例を示す。Avenueは16clipの学習データ、21clipのテストデータからなる。図に示すように、定点の監視カメラ画像を収録しており、正常データは通常の速度で歩行する様子、異常データは走る、荷物を投げる等の異常な様子が収録されている。このAvenueに対して、Frame-levelのReceiver Operating Characteristic (ROC) 曲線に対するAUROCによる定量的評価を行った。表1にAvenueに対する結果を示す。なお、Ours only STNは map による重みづけをしない場合の結果である。AUROCを用いた評価により、提案手法の有効性を確認した。また、 map を融合することによる有効性も確認した。



Fig. 3 Avenue の例. 正常例 (上) と異常例 (下)

Table 1 Avenue の結果

	Ours	Ours only STN	STAN[3]	Future frame prediction[7]	SRNN[1]
AUROC↑	0.897	0.868	0.872	0.851	0.817

4. 結論

本稿では、動画画像を効率的にモデル化するため、フレーム予測型の時空間敵対的生成ネットワークを提案した。敵対的学習の不安定性を改善するため各種安定化手法を用い、U-Net Discriminatorの出力を融合することで効率的な異常検知手法を確立した。Avenue データセットにおいて、既存手法を上回る結果を確認した。今後の展望として、異常の局所的な検知のためのフレームワークを導入する。

文 献

- [1] Weixin Luo, et al., “A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework,” ICCV, 2017.
- [2] Weixin Luo, et al., “Remembering history with convolutional LSTM for anomaly detection,” ICME, 2017.
- [3] Sangmin Lee et al., “STAN: Spatio-Temporal Adversarial Networks for Abnormal Event Detection,” ICASSP, 2018.
- [4] 橋本 慧志, 工藤 謙一, 高橋 孝幸, 梅田 和昇, “GANを活用した動画画像異常検知手法の構築と労働災害防止へ向けた応用の検討”, 精密工学会画像応用技術専門委員会サマーセミナー, 2020.
- [5] Mahdyar Ravanbakhsh, et al., “Abnormal Event Detection in Videos using Generative Adversarial Nets,” ICIP, 2017.
- [6] Hung Vu, et al., “Robust Anomaly Detection in Videos Using Multilevel Representations,” AAAI, 2019.
- [7] Wen Liu, et al., “Future Frame Prediction for Anomaly Detection – A New Baseline,” CVPR, 2018.
- [8] Mohammad Sabokrou, et al., “AVID: Adversarial Visual Irregularity Detection,” arXiv, 2018.
- [9] Muhammad Zaigham Zaheer, et al., “Old is Gold: Redefining the Adversarially Learned One-Class Classifier Training Paradigm,” CVPR 2020.
- [10] Edgar Schonfeld, et al., “A U-Net Based Discriminator for Generative Adversarial Networks,” CVPR, 2020.
- [11] Xingjian Shi, et al., “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” NeurIPS, 2015.
- [12] Sangdoo Yun et al., “Cutmix: Regularization strategy to train strong classifiers with localizable features,” ICCV, 2019.
- [13] Martin Heusel, et al., “GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium,” NeurIPS, 2017.
- [14] Han Zhang, et al., “Self-Attention Generative Adversarial Networks,” ICML, 2019.
- [15] Takeru Miyato, et al., “Spectral Normalization for Generative Adversarial Networks,” ICLR, 2018.
- [16] Cewu Lu, et al., “Abnormal Event Detection at 150 FPS in MATLAB,” ICCV, 2013.

セマンティックセグメンテーションに基づく線分の属性判定と画像をクエリとする位置推定手法への応用

渡瀬 渉太¹ 高橋 淳二² 余 永²

¹ 鹿児島大学工学部機械工学科 〒890-0065 鹿児島市郡元 1 丁目 21-40

² 鹿児島大学大学院理工学研究科機械工学専攻 〒890-0065 鹿児島市郡元 1 丁目 21-40

概要

背景と目的：我々が提案している位置推定手法(Universal Map: UMap)では、クエリと DB の線分画像の類似度を評価して位置推定を行っている。しかし、線分情報のみを用いた類似度評価では異なる位置が推定される場合がある。本研究では、新たなマッチング手法を提案し位置推定精度の向上を目指す。

方法：線分に壁やドアなどの属性を付与した属性付線分画像を用意し、属性を考慮した類似度評価を行うことで、位置推定精度を向上させる。Semantic Segmentation (SS)により画像を構成する属性の領域を判別できる。これを用いて、クエリ画像の線分に適切な属性を付与した属性付線分画像を生成する。

結果：SS の精度は学習させるデータに大きく影響する。また、どのような画像に実行するかによって精度が変わってくる。属性付線分画像の生成においては、線分にはセグメンテーション画像に対応して適切な属性を付与することができた。属性付線分画像を用いての位置推定精度の評価はまだ行っていない。

考察：SS の精度を 100%にすることは難しいと思われる。しかし、学習データの数を増やしより多くの種類の建物のデータを学習させれば、SS の精度は向上できると考えられる。また、十分な SS の精度を出せば、線分に正しい属性を与えることができると考えられる。

結論：UMap の位置推定精度向上に向けて属性付線分画像を生成する仕組みを構築した。具体的には SS の実装と属性付線分画像生成プログラムの作成を行なった。

キーワード：Semantic Segmentation ビジョンを用いた位置推定

1. はじめに

我々は、屋内で移動ロボットや IoT デバイスを位置推定するクラウド型位置推定インストラクチャシステム Universal Map (UMap)を提案している[1]。UMap の位置推定手法を図 1 に示す。UMap ではクライアントが撮影した写真を線分化して得られるクエリ画像と事前マップより生成する座標付き DB 画像群との類似度を評価し、クエリ画像の撮影視点を推定する。UMap の利点は、クライアントにおける演算負荷が小さいこと、環境に手を加える必要がないこと、建物の CAD データさえあれば容易に導入できることなどがある。

UMap の類似度評価では、クエリ画像と DB 画像の線分情報のみを利用するため、まったく異なる位置の画像同士であっても線分が重なり合えば類似度が大きくなり誤った位置が推定される（以下、これを誤マッチングと呼ぶ）。また、クエリ画像に掲示物や人などの DB 画像に存在しないものが映り込むと、それらが余計な線分として検出され、誤マッチングを起こすという問題がある。この問題に対し、幾何学的情報でしかない線分にその出どころである壁やドアなどの属性情報を付与し、属性を同一とする場合のみ加算されるよう類似度評価関数を設計することが有効である。本研究の目的は、撮影により得られたクエリ画像に対して領域分割を行い、その結果に基づき線分に属性を付与するアルゴリズムの開発である。

2. Semantic Segmentation

クエリ画像の線分に適切な属性を与えるために Semantic Segmentation [2]を用いる。これは、画像内の全画素にラベル付けを行う画像アノテーションの手法である。教師データを学習させることで、与えられたラベルに応じて色付けされたセグメンテーション画像を生成できる。

主に鹿児島大学工学部機械工学 2 号棟で撮影した画

像データから属性ラベル付きの学習データを 515 個作成し、学習回数 12000 回でニューラルネットワーク (NN)を学習させた。学習させた NN について、137 個のテストデータによる検証結果を表 1 に、2 枚の画像の属性判定を行った結果を図 2 に示す。表 1 の mIoU (mean Intersection over Union)は SS の評価指標である。図 2 上側は機械工学 2 号棟で撮影した画像であり、精度よく判定できていることが確認できる。一方で、下側の青山学院大学 O 棟内で撮影した画像は、あまりいい精度ではない。例えば、床に反射して写る壁の領域を床ではなく壁と判定するなど精度向上に向けての改善の余地がある。失敗した箇所を補正するような学習データを追加することで精度を向上させることが可能であると考えられる。

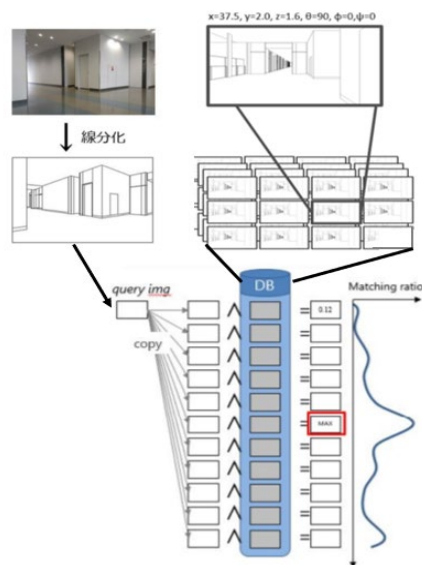


図 1 UMap による位置推定の概要

表 1 Semantic Segmentation の mIoU

属性	全体	壁	床	天井	エレベータ
mIoU(%)	45.0	73.5	81.4	68.3	65.8
属性	ドア	窓	階段	掲示板	消火栓
mIoU(%)	46.2	60.3	62.6	43.5	66.5

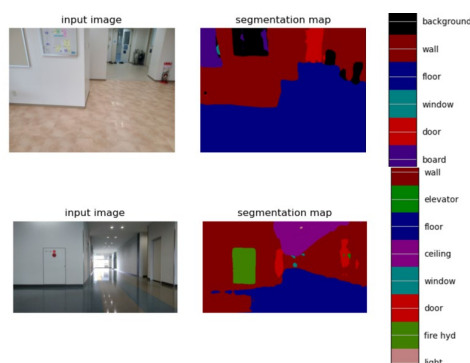


図 2 Semantic Segmentation の実行結果

3. 属性付線分画像生成プログラム

3.1 属性付線分画像生成の必要性

SS は分割した領域の内部のピクセルをラベリングするものである。一方、UMap では線分画像を用いるため、分割した領域の境界線となるピクセルにラベリングをする必要がある。そのため、SS で得られた画像を基に線分に適切な属性を付与する新たなアルゴリズムを開発する必要がある。

3.2 線分属性の決定方法

図 3 を用いて線分に付与する属性の決定方法を解説する。左図はクエリ画像と検出される線分（白線）を重ねたもの、右図はセグメンテーション画像であり属性の色は赤がドア、茶が壁、紫が床に対応している。左図におけるドアを構成する右の線分について、どのように付与する属性を決定するかを考えたい。初めに、対象の線分（右図実線）が何と何の属性の境界線なのかを判別する必要がある。そのため垂直な二方向にずらした平行線分（破線）を参照する。平行線分について、線分に重なる属性の数を集計し、それぞれの線分で最多属性を求める。図の場合は、右の線分では壁、左の線分ではドアが最多属性となる。得られた属性のうち、定められた優先順位の高い方を対象の線分に付与する。属性の優先順位は、高い順に DB に存在しないもの（床や天井など）、消火栓、掲示板、エレベータ、窓、階段、ドア、壁、とした。図においては、線分にはドア属性が付与される。この作業を各線分で行い、すべての線分に属性を付与する。



図 3 線分属性の決定方法

3.3 属性付線分画像生成プログラム

図 4 に示すように、クエリ画像と SS により生成さ

れるセグメンテーション画像を入力として属性付線分画像を生成するプログラムを作成した。比較として図 4 右上に線分検出のみで生成される線分画像を掲載した。プログラムの具体的な内容は以下のとおりである。

クエリ画像に FLD(Fast Line Detector)を用いて線分検出を行い、検出した各線分の端点の座標の組を取得する(1)。続いて、セグメンテーション画像上で平行線分を参照するため、元の線分の端点の座標より幾何学的に平行線分の端点の座標をそれぞれ取得し(2)、端点の座標を基に各線分を構成する全ピクセルの座標を取得する(3)。これについては、 $x \cdot y$ 軸に平行な線分は連続的なピクセルの並びから容易に求められ、その他の斜めの線分についてはブレゼンハムのアルゴリズムを用いることで線分を構成するピクセルの座標を求めることができる。取得した座標を用いてセグメンテーション画像上で平行線分に重なる色（属性）をそれぞれ集計し、最多属性を 2 つ取得する(4)。属性付線分画像では、付与された属性は画像の輝度を用いて表現される。よって、2 つの属性のうち優先順位の高い方に対応する輝度で元の線分を描画する(5)（ただし DB 画像に存在しない属性の線分は描画しない）。1~5 を各線分で行うことで属性付線分画像を生成する。

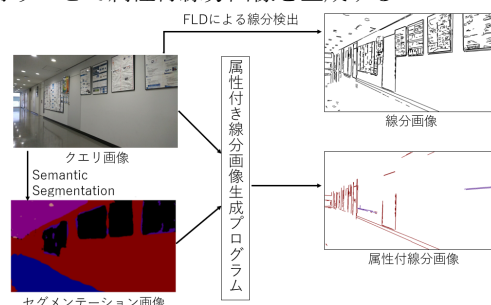


図 4 属性付線分画像生成の流れ

4. 結論

本研究では、UMap の位置推定精度向上に向けて線分の属性を考慮した類似度評価を行うため、属性付線分画像を生成するプログラムを作成した。属性付線分画像の生成の際は SS を活用した。SS の精度について、学習データを増やすことで向上しているが、全体の mIoU では 45.0%と十分な精度は出ていない。そのため、より多くの種類の建物のデータを学習させ、各属性の mIoU が 80%を超えるほどに高い精度を出せるよう改善していきたい。属性付線分画像を生成するプログラムについては、検出した線分に対してセグメンテーション画像に対応した適切な属性を与えることができた。今後は、SS の精度向上や属性付線分画像を用いての評価実験を行う。

文 献

- [1] Junji Takahashi, Masato Kawabe, Seiya Ito, Naoshi Kaneko, Wataro Takahashi, Toshiaki Sakamoto, Akihiro Shibata, and Yong Yu, "Image-retrieval Method Using Gradient Dilation Images for Cloud-based Positioning System with 3D Wireframe Map," Sensor and Materials, 32(2):611-23, 2020.
- [2] Chen LC., Zhu Y., Papandreou G., Schroff F., Adam H., "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," the European Conference on Computer Vision (ECCV), 2018, pp. 801-18.

ナッジ理論を応用した AI による省エネ Web サービスと人間活動への影響

岩崎 哲¹ 阪井 遼² 原 祐太² 益川陽平²

¹ 株式会社アイ・グリッド・ラボ 〒102-0083 東京都千代田区麹町 3-7-4

² 株式会社アイ・グリッド・ソリューションズ 〒102-0083 東京都千代田区麹町 3-7-4

概要

事業所の電力使用量を AI で予測し、電力ロードカーブに応じた省エネ活動のレコメンドを表示する Web サービスを開発した。行動経済学のナッジ理論を応用し、各事業所の従業員に対して、省エネ活動が自発的に推進される仕組みを構築している。AI とナッジによる省エネ活動のレコメンドが各事業所の従業員の意識や活動に対して与える影響と、当該事業所における実際の省エネ効果について評価した。

キーワード： 省エネルギー AI ナッジ理論

1. 背景と目的

昨今、「SDGs」「ESG 投資」「EV シフト」といった、経済活動の中で環境問題を論じるワードがメディアで喧伝され、企業活動における環境対策が社会的に求められている。官民含め「脱炭素化」の動きは徐々に始まってきているが、2050 年のカーボン・ニュートラルの実現にはまだまだ課題が多い。企業活動においても、自社で何を行えば省エネルギーにつながるのかを十分に認識し実行できている企業は少ない、というのも、省エネ設備を導入するだけでは不十分であり、自社の従業員の行動変容が必要だからである。

そうした省エネルギーにおける行動変容については、環境省主導の日本版ナッジ・ユニットなどの事例[1]はあるものの、企業活動という観点で十分に分析された例は少ない。

そこで本研究では、スーパーマーケットの店舗において、AI と行動経済学のナッジ理論[2]を応用して、適切な省エネアクションをレコメンドするシステムを利用することで、各店舗の従業員に対して自発的な行動変容を促し、企業活動における省エネルギーに寄与できることを明らかにする。

2. 前提

当社(株)アイ・グリッド・ソリューションズは創業以来、顧客企業のエネルギーマネジメント、とりわけ電力に着目して省エネコンサルティングを行ってきた。そうした経緯から、これまでのエネマネ事業の中で培った省エネノウハウと電力ビッグデータを AI に落とし込む形で「エナッジ®」(以下「本サービス」)を開発し、2018 年 6 月にリリースした[3][4]。本サービスは、各需要地点の電力使用量を AI で予測する(図 1)とともに、予測した使用量に応じた省エネ活動の提案までを行う。

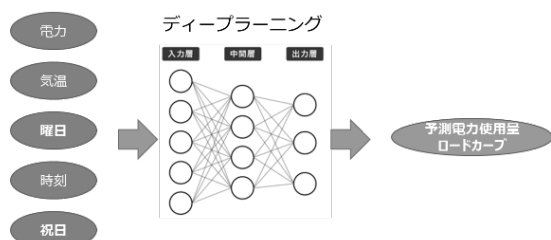


図 1 本サービスの AI 予測

また、AI 予測値は 30 分ごとの値として算出されており、時間帯ごとのロードカーブとして画面上で表示されている。この予測は、メインの予測値に加え、上振れ値と下振れ値を算出して予測に幅を持たせたものとなっている(図 2)。

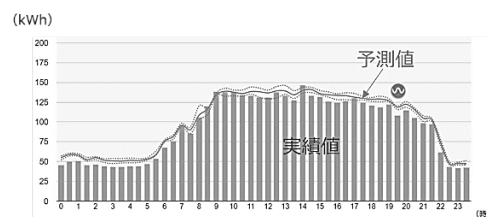


図 2 電力使用量の予測値と実績値

加えて、本サービスの省エネ活動提案はナッジ理論を用いた UI を採用しており、ユーザーのモチベーションを喚起する要素を盛り込むなど、人手不足の中でも適切な省エネが可能となる設計となっている(図 3)。



図 3 本サービスのユーザー画面

本サービスのナッジ理論を用いた UI の一つとして、AI 予測をもとにシステムが提案した省エネアクションをユーザーが画面で確認する際、身体動作を伴って確認するために「OK ボタン」を押下する UI となっている。この OK ボタンの押下を「活動」と定義する。

3. 提案手法

本研究では、本サービスを利用し活動回数が多くなるにしたがってユーザーに省エネアクションが定着するという行動変容が起き、それにより省エネルギー効果が生じるという仮説を立て、実際の本サービスの顧

客企業のデータを分析した。

3.1 本サービス利用による省エネアクションの定着

スーパーマーケットなどの商業施設は、曜日や時間帯による来客数の多寡はあるものの、日々のオペレーションが安定していれば週間ごとの電力使用量も安定し、ムラの無い店舗運営が可能となる。

この点、本サービスが提案する省エネアクションは、局所的・一時的な内容ではなく、その季節中は常に実行可能なものや年間通じて注意すべきものが数多く含まれている。これは、OK ボタンを押下して省エネアクションを実行することにより、省エネアクションがユーザーに定着し、店舗の電力使用傾向が安定化することを企図している。

そこで、活動回数の増加に伴って省エネアクションが従業員に定着し電力使用状況のムラが小さくなっていくかを、1週間単位 of データを用いて分析した。

3.2 分析手法の詳細

本サービスの実顧客のうち、2019年11月から2020年10月までの期間、定常的に本サービスを利用していったスーパーマーケット(48店舗)を分析対象とした。

次に、30分ごとのAI予測値と使用電力量実績値との関係を以下のように定義し、計測した。

- a) num_actions : 本サービスの画面上で OK ボタンを押下した活動回数。
- b) upper_count : 使用電力量実績値が予測上振れ値よりも大きくなった期間(30分単位)の個数。
- c) lower_count : 使用電力量実績値が予測下振れ値よりも小さくなった期間(30分単位)の個数。
- d) expected : 予測の上振れ値・下振れ値の範囲内に使用電力量実績値が収まった期間(30分単位)の個数。
- e) target : 予測の上振れ値・下振れ値の範囲内、もしくは下振れ値を下回った期間(30分単位)の個数。
「エネルギー消費適切度」
$$\text{target} = \text{expected} + (\text{lower_count} - \text{upper_count})$$
- f) diff : 使用電力量予測値と実績値との差(kWh)。
「予実差」

こうした定義と計測の上で、OK ボタンを押下した活動回数とエネルギー消費適切度との間、および、エネルギー消費適切度と予実差との間の相関関係を分析した。

4. 分析結果と考察

4.1 活動回数とエネルギー消費適切度との相関関係

実店舗におけるデータ分析の結果、活動回数(num_actions)とエネルギー消費適切度(target)との間に正の相関があることがわかった(図4. ピアソンの積率相関係数: 0.434)。

このことから、店舗において本サービス上での活動回数が多いほど、使用電力量実績値がAIの予測した使用電力量の範囲内に収斂され、想定外の実績を記録する回数が減ると評価できる。

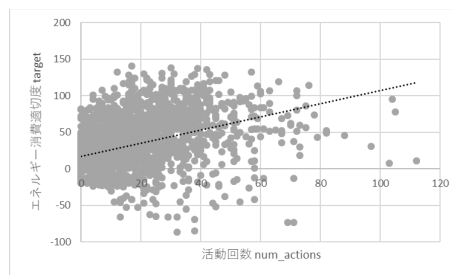


図4 活動回数とエネルギー消費適切度の相関関係

つまり、本サービスの利用を通して、各店舗の日々の電力使用におけるムラが小さくなり、無駄なエネルギー消費が抑制されることが分かった。

4.2 エネルギー消費適切度と予実差との相関関係

次に、エネルギー消費適切度(target)と予実差(diff)との相関分析から、両者に負の相関があることがわかった(図5. ピアソンの積率相関係数: -0.362)。

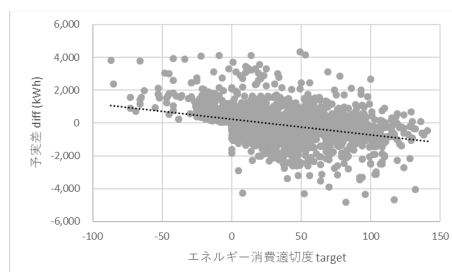


図5 エネルギー消費適切度と予実差の相関関係

すなわち、エネルギー消費適切度が高まり店舗のオペレーションが安定することは、無駄なエネルギー消費の阻止に留まらず、標準的な使用電力量として過去の使用量実績の傾向を元にAIが予測した値よりも少ないエネルギー消費での店舗運営が可能といえる。

5. 結論・今後の展望

本研究では、実店舗でのデータ分析の結果から、ナッジ理論を応用したAIによる省エネWebサービスを利用することで、ユーザーの行動変容を促し、電力使用におけるムラが小さくなり無駄なエネルギー消費が抑制される効果や、標準的な使用電力量としてAIが予測した値よりも少ないエネルギー消費での店舗運営を可能とする効果が得られることが示唆された。

今後の展望として、AI・IoTと連携した空調等設備の自動制御による店舗環境のアダプティブな制御や、ウェアラブル機器との連携による人間の快適性まで含めたエネルギー制御の全体最適化などを検討している。

文 献

- [1] 環境省「日本版ナッジ・ユニット(BEST)について」 <http://www.env.go.jp/earth/ondanka/nudge.html>
- [2] Thaler, Richard H., and Cass Sunstein., "Nudge: Improving Decisions About Health, Wealth, and Happiness." New York: Penguin, 2009 (訳書『実践行動経済学』遠藤真美訳, 日経BP社, 2009年)
- [3] エナナッジサービスサイト <https://enudge.igrid.co.jp/>
- [4] ネイチャーインタフェース 第76号, 2019年8月, 24-25頁

看護倫理学会誌の特徴語から抽出される看護倫理の特性の考察—テキストマイニングを用いて—

平木早紀¹ 川原靖弘¹

¹ 放送大学大学院文化科学研究科 〒261-8586 千葉県千葉市美浜区若葉2-11

概要

看護倫理は従来の規範的な倫理理論よりも臨床実践を重視するという特徴をもつが、構成概念が曖昧で理解が難しいという指摘がある。本研究では、構成概念を可視化することで、看護教育の際に看護倫理の特徴について説明するための一助となることを目指す。方法として、『日本看護倫理学会誌』と『生命倫理』に投稿された論文188報を対象にテキストマイニングを行った。論文1報を単位とし、文書同士が意味のまとまりをもつ28のクラスターに分類した。クラスターとジャーナル内の単語について対応分析を行った結果、看護倫理の特徴は「実践重視の姿勢」と「倫理的感受性の涵養」であることが明らかになった。また、『生命倫理』との比較により、法律や周辺制度の考察が少ないこともわかった。本研究により、質的研究で説明されてきた看護倫理の独自性は数量化によっても説明可能であると示された。また、ジャーナルの比較という先行研究の手法を踏襲し、看護倫理の研究が現在は射程としていない領域について指摘することができた。

キーワード： テキストマイニング 看護倫理 対応分析

1. はじめに

医療現場において専門職者は高度な倫理観を求められており、職能団体が定めた倫理綱領が存在する。実際の臨床実践の際には、職業規範的な側面をもつ綱領のみならず、BeauchampとChildressによる医療倫理の4原則「自律尊重」「無危害」「善行」「正義」[1]や、看護倫理の構成概念である□患者のニーズを代弁する行為「アドボカシー」、□患者の健康に対して責任をもつこと「責務と責任」、□多職種間で連携を図ることとして「協働」、□患者-看護師関係に基づいて患者を尊重する態度「ケアリング」[2]がある。この複雑な意義をもつ臨床実践の内容を学生に伝えるには困難があるとされ、また教育者側の教材や情報の不足も指摘されている[3]。

2. 先行研究との比較

2-1. 質的研究

概念分析のために従来使われてきた手法は、文献調査やインタビューといった質的研究法である。しかしながらこれらの手法は手作業で分類・再構成を行うため労力と時間を要することや、得られたデータの解釈を文献間で比較しづらいというデメリットが存在する。

2-2. テキストマイニング

質的研究法のデメリットを補うために、近年では自然言語処理技術を用いたテキストマイニングが選択されることがあり、看護研究にも応用されている[4]。テキストマイニングによる文献比較は様々な分野で行われており、鈴木らは経営関連の学会の文献抄録を分析し、類似学会の棲み分けや投稿年次による学術誌の特徴語の変化を可視化した[5]。本稿はテキストマイニングを用いて『日本看護倫理学会誌』と『生命倫理』を比較することで、看護倫理で用いられる特徴的な概念を可視化することを試みる。

3. 方法

3-1. 分析対象

国内で医療倫理を扱う学会は「看護倫理学会」「生命倫理学会」「医学哲学・倫理学会」「臨床倫理学会」「臨床死生学会」等が存在するが、ジャーナルがオープンアクセスであることや、言語が日本語であり質的に比較しやすいことから、日本看護倫理学会『日本看護倫理学会誌』と日本生命倫理学会『生命倫理』を選択した。文献比較にあたって1報あたりの文字数・構成を均質にするため、両ジャーナルに掲載されている原著

論文・短報・報告論文を使用した。また、学会同士の文献数を同等にするため、看護倫理学会の全年度分(2010-2020)90報に対して生命倫理学会2014-2019年度分98報を使用した。

3-2. 前処理・単純集計

分析にはKH coder ver.3[6]を使用した。各文献のうち、解析には本文のみを使用し、謝辞・利益相反・脚注・参考文献リストは除外した。また、研究倫理に関する文献が存在しているため、各文献内の研究対象に対する「倫理的配慮」の項が文献分類に影響を与える可能性があると考え、事前に除外した。次に、形態素解析器ChaSenにより全文の語句の品詞解析を行い、分析対象外となる数詞・助詞・記号を除去した。「看護師」や「意思決定」などのテクニカルタームは、出現頻度の高い順から複合語として検出した。前処理した語彙のうち、分析に使用した総語数は432,794語、語彙の種類(異なり語数)は16,366語であった。

3-3. 文獻クラスタリング

文献同士の類似性を確認するため、全文中100回以上出現した809種類の語を用い、TF-IDF(Term Frequency - Inverse Term Frequency)値とCosine類似度に基づいてクラスター分類を行った。クラスター間の距離計算はWard法を用い、併合水準を確認しながら28のクラスターを作成した。それぞれのクラスターがもつ意味のまとまりをもとに、手作業でクラスター名を付与した(図2・図3内で赤字表記)。

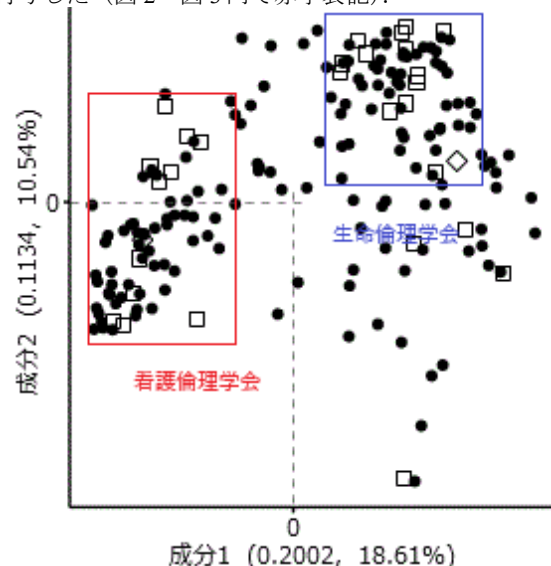


図1. 対応分析 全体図(原点付近を拡大したもの)

語彙 (●) とクラスター名 (□) は省略している。
赤枠 (看護倫理学会を主とした部分) を図2, 青枠内 (生命倫理学会を主とした部分) を図3として切り出した。

3-4. 対応分析

3-3 で形成した文書クラスターとそれぞれの学会を外部変数とし, 全文献中で計 300 回以上出現した名詞 160 語を用いて対応分析を行った。対応分析では, 文書クラスターと名詞をクロス集計したデータ行列を使い, クラスターの数と同数の次元を作成した。図 1 の成分 1・2 はそれぞれもっとも寄与率の高い次元を軸として用いたものであり, 原点からの方向に数値上の情報は無い。軸ラベルの数値は, それぞれの成分の固有値の総和と寄与率を表している。

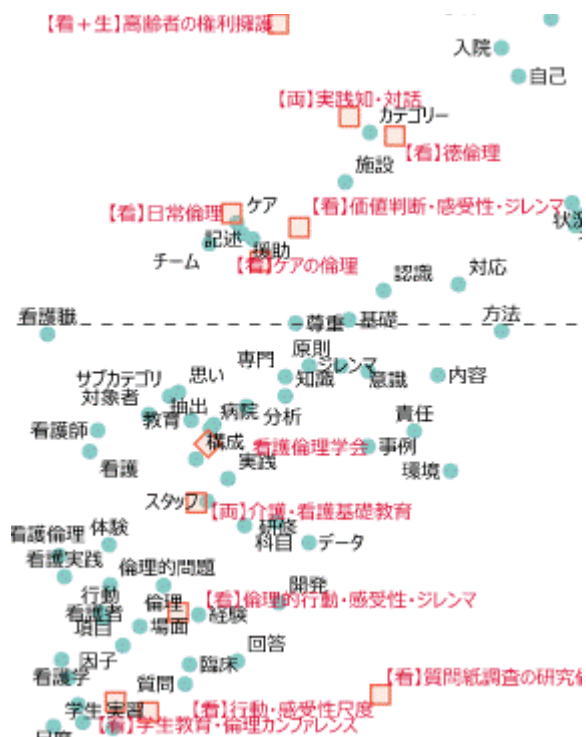


図 2. 図 1 赤枠部分の拡大図

【看】看護倫理学会を主とした文献, 【生】生命倫理学会を主とした文献, 【両】両学会同数程度, 【A+B】A学会文献にごく少数のB学会文献が加えられていることを示す。



図 3. 図 1 青枠部分の拡大図

3-5. 著者属性

著者背景を推定するため, 文献内の所属表記・Researchmap・機関公式サイト等を検索のうえ背景属性を振り分けた。看護倫理学会では医学・看護・介護福祉の国家資格ないし臨床経験をもつと思われる著者による文献が 90 報中 83 報 (92.22%) であり, 生命倫理学会では 98 報中 47 報 (47.96%) であった。非臨床系著者の専門分野は, 人文・社会科学系 (法学, 哲学・倫理学, 宗教学等) であった。

4. 考察

4-1. 看護倫理の独自性

図 2 上部では「徳倫理」「実践知」「日常倫理」「ケアの倫理」といった看護倫理の理論的考察の文献が近く付置された。このことから, 看護倫理における実践重視という特徴と, よい看護師という職業規範 (徳) を含んだ観念が結びつきをもって記述されていることがわかる。また, 図 2 下部では, 倫理教育と倫理的問題に対する感受性について言及した文献が多く, 看護倫理の教育が感受性の涵養と結びついていることが伺える。

4-2. 生命倫理と看護倫理の関係性

図 3 では, 法制度や終末期医療・安楽死の問題や公衆衛生上の義務について議論が行われていることがわかる。看護倫理学会からの文献は, 「意思表示困難者の支援」のみが生命倫理学会と同程度付置された。この傾向について国際的な動向と比較すると, Barbara ら[7]による報告が一致する。看護倫理の研究において, 安楽死に対する実践的立場からの考察が行われている文献は 2019 年時点では存在しない。このことから, 看護倫理の特徴である実践知や日常倫理は法や制度との間に溝のある概念であることが予想される。

4-3. 看護倫理の展望

看護倫理の独自性が「実践の重視」と「倫理的感受性の涵養」にあることは, 倫理教育にあたって教育者が理論の概説をするだけでなく, ケーススタディや実習の振り返りから学ぶ方法が有効であることを示唆する。また, 現在の看護倫理研究の弱みを, 法制度や公衆衛生学的な観点からの報告の少なさとして発見した。これは, 現在の倫理的な臨床実践が制度変革や改善に結びつきにくい状況であるとも説明できる。今後の看護倫理研究に公衆衛生や医療政策視点のテーマを加えることで, 看護倫理の更なる発展に寄与することができるといえる。

文 献

- [1] 赤林明, “〔改訂版〕入門・医療倫理I”. 勁草書房, pp. 58-72, 2017.
- [2] S.T.Fry, M-J.Jhonstone, “看護実践の倫理【第3版】倫理的意思決定のためのガイド”, 日本看護協会出版会, pp. 49-58, 2010.
- [3] 遠藤由美子, “教育者側に焦点を当てた看護倫理教育に関する研究の動向と課題”, 医療保健学研究 vol3, pp. 125-135, 2012.
- [4] L.Hyeyong, R.Shimotakahara, and A.Fukada, “Impact of differences in clinical training methods on generic skills development of nursing students: A text mining analysis study”, Heliyon vol.5(3), pp. 1-21, 2019.
- [5] 鈴木啓, 大内紀知, “テキストマイニングによる学会の特徴分析”, 経営情報学会 全国研究発表大会要旨集, pp. 79-82, 2016.
- [6] 樋口耕一, “テキスト型データの計量的分析—2 つのアプローチの峻別と統合—”, 理論と方法 19(1), pp. 101-115, 2014.
- [7] P.Barbara, M.Greig, and R.Janke, “Nursing and euthanasia: A narrative review of the nursing ethics literature”. Nursing Ethics vol.27(1), pp. 152-167, 2020.

複数人対話での共通基盤構築における 非言語行動因子および社会的対話行動因子の分析

山口 留実¹ 渡辺 巧登¹ 桑原 多瑛² 古谷 優樹³ 光田 航⁴ 石井 亮⁴ 東中 竜一郎^{4,5} 高汐 一紀³

1 慶應義塾大学大学院政策・メディア研究科 〒252-0882 神奈川県藤沢市遠藤 5322

2 慶應義塾大学総合政策学部 〒252-0882 神奈川県藤沢市遠藤 5322

3 慶應義塾大学環境情報学部 〒252-0882 神奈川県藤沢市遠藤 5322

4 NTT メディアインテリジェンス研究所 〒239-0847 神奈川県横須賀市光の丘 1-1

5 名古屋大学大学院情報学研究科 〒464-8601 名古屋市千種区不老町

概要

複数人対話において、参与者同士は発話内容を互いに理解しているかどうかを確認し、共通基盤構築を行いながら対話を進める。共通基盤構築に関する研究は主に言語的な行動に着目して行われているが、我々は言語的因子のみならず非言語行動因子および社会的対話行動因子が共通基盤構築に影響すると考えている。我々は、共通基盤構築過程における各因子を観測できる課題として旅行計画立案課題を採用し、遠隔対話状況下において動画コーパスの収録を行っている。その目的は、遠隔対話と直接対話におけるこれら行動因子の差異、さらには共通基盤構築や合意形成への影響に関して理解を深めることにあるが、本稿では、予備的分析において観察された社会的関係性による配慮行動や姿勢の同調行動について報告する。

キーワード： 遠隔対話 共通基盤構築 非言語行動 社会的対話行動 HRI/HAI

1 はじめに

複数人対話において、しばしば参与者同士は相手の発話が理解できたことを何かしらの行動で伝えている。例えば、「はい」や「なるほど」という発話を伴う相槌や、発話を伴わず首を縦に振る非言語行動が挙げられる。Clark らはこのプロセスを基盤化 (grounding)、参与者間で共有された認識を共通基盤といい、共同作業を支えるものと述べている[1]。Clark らは、議論や計画立案を進める際、参与者の発話内容に対して、異なる参与者たちがその内容を十分理解したという相互の信念を互いに確立することが対話への貢献であると主張している。また、川端らは『日本語地図課題対話コーパス』を用いた非対面での共同作業における共通基盤の分析を行っている[2]。具体的には、giver が出発地点から目標地点までの経路を非対面状況下で説明し、その結果を follower が紙の地図に描写するというタスクにおいて、共通基盤が構築される過程を分析している。一方、HRI (Human-Robot-Interaction) において、Chai らは対話システムの機能として、ユーザの発話に対する受理や拒否を示す意思表示だけではなく、認識している情報をユーザに伝えるという協調的な動作を実装している。ロボットが協調的な動作を冗長に行い、ユーザとのやり取りを多く重ねることで、認識の誤りが減ったという結果が示されている[3]。これらの既存研究は特定の状況下（例えば、対面や音声のみのコミュニケーションなど）における共通基盤構築の過程を扱っており、報告されている知見が異なる状況化でも有効であるかは明らかではない。特に、対話参与者間の社会的関係性や非言語行動の影響については議論されていない。

我々はこれまで割り込み場面での合意形成過程のコーパスを収集しており、参与者の1人が他2人の対話中に割り込みを行う過程の非言語行動、参与者同士のインタラクションや割り込まれた際の印象から合意形成に与える影響について分析を行った[4]。本稿では、COVID-19 で遠隔コミュニケーションが普及・発達しているという社会的状況を考慮し、遠隔状況下での動

画を用いた対話において、非言語行動及び社会的行動が共通基盤構築に与える影響を明らかにすることを目指す。遠隔対話では、対面とは異なる非言語行動が出現すると考えられ、これらの非言語行動及び社会的行動が共通基盤構築に与える影響を明らかにすることは、今後遠隔コミュニケーションを支援するエージェントやロボットを設計する上で重要だと考えられる。本稿では、予備的な調査として、2者間で共通基盤を構築する協働型合意形成タスクを設定し、そのコーパス収集と分析を行った結果について報告する。

2 動画コーパスの設計

参加者は3名の学生（全て男性）で、3名とも同じ研究室に所属する学生である。参加者をR-3の観測のため年齢差が生じる2組のペアに分けた。収録する動画コーパスでは、合意形成における非言語行動および社会的対話行動が観測できるタスクを設定する必要がある。課題の要件は以下の通りである。

- R-1. 合意形成が瞬間的でなく、議論を積み重ねる必要があること
- R-2. 合意形成の結果を確認でき、タスクの終了が明確であること
- R-3. 対話における非言語行動および社会的行動を観測できること

- (1) 画面共有された地図から集合場所、解散場所を選んでください。
 - (2) 2人で訪れたい場所を3箇所相談して決めてください。
 - (3) 訪問する順番を2人で決めてください。

決まり次第画面共有された地図に決めた場所と道順を書き込んで「終了しました」と伝えて下さい。但し、

 - ・ 12:00～17:00の間でのスケジュールを決める。
 - ・ 共有された地図をメモ代わりにペンや矢印で書き込んで良いものとします。

図1 タスクシート



図2 収録した動画コーパス
(左:収録映像 右:交互に鼻を掻く40秒間のアノテーション)

上記の要件を満たす協働型合意形成タスクとして、旅行計画立案課題を設定した。具体的には、画面共有された地図を用いて2者で訪れたい場所を決めるといタスクである。参加者に対して図1を提示し、タスクの映像収録を行った。参加者の姿勢を分析できる全体映像、カメラオンにて相手の画面にも映る参加者の視線や表情を記録するバストアップ映像、地図が共有されている画面を収録し(図2左)、タスクの終了は地図上に書き込んだスクリーンショットを提出してもらい両者の最終的な合意形成とみなした。

3 動画コーパスの予備的分析

ELANを用いて、取得した動画コーパス3本(カメラオン2本、カメラオフ1本)に対して予備的分析を行った。アノテーションの項目は、発話区間および発話内容と、発話内容を踏まえ共通基盤を構築している箇所やタスクシートの(1)～(3)の回答である最終的な合意形成に伴う視線、手、姿勢の変化(前傾か後傾か)である。対話中の非言語行動および社会的対話行動としての観点から、共通基盤構築に影響を与えていると考えられる参加者の行動に着目すると、以下のような様子が見られた。

発話量の分析を行ったところ、先輩の方が後輩よりも発話量が多いことが明らかになった。要因としては、タスクが始まってから1番最初に相手に話しかける場面や、沈黙が発生したときに先輩が後輩に対して提案を行うなど、積極的に話の切り出しを先輩側から行ったためといえる。一方で5年程度同研究室に所属していた学生間(先輩・後輩関係)では、後輩自身が行きたい観光地を明確に表現していた。対して、研究室に所属している期間が半年である後輩と5年程度の先輩間では、後輩が先輩に対して「この〇〇が気に入ります」等、相手に遠慮するような言動が起こり、結果的に先輩の意見が採択されるといった、社会的関係性に配慮する行動がみられた。

カメラオンの条件下において、お互いの顔を見ずとも仕草がつかれるシーンが幾度と有った。例えば、研究室に所属している期間が半年である後輩と5年程度の先輩間でのタスクでは、40秒間で相手の映る画面を注視すること無くお互い交互に2回ずつ鼻を掻くシーンが見られた(図2右)。この間は後輩が今までの話の内容をまとめていたことから発話量は比較的后輩の方が長くなり、行動の伝染や同調が生じているときは必ずしも年齢は関係なく発話量が均衡に近くなる傾向にあった。これはFrankら[5]の研究でも明らかになっている。遠隔対話下でも「姿勢の同調行動」が起こっている可能性がある。直接対話での同種行動と比較し、その影響と共通基盤構築との関係を詳細分析したい。

また先輩は後輩が表示されているディスプレイを注視する時間がほとんどないが、後輩は先輩の表示されているディスプレイを注視する時間が比較的長かった。沈黙の時間が続いたときに後輩が画面向こうの先輩の様子を伺う様子や、最終的に訪問する箇所・ルートについて先輩の映る画面を見ながら後輩が確認するなど、後輩が先輩の意見や反応を配慮する様子が観察できたことから、共通基盤構築への影響、合意形成への影響に関して、今後詳細を分析したい。

今回は、研究室内での先輩・後輩関係であったが、より参加者間の社会的関係性の差が大きい場合は、社会的立場による圧力が発生し、片方がルートを一方的に指示するような流れで共通基盤が構築される可能性もある。

4 まとめ

本研究では、複数人対話での共通基盤構築に伴う言語行動、非言語行動および社会的対話行動の分析を行うために、旅行計画立案課題を用いて動画コーパスの収録を行った。収録した動画に対して予備的分析を行ったところ、社会的関係性による配慮行動や姿勢の同調行動が共通基盤構築過程にみられた。

今後は、動画コーパスの収録本数を増やしつつ、遠隔対話と直接対話の比較、さらには、カメラのオンオフをはじめとする遠隔対話独特のモダリティの変化が共通基盤構築過程にどのような影響が生じるかを、非言語行動と社会的対話行動の側面からより詳細な分析を行いたい。併せて、これらの分析から得られた知見を、より円滑な対話を実現するためのHRI、HAI設計へ適用する予定である。

文 献

- [1] Clark, Herbert H., Edward F. Schaefer, "Contributing to discourse." *Cognitive science*, 13(2):259-94, 1989.
- [2] 川端良子, 松香敏彦, 土屋俊, "地図課題対話における共有信念更新のメカニズム." *認知科学* 24(2): 153-68, 2017.
- [3] Chai, J. Y., She, L., Fang, R., Ottarson, S., Little, C., Liu, C., & Hanson, K., "Collaborative effort towards common ground in situated human-robot dialogue." *9th ACM/IEEE International Conference on Human-Robot Interaction*, 2014.
- [4] 堀江拓実, 山口留実, 桑原多瑛, 渡辺巧登, 高汐一紀, "HRIにおける割込み時の会話順合意形成コーパスの構築にむけて" *電子情報通信学会 クラウドネットワークロボット研究会 (CNR)*, pp.41-6, 2020.
- [5] Bernieri, F. J., Gillis, J. S., Davis, J. M., Grahe, J. E., "Dyad rapport and the accuracy of its judgment across situations: A lens model analysis," *Journal of Personality and Social Psychology*, 71(1):110-29 1996.

心拍数からの勉強中の集中、眠気、疲労状態 検出手法の検証

川崎勇佑¹ 横窪安奈¹ ロペズギョーム¹

¹ 青山学院大学理工学部情報テクノロジー学科 〒252-0206 神奈川県相模原市中央区淵野辺 5-10-1

概要

精神状態を反映している自律神経系の活動を定量評価するため、心拍変動解析がよく用いられているが、心拍数のみでは不十分とされている。本研究では、スマートウォッチで計測可能な心拍数のみを用いて、勉強中の精神状態検出の可能性を検証した。集中、眠気、疲労のそれぞれのレベルを3段階に設定し、勉強開始と勉強中にいずれかのレベルが変わったときに状態ラベルを記録した。得られたデータセットを学習用とテスト用にランダムに分けたうえで、複数の機械学習手法を比較検証した結果、各精神状態の度合いを74%以上の平均精度で検出可能と分かった。

キーワード: 勉強 集中 眠気 疲労 心拍センサ

1. はじめに

作業効率および学習効果を高めるため、作業時および勉強時の集中はとても重要である。近年、ウェアラブルセンサの普及に伴い、生体情報を用いた集中状態の定量評価に関する研究が盛んである。大久保らは、椅子に加速度センサを付けて、集中状態を評価するシステムを提案しているが、精度は低く、個人差の影響も大きかった[1]。脳波計を用いる方法はより正確であるが、デバイスの大きさおよび電磁ノイズの影響に敏感なことから日常生活において利用することが困難である[2][3]。さらに、いずれも集中状態を推定する研究であり、仕事・勉強を始めとしたデスクワーク時にも起こる眠気や疲労状態の測定ができていない。デスクワーク時における作業効率には休憩の取り方が関係している。一度も休憩を挟まない長時間学習に比べ、一定時間ごとに短い休憩を挟む積み上げ学習の方では作業効率が向上することが示されている[4]。角田ら[5]および、東海林ら[6]のように、心拍変動の特徴量と精神疲労水準の間に相関があることを確認し、精神疲労における逐次推定は可能であることが多くの研究にて示唆されている。しかし、心拍変動を計測するには連続的に心拍間隔を高精度に計測する必要があり、日常生活で用いられるデバイスでは、未だに不可能である。

本研究では心拍センサ搭載のスマートウォッチ・スマートバンドを用いて、勉強中のメンタル状態を連続的に推定し、短期的なアクションおよび長期的なフィードバックにより、パフォーマンス向上を支援するシステムを提案している(図1)。

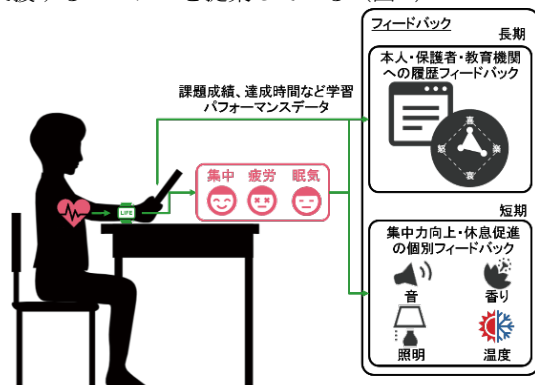


図1. 心拍数を用いた勉強中のメンタル状態推定とフィードバックシステムの概念図

精神状態を反映している自律神経系の活動を定量評価するため、心拍変動解析がよく用いられている。しかし、自律神経系の活動を評価するには心拍数のみでは不十分とされているため、既存のスマートウォッチでは評価できない。まず、スマートウォッチの心拍センサから心拍数関連の特徴量を用いて、集中、眠気、疲労の度合いが十分に検出可能かどうかを検証した。

2. 勉強中における心拍データの計測

2.1 実験の概要

実験参加者の心拍数を抽出するために左右いずれかの腕に、光電式容積脈波センサ搭載スマートウォッチ(FitBit versa2)を装着してもらった。集中、眠気、疲労のそれぞれの度合いを低・中・高の3段階に設定し、勉強開始時と勉強中にいずれかの精神状態ラベルの度合いが変わったときに記録した。勉強に極力影響しないように、スマートウォッチの画面をタッチするだけで変更された精神状態のラベルと度合と時刻が記録できるようなアプリケーションソフトウェアを開発した。

塾に通っている高校生男女6名に協力してもらい、自宅学習の勉強中の連続心拍データと精神状態ラベルの度合いデータを取得した。以下の手順に従い、被験者に通常通り勉強をしてもらった。

- (1) 勉強を開始する時に時間を記録する。
- (2) 勉強中に集中、眠気、疲労状態のいずれかが変化した場合にその時刻を記録してもらう。
- (3) 勉強が終了した時に時間を記録する。

2.1 取得したデータのまとめ

計測した心拍データはある程度の時間でまとめてその間の変化を特徴として、精神状態と関連付けることが必要になる。その最適な時間を特定するため、3分、5分と、10分の時間窓ごと心拍データと精神状態ラベルをまとめることにした。また、それぞれの解析窓を20秒、30秒、60秒シフトさせてデータセットを作成した。

各精神状態度合ごとに収集した心拍とラベルデータを表1にまとめている。精神状態ごとのレコード数はほぼ同じ数取得できているが、全く制御されていない計測環境のため、精神状態度合い毎のアンバランスが生じた。

表1 解析窓サイズによる各精神状態度合ごとに収集した心拍とラベルデータ数

精神状態	解析窓(分)	精神状態度合		
		低	中	高
集中	3	1239	10585	4789
	5	803	6814	3101
	10	365	3119	1431
疲労	3	11107	4354	1152
	5	7142	2825	751
	10	3246	1311	358
眠気	3	8977	6555	1081
	5	5775	4240	703
	10	2627	1959	329

3. 心拍数を用いた精神状態度合の推定

3.1 窓時間と学習モデルの評価

分析をするために心拍数データから次の七つの特徴量を抽出した：平均値，最大値，最小値，最大値 - 最小値，標準偏差，尖度，歪度。ランダムにデータセットの75%を学習用に利用した。Support Vector Machine (SVM)，Gradient Boosted Regression Trees (GBRT)，Random Forest (RF)，K 近傍法 (KNN) の四つのモデルを学習残りの25%のデータで精度評価した。そのプロセスを10回繰り返し，各精神状態度合のF1値(%)と標準偏差を算出した。どの学習モデルにおいても，一番高い精度は10分の窓サイズにて一番高い精度を得た。また，Random Forest モデルがどの精神状態度合においても，高い精度になった。

その組み合わせの詳細評価結果を表2にまとめている。集中，眠気，疲労のいずれの精神状態を74%以上の平均精度で推定できた。精神状態の度合いに関して，60%を下回る項目はあったものの，各状態度合のデータ数と照らし合わせると，300レコード程度の低集中，高疲労と，高眠気がそれぞれ一番低い精度となり，レコード数の多かった中集中，低疲労と，低眠気はそれぞれ85%を超える高い精度となった。そのことから，更なるデータを収集し，ラベルごとのアンバランスを縮小すれば，精度の向上が期待できると考えられる。

表2 10分窓と1分シフトのデータセット特徴量を用いて，RFでの学習精度のまとめ

	低	中	高	平均
集中	63.5±3	86.5±1	72.7±2	74.5±11
疲労	85.7±1	79.9±1	71.4±3	77.1±9
眠気	89.3±1	75.8±2	57.5±3	71.7±16

3.2 未知のユーザに対するロバスト性評価

次に，前述の10分窓の心拍数特徴量とRandom Forest 学習モデルを用いて，未知のユーザデータに対する精度評価を行った。前述6名のデータのうち，一名のデータを順番に除いてモデルの学習を行ったうえで，除いたユーザのデータでテストを行った。評価方法として合致率，適合率，再現率，F1値を導き出した。眠気状態のテストデータの高ラベルのデータが0個であり，必ず0%になるので表には斜線を引いた。その結果を表3に示している。明らかに1つ目の結果と比べて精度が落ちていることがわかる。ほとんど60%を超えていたラベルが，10%を下回る結果になり，80%を超えていたラベルも大幅に精度を落としている。

表3 未知のユーザに対する学習精度の評価結果

	合致率	適合率			再現率			F1 値		
		低	中	高	低	中	高	低	中	高
集中	39.2	3.5	39	43.3	0.8	64.1	26.3	1.3	48.5	32.7
疲労	32.4	79.6	16.6	35	29.2	82.6	3.7	42.7	27.7	6.6
眠気	46.2	50	30.1	/	87.2	3.2	/	63.6	5.8	/

4. まとめと展望

本研究では心拍数のみを用いて勉強中の集中，眠気，疲労の度合を3段階を予測するモデルを作成し，その精度を検証した。心拍数の特徴量抽出窓とシフト時間として，それぞれ10分と1分がどの学習モデルにおいても最も高い精度が得られることが分かった。個人に依らないデータセットで学習したモデルでは，10分の窓サイズにて一番高い精度を得た。また，Random Forest モデルがどの精神状態の度合推定においても，一番優れた結果を得て，平均74%以上の精度であったことから，心拍数のみでもある程度の精神状態の度合い予測ができることが示唆された。

一方，汎用モデルは未知のデータに対しての予測は全く機能しないことも示された。各学習ラベルのデータ数がまだ少なく，ラベル間のアンバランスが顕著であるが，このような結果になった一因だと考える。また，心拍数のベースラインと変動幅は個人により大きく変わるが，Nkurikiyeyezu らによると，少ない個人サンプルで汎用モデルの強化が見込める[7]。

今後の展望として，データをより多く集めることや，学習モデルの最適化，特徴量エンジニアリングの実施などを行うことなどをして改善を進めていく。

謝 辞

本研究の実験データは株式会社 JSOL と提携の学習塾にて収集され，株式会社 JSOL と著者らの共同研究の一環として解析を行っている。

参 考 文 献

- [1] 大久保雅史，藤村安耶：加速度センサを利用した集中度合い推定システムの提案，WISS 2008.
- [2] 趙丹俊，松田香菜，片山喜規，伊良皆啓治：脳波・NIRS・心電図による集中状態の評価，電子情報通信学会技術研究報告，111(423):117-20 (2012).
- [3] S. Tanabe, R. Nagase, K. Fujita, H. Touyama and Y. Tobe: HOKAGE: Providing the Attention Level of a Remote User Using EEG, IWSSS 2017.
- [4] Y. Watanabe, Y. Ikegaya: Effect of intermittent learning on task performance: a pilot study, Journal of Neuronet, 38:1-5 (2017).
- [5] 角田啓介，千葉昭宏，千明裕，吉田和広，渡部智樹，水野理：心拍変動を用いた認知能力水準の逐次推定，情報処理学会論文誌，58(2):568-79 (2017).
- [6] 東海林可奈，石井峻，横濱安奈，ロペズギョーム：生体信号を用いたデスクワークにおけるメンタル疲労度推定，DICOMO 2019.
- [7] Nkurikiyeyezu, Yokokubo, Lopez: The Influence of Person-specific Biometrics in Improving Generic Stress Predictive Models, Sens. Mater. 32 (2):703-722 (2020).

感情制御を伴うテキスト対話システムの実現

片山 晋¹ 米澤 拓郎¹ 大越 匡² 中澤 仁³ 河口信夫^{1,4}

1 名古屋大学大学院工学研究科 〒464-8601 愛知県名古屋市中種区不老町

2 慶應義塾大学大学院政策・メディア研究科 〒252-0882 神奈川県藤沢市遠藤 5322

3 慶應義塾大学環境情報学部 〒252-0882 神奈川県藤沢市遠藤 5322

4 名古屋大学未来社会創造機構 〒464-8601 愛知県名古屋市中種区不老町

概要

対話システムにおいて感情表現を行うことは、ユーザの生活満足度の向上などに繋がり、メンタルヘルスケアやカウンセリングにも応用可能であるとされている。しかし、現在の対話システムはユーザの感情や状況に適した感情表現を理解していない問題が存在する。そこで本研究では、ユーザの発話入力テキストから応答時の感情カテゴリを推定する感情制御を行い、感情的な対話応答の生成を行うテキスト対話システムを提案する。本提案手法は入力テキストから意味的文脈と感情的文脈を分散表現の抽出によって応答時の感情カテゴリの推定を行う感情制御部と、注意機構付き Seq2Seq に感情カテゴリラベルを付与することで感情的な対話応答生成を行う感情応答生成部の二つのニューラルネットワークによって構成される。推定された感情カテゴリに即して生成された感情的応答が適切なものであるかを評価するために、クラウドソーシングを使用して集めた 100 人の被験者による人手評価によって既存手法と提案手法の比較実験を行った結果、提案手法が生成精度や対話の自然さ、意味的整合性などの指標で上回ることを示した。

キーワード：感情制御 対話システム アフェクティブコンピューティング

1. はじめに

Apple 社の「Siri」や Amazon 社の「Alexa」などに代表されるスマートフォン上やスマートスピーカーなどのデバイス上で動作する対話システムが我々の日常生活に広く普及している。対話システムでより円滑なコミュニケーションを実現するために感情を扱う研究が行われており、感情を扱うことで、ユーザの生活満足度の向上や肯定的なインタラクション増加への貢献が認められている。また近年では、ニューラルネットワークを用いて感情的な応答を生成する手法[1]も提案されており、従来の生成手法で存在した汎用的な応答が多く生成されてしまう問題を解決し、より自然で感情的な応答生成手法を確立している。しかし、これらの感情的な対話応答生成の手法では、生成時に特定の感情カテゴリを付与することで、その感情に基づいた応答内容の生成に焦点が当てられている。一方で、人間同士の対話コミュニケーションでは、ユーザの感情状態や発話内容を考慮した上で適切な対応の決定をし、応答を行う感情制御を行っている。例えば、対話相手が悲しい感情状態である時に、ペットが亡くなって悲しんでいる相手に対しては共に悲しむ共感の応答を行うが、テレビゲームに負けて悲しんでいる相手に対しては楽しい感情で気持ちを切り替えさせるような応答を行うことが考えられる。そのため、対話システムがユーザの感情状態や発話内容を考慮する感情制御を行った上で対話応答を行うことは円滑でより人間らしい対話の実現に効果的であると考えられる。

そこで本研究では、ユーザの発話入力テキストから感情的文脈と意味的文脈を考慮した感情制御を行い、感情的な内容で応答生成を行う対話システムを提案する。分散表現を用いて感情的文脈と意味的文脈を考慮し、応答時の感情カテゴリの推定と感情的な対話応答生成を行う二つのニューラルネットワークを用いることで、入力テキストに対して感情的内容で対話応答を行う end-to-end なテキスト対話システムを実現する。

クラウドソーシングを用いた被験者 100 人による人手評価を行い、既存の感情的対話生成手法と比較して提案手法が対話の滑らかさや満足度、意味的整合性などの指標においてより優れていることを示した。

2. 関連研究

テキスト対話生成の手法は従来、ルールベース手法や統計的抽出ベースの手法が主流であった。近年では、深層学習の発展による大規模な対話コーパスを用いて柔軟な対話の生成を低コストで行う Seq2Seq などの生成ベースの手法が主流となっている。感情的な対話応答生成の研究[1]も行われているが、これらの研究の問題点として、感情的な応答の生成を行う際に、ユーザが最適な応答感情カテゴリを手動で選択しなければならない問題がある。

そのため、本研究では対話相手の感情状態や発話内容を考慮した上で応答時の感情カテゴリを推定する感情制御を行い、感情的な応答の生成を可能にする。

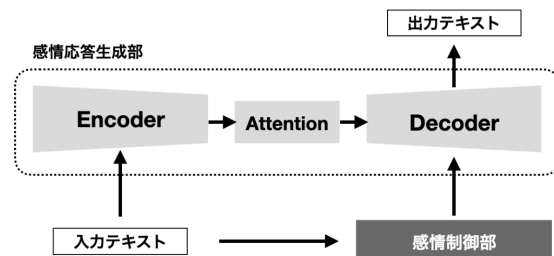


図1 システム構成図

3. 提案手法

本研究では、入力された発話テキストから応答時の感情カテゴリを推定し、その感情カテゴリに即した内容で対話応答を生成するテキスト対話システムを提

案する。提案手法は応答時の感情カテゴリを推定する感情制御部と感情的な対話応答を生成する感情応答生成部の二つのネットワークから構成される。提案手法のシステム構成図を図1に示す。

感情制御部では、入力テキストから fastText[2]による意味的分散表現と感情分散表現[3]を用いることによって意味的表現と感情的表現を抽出する。これらの分散表現を用いて、Transformer[4]のエンコーダに使用される Self-Attention 層と Feed Forward 層を用いてエンコードを行い、その出力を足し合わせて全結合層に送りソフトマックス関数によって one-hot ベクトルで表現された最適な感情ベクトルの出力を行う。

感情応答生成部では、入力されたテキストデータなどの時系列データを対応する時系列データに変換する Encoder-Decoder から成る Seq2Seq フレームワークと注意機によって構成される。Encoder では GRU ネットワークによって入力テキストの情報を隠れ層のベクトルへ圧縮し、Decoder では、感情制御部にて出力された感情ベクトルを注意機構と結合することによって感情的内容を含んだ応答の生成を行うことを可能にする。

これらの二つのネットワークを用いて、入力テキストの意味的文脈と感情的文脈を考慮した感情制御を行い、感情的な対話応答テキストの生成を可能にする。

4. データセット

本システムで使用する対話データセットの要件として、発話分と応答分がペアになっている大規模対話データ、ならびにそれぞれに感情カテゴリラベルが付与されたデータである必要がある。本研究では DailyDialog データセット[5]と、OpenSubtitles の字幕データセットを用いる。本研究で対象とする感情ラベルは Ekman の基本感情として定義されている Anger, Disgust, Fear, Joy, Sadness, Surprise の 6 クラスとし、DailyDialog データセットの中から Ekman の基本感情に該当する対話分を抽出する。OpenSubtitles は、約 20,8000 件の映画やテレビの脚本から抽出した多言語字幕データセットであり、本研究では英文字幕データセットを用いる。OpenSubtitles の字幕データは感情ラベルが付与されていないため、事前学習済みの BERT モデルによる感情分類器[6]を用いて感情カテゴリラベルの付与を行う。以上の過程によって 1,302,991 組の対話ペアデータを訓練データとして作成し、13,106 組の対話ペアデータをテストデータとして作成した。

5. 評価実験

提案手法を用いた出力テキストの有効性を明らかにするために、本研究ではベースラインと比較した自動評価と人手評価を行った。データセットにおけるテストデータから、6 つの感情における入力テキストをそれぞれ 10 件ずつ、合計 60 件サンプル抽出し、注意機構付き Seq2Seq モデル(Att-Seq2Seq)[7]と、感情的な応答を可能にする ECM[1]の出力テキストと比較する評価実験を行った。ECM は手動で感情カテゴリの選択を行うため、本実験ではランダムな感情カテゴリで出力を行うように設定した。クラウドソーシングサービスを用いて募集した 100 人の被験者によって 60 件の入力テキストに対して計 180 件の出力テキストを流暢さ、満足度、意味的整合性、感情の豊富さ、感情の

適切さの指標において 5 段階のリッカート尺度で評価をしてもらった。

図2に評価結果を示す。全ての指標において、提案手法がベースライン手法を上回っていることがわかる。特に、満足度においては Att-Seq2Seq よりも 0.817、ECM よりも 0.526 の値で上回っており、提案手法の出力テキストが比較的高い満足度を提供していることが分かる。

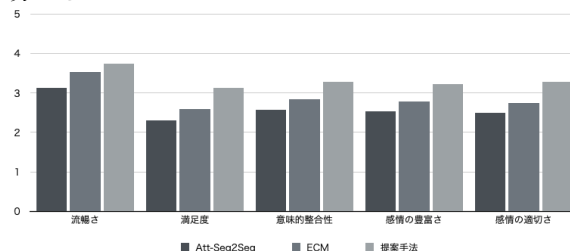


図2 評価結果

6. おわりに

本研究では、ユーザの発話内容のテキストから分散表現を用いて意味的要素と感情的要素を抽出し、感情適応を行うことで感情情報を含んだ対話応答の生成を行うテキスト対話システムを提案した。Opensubtitles から取得した対話データと DailyDialog データセットを用いて、感情カテゴリを推定する感情制御部と感情的な応答生成を行う感情応答生成部の二つのニューラルネットワークを構築した。クラウドソーシングサイトを用いて集めた 100 人の被験者による人手評価を行いベースラインの手法と比較した結果、自動評価と人手評価の両方においてベースラインの手法を上回る結果を示した。

今後の展望としては、マルチターンの対話における感情制御や感情制御のパーソナライズ化を行うことで、より人間らしく円滑な対話の実現を目指す。

文 献

- [1] Zhou ,Hao, et al. "Emotional chatting machine: Emotional conversation generation with internal and external memory." arXiv preprint arXiv:1704.01074 (2017).
- [2] Bojanowski, Piotr, et al. "Enriching word vectors with subword information." Transactions of the Association for Computational Linguistics 5 (2017): 135-146.
- [3] Agrawal, Ameeta, Aijun An, and Manos Papagelis. "Learning emotion-enriched word representations." Proceedings of the 27th International Conference on Computational Linguistics. 2018.
- [4] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems. 2017.
- [5] Li, Yanran, et al. "Dailydialog: A manually labelled multi-turn dialogue dataset." arXiv preprint arXiv:1710.03957 (2017).
- [6] Yang, Kisu, et al. "EmotionX-KU: BERT-Max based Contextual Emotion Classifier." arXiv preprint arXiv:1906.11565 (2019).
- [7] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014).

第37回 人間情報学会講演集

2020年 12月 3日発行

発行所 人間情報学会

〒112-0012 東京都文京区大塚 2-20-1
お茶の水女子大学 国際交流プラザ407

Tel. 03-6912-1361 Fax. 03-6912-1362

E-mail: ahi-conf@npowin.org

URL: <http://www.ahi-soc.info>