イスの揺れからの人物と活動の識別

望月康平 †1　細川健人 †2　豊浦正広 †3
山梨大学　山梨大学　山梨大学

1. はじめに

人間の行動をセンシングすることでデータ化するときに，得られたデータを人間が理解できる形に解釈することによって，人間の活動を支援する材料に昇華することができるケースがある．IoT センサの性能の進化とコモディティ化によって，多種多様なセンサを利用できるようになってきたものの，人間の行動への解釈の側面においては，まだ不十分な点も多い．

我々は，人間の活動の手がかりになる振動を観測できるイスに着目する．人間の活動の推定はカメラ映像の解析によって活発に行われているものの，カメラ映像にはプライバシに関わる余計な情報まで取得してしまう問題がある．この点において，人間が触れたときにだけ人間の活動を観測できるイスは有望である．EU で 2024 年の施行が計画されている AI 規制法案では，監視カメラの利用を制約する条項を含み，公権力によるむやみなカメラ映像の取得が制限される．日本でも，個人情報保護法の規制によって，監視カメラの映像を無加工で保存したり転送したりすることは，問題になることがある．これらは，そもそも人間が撮影されることに忌避感を覚えることが発端であり，可能であるならばこの状況は避けるべきである．この点において，振動センサの信号に対しては，現在のところ規制する法律はなく，また，人間が感じる忌避感も映像に比べると小さい．さらに，イスに取り付けた振動センサであれば，人間が触れていないときにはデータが発生せず，余計な情報を取得しなくて済む利点がある．

本研究では，イスに取り付けた 9 軸センサから得られる信号により，活動に加えて人物も識別することを目指す．人物の識別では，特定の個人であることを識別するものではなく，別のデータと同一の人物から発生したものであることを識別するものとし，データ間や活動間の紐づけを行うために利用できる．

本研究の貢献は以下の通りである．

1. Transformer を含む深層学習ネットワークによって，イスに取り付けた 9 軸センサの信号から，人物と活動の両方を推定する手法を提案する．
2. イスに座る人物を撮影し，同時にセンサの信号を収録することによって，センサ信号と人物および活動が紐づけられたデータセットを構築する．
3. 着席中の信号によって活動を認識するのとは別に，離着席時の信号によって人物の認識を高い精度で行えることを示す．

以下，2. では関連研究を紹介する．3. では着席中に観測される 9 軸センサの信号からからイスに座った人の活動認識する手法を提案する．4. ではさらに，離着時の大きな信号に着目することで，人物を高い精度で認識する手法を提案し，実験結果と考察を加え，5. で本研究をまとめる．

2. 関連研究

Mizumoto ら [1] は，8 つの加速度センサを内蔵したイスの設計を行うことで，業務を妨げることなく人物の姿勢を認識する手法を提案した．この手法により，座っている人の姿勢の検出と共に，在席時間や場所から着席時の姿勢の改善に活用できる．しかし，コストの面から必要なセンサの数はより少ないことが望ましい．比較的安価なセンサを使用した研究として，Yoo ら [2] らによる，近接センサを使用した研究がある．近接センサは感応式のセンサであり，スマートフォンやパソコンといったタッチパネルに利用される．Yoo らは，背もたれに取り付けられた近接センサから座姿勢を推定し，ヘッドレストに取り付けられたスピーカーから警告を出すイスを提案した．提案手法は，通常のイスと比較して，視覚ディスプレイによる作業中の人物の姿勢を改善する効果があることが示された．また，Sztyler ら [3] は，加速度計，ジャイロ，および地磁気のセンサデータを使用し，ウェアラブルデバイスから位置情報を含む人の行動を認識する新しいモデルを提案した．センサによる行動認識は高い精度で可能であることが示され，特に腰部に取り付けた場合に高い効果が得られるという結果が示された．

深層学習を人物の行動認識に利用した研究に，Chen ら [4] や Wang ら [5] の研究がある．Chen ら [4] は，単一の加速度センサから行動を表す特徴を抽出し，ニューラルネットワー

User and Action Identification from Chair Viblation

†1 Kouhei MOCHIZUKI, University of Yamanashi
†2 Taketo HOSOKAWA, University of Yamanashi
†3 Masahiro TOYOURA, University of Yamanashi

クを訓練することで行動の分類を行った．実験から，従来手法に比べて高い精度で人物の行動の認識が可能となったことが示された．Wang ら [5] は，ウェアラブルセンサから得られたデータを使用して，深層学習モデルである LSTM を訓練させることで行動識別を行った．HAR データセットを使用した実験では，高い精度での行動の識別が可能であることが示された．Chen らや Wang らの研究成果により，スマートフォンやウェアラブルデバイス等のモバイルデバイスへ拡張の幅が広がった．画像処理の分野で Hu ら [6] は，航空写真のマルチスペクトル画像から空間的な特徴を抽出するために，Convolutional Neural Network(ConvNet) と LSTM を結合したモデルである ConvLSTM を提案し，より高度な土地利用状況の認識を可能とした．

Shinmoto ら [7] は，対象者の腹部に大きな RFID タグを貼り付けることで，バッテリーを装着することなく，室内の対象者の行動を認識する装置を提案した．壁に配置した複数の RFID リーダによって，各時刻での RFID の位置と姿勢を推定することができるようになった．室内にいさえすれば位置や姿勢がおおまかにわかることは望ましい性質ではあるが，腹部に取り付けた RFID の物理的な大きさは無視することができず，利用できない場合も考えられる．イスに取り付ける装置では，対象者にバッテリーの管理をさせる必要はなくなり，明るい室内であればソーラーバッテリなどでも駆動することが可能である．

Ishac ら [8] は，イスに取り付けた 9 つの圧力センサによって，座っている人の前傾や後傾などの姿勢を推定し，健康に悪い姿勢であった場合に警告を出すシステムを提案した．人によっては背もたれから離れて座ることもあり，その場合には正しく姿勢が推定できなかった．本研究では IMU センサを用いるために，イスのどこかに体が接していなければいけないといったような制約を置く必要はなくなる．

Enokibori ら [9] は，胴体の姿勢を推定するために導電性繊維を織り込んだベストを作り，電気容量の変化によって姿勢を推定した．姿勢の変化によっては活動の推定をすることもできるものの，ベストを常に身に着けるのは活動の記録に対する費用対効果が小さい．

スマートフォンに内蔵されたセンサを用いて活動を認識する手法は多く提案されている [10]．スマートフォンを持つ位置に頑健に活動を推定する方法は提案されているものの，スマートフォンを常に身に着ける必要はあり，介護対象者や就学児などには適用が難しい．イスに取り付けたセンサであれば，イスに座ったときの信号を取り逃すことはない．

また，人物や行動を識別する手法として，映像を使用する手法も活発に提案されている [11][12]．Bremond ら [11] は，カメラ映像を使用した人物の行動認識として，映像の各フレームを静止画像として特徴抽出を行い，その後分類されたフレームをつなぎ合わせることで，一連の行動として認識を行った．人物の行動を自動的に認識する Bremond らの手法は，自動運転の安全性の向上などに利用されている．しかし，カメラ映像を使う場合には，プライバシ保護の問題が起こる．

以上から，本研究では，単一のセンサをイスに取り付けて使用することで，簡易で安価な，プライバシに配慮した人物と活動の識別を目指す．

3. 着席中の信号からの人物と活動の識別

3.1. 識別ネットワーク

Transformer は，Vaswani ら [13] によって発表された自然言語処理や時系列に有効とされるモデルである．時系列信号であるセンサデータを効率的に扱うため，本研究では Transformer を採用した．Transformer は意味抽出を行う Encoder 部と，抽出した意味から新たな出力へと変換する Decoder 部から構成される．特に Encoder 部の中で，複数の異なる Attention 機構を用いて入力データの前後関係の情報を取り出す Multi-Head Attention 機構が，意味抽出において大きな役割を果たしている．そこで本研究では，時系列の前後関係を抽出するために Multi-Head Attention 機構を含む Transformer の Encoder 部を使用した．Encoder 部を含むネットワーク全体は，図 1 のように構成される．入力は 3 軸加速度・3 軸角速度・3 軸地磁気の 9 次元の時系列である．図 1 の右図は Transformer の Encoder 部の詳細である．Multi-Head Attention と Feed-Forward 層に，それぞれ残差接続と正規化処理が行われる．Encoder 部は，多層パーセプトロン，Dense 層へと接続される．本モデルは，人物と活動を同時に学習するマルチタスク学習を行った．そのため，出力は，人物と活動それぞれの Softmax 関数から得られる識別クラスの尤度である．

3.2. 着席中の行動のセンシング

着席中のイスのセンシングによる人物と活動の識別を行った．まず，イスを利用したデータ収集のために，着席して複数の活動を行ってもらう実験を行った．実験の様子を図 2 に示す．イスの座面裏に 3 軸加速度・3 軸角速度・3 軸地磁気を計測する 9 軸 IMU センサを設置した．イスの座面裏に取り付けることで，活動の妨げにならず，センサの存在を意識せずに使用することができる．センサの方向関係は，X 軸が正面方向，Y 軸が左右方向，Z 軸が上下方向とした．本実験では，9 軸 IMU センサに ATR-Promotions 製の TSND121 を使用した．同製品は，サンプリング周波数が 50Hz であり，専用のソフトウェアである SensorController を介して信号を取得する．また，9 軸センサと合わせてウェ

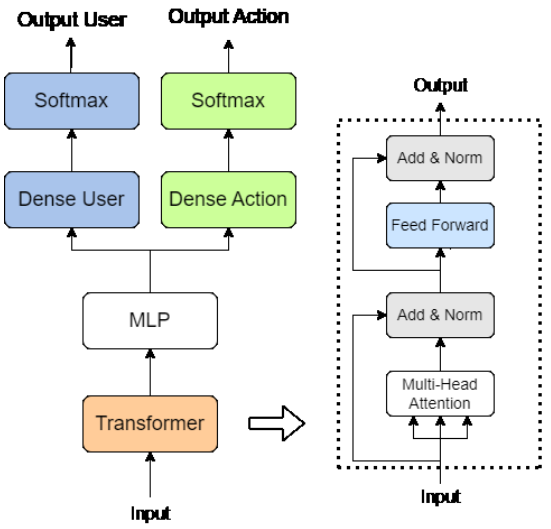

図 1　Transformer を含む識別ネットワーク

Figure 1　Proposed identification network including Transformer

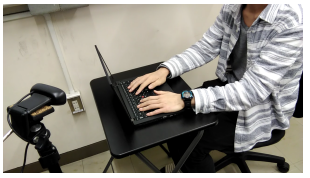

図 2　実験 1: 活動計測実験の様子

Figure 2　Experiment 1: Activity Measurement

ブカメラによる撮影を行った．録画映像は，DNN の訓練時の正解データの作成に使用した．

3.3.　データセット作成

男性 14 人，女性 8 人を対象とした実験を行った．実験協力者には，3.2 の実験環境下で，表 1 のスケジュールで決まった手順で活動を行ってもらった．活動は，座る・立つといった着席前後の 2 種類の活動，および，PC 作業・ボールペンによる書き作業・静止といった着席時の 3 種類の活動の全 5 種類である．計測は，各人ごと 4 セット行い，各セットを学習・検証・テスト用に分割して使用した．また，録画映像から人物と活動のタグ付けを行った．図 3 は，9 軸センサから得られた 1 セット分のイスの回転方向の角速度の例である．座る・立つといった離着席動作と比較して，PC 作業・静止・書き作業はパワーが小さいことが見て取れる．本研究では，着席中の活動に対して識別を行うため，PC 作業・書き作業・静止の 3 種類の活動の信号のみを使用した．映像によるタグ付け後，3.1 の Transformer を含むネットワークを訓練させ，人物と活動の識別を行った．

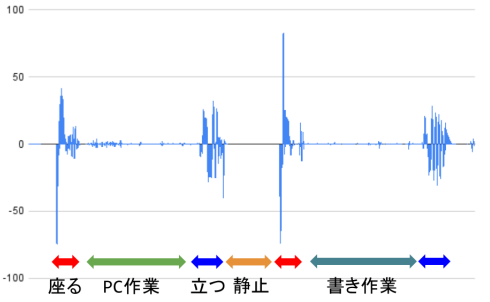

図 3　実験 1: 1 セット分の信号例

Figure 3　An example of Signal in Experiment 1

表 1　実験 1: 実験手順

Table 1　Overview of Experiment 1

経過時間	活動
0:00	計測開始
0:10	座る
0:20	PC 作業開始
1:20	PC 作業終了
1:30	立つ
1:40	座る
1:50	書き作業開始
2:50	書き作業終了
3:00	立つ
3:20	座る
3:30	計測終了

3.4.　実験結果

訓練したネットワークによって，22 人の人物を 0.1565，3 種類の活動を 0.8142 の正解率で識別することができた．時系列に有効とされる Transformer を深層学習ネットワークによる識別は，着席中の信号による数種類の活動に対して有効であると示された．一方で，識別するクラス数の多い人物の識別は困難であるという結果となった．これは，着席中のパワーの小さい信号では，人物のような多クラスの識別の場合，クラスごとの特徴がデータの差として現れづらいことが原因であると推察される．

4.　離着席動作による人物識別

着席中の信号を用いた人物識別の正解率の低いことを受けて，手法の改善を行った．3.3 で取得したデータから，着席中のパワーの小さな信号より，立つ・座るといった，離着席時の大きな信号の方が人物の特徴が表れているのではないかと考えた．そこで，着席中の活動の小さな信号ではなく，イスを離着席する動作の大きな信号に注目して識別を

(a) 実験協力者 A: 1 回目　(b) 実験協力者 A: 2 回目　(c) 実験協力者 B: 1 回目

図 4　実験協力者の離着席時の計測データの観察 (イスの回転方向の角速度)

Figure 4　Observed sensor data of participants from chairs (Angular velocity in the direction of chair rotation).

行った．

4.1.　データ分析

3.3 で使用したデータを分析することで，離着席動作が人物識別に有効であるか検証した．図 4(a)，(b) から，実験協力者 A では 2 回の計測で着席時には大きな信号が見られ，離席時は比較的小さな信号が共通して見て取れる．また，図 4(c) と比較して，実験協力者 B では離着席時に正負の片方に対して大きな信号が見られる．また，信号の大きさについても，実験協力者 A は離着席動作で差が見られるが，B は離着席共に平均 ±100[rad/s] 程度であるという違いも見られる．この違いは，離着席時にイスをどの程度回転するかという，使用者ごとの特徴に起因すると考えられる．着席中の信号と比較して，離着席動作には，このような人物ごとの特徴が現れやすいことから，離着席動作による人物識別は可能であると判断した．

4.2.　離着席動作のセンシング

離着席動作におけるイスのセンシングによる人物とイスの識別を行った．まず，データ収集のために，複数種類のイスで離着席動作を行ってもらう実験を行った．実験の様子を図 5 に示す．本実験では，9 軸 IMU センサ WitMotion WT901SDCL を使用した．WT901SDCL は，200Hz までの 3 軸加速度・3 軸角速度・3 軸地磁気を計測することができる．本実験では，信号の周波数はデバイスの上限である 200Hz とした．同製品は，SD カードを挿入することで計測を開始し，SD カードを取り出すことで計測を終了する．計測した信号は，SD カードにバイナリデータとして記録され，WitMotion の設定を行うソフトウェア (MiniIMU) 等により，各時刻ごとの 9 軸の信号を確認することができる．センサは，3.2 と同様に，イスの座面裏に取り付けた．

4.3.　前処理

図 6 は，1 回の離着席動作を行ったイスの上下方向の加速度の信号例である．最初の着席時は負の方向に，その後の

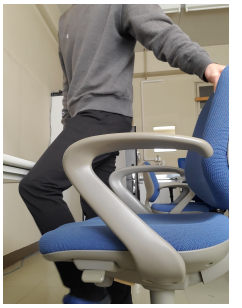

図 5　実験 2: 離着席動作

Figure 5　Experiment2: Seating and leaving actions

離席時は正の方向に大きな信号が得られていることがわかる．この様な信号に対して，前処理として，まず欠損値の除去を行った．また，センサから得られる信号には，少なからずノイズが含まれる．そこで，ノイズによる影響を低減するために，バターワースフィルタによる平滑化を行った．また，加速度・角速度・地磁気はそれぞれ取り得るデータの範囲が異なるので，データの標準化を行った．その後，離着席動作を行う部分に対して人物とイスのタグ付けを行った．タグ付けは，9 軸の信号の内，イスの上下方向である加速度 Z 軸と，回転方向である角速度 Z 軸に特徴が含まれていると考え，その 2 つの二乗和が閾値を超えたときにイスを使用しているとして判定した．

4.4.　計測実験

男性 22 人，女性 2 人を対象とした実験を行った．実験協力者には，静止したイスに着席し，数秒静止した後，イスから離席するという動作を 20 回行ってもらった．実験協力者には，着席時にイスの正面の机で作業を行う想定で座ることと，離席時にイスの向きを正面に戻すことの 2 点を指示した．実験協力者には，規定回数の離着席動作後に終了であることを伝えた．また，実験の疲れによるデータへの影響を無くすため，各イスが終了するごとに休憩時間を設けた．計測した 20 回分のデータは，4.3 で述べた前処理を行

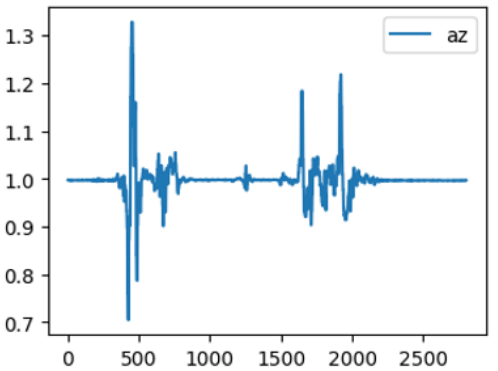

図 6　1 回分の離着席動作の信号
Figure 6　Signal of a pair of seating and leaving actions

図 7　実験 2 で使用した 3 種類のイス
Figure 7　Three chairs used in Experiment 2

い，図 6 の様に 1 回分ごとに分割して使用した．

図 7 は実験に使用した 3 種類のイスである．イス 1 と 2 は異なる商品だが，イス 1 と 3 は同じ商品である．1 と 3 で同じ商品を使った理由は，同じ商品でも，スプリングやキャスターの動かしやすさ等の違いから識別できると考えたからである．また，システムを導入することを想定した時，教室や会社では同じイスを大量に使用することが考えられるため，今回イス 1 と 3 は同じ商品を使用した．

4.5.　区間内投票による識別

本実験では，時系列を扱うことから，3.1 で述べた Transformer を含むネットワークを使用した．出力は，人物とイスのマルチタスク学習により，各クラスの Softmax 関数の出力が得られる．また，新たに，各クラスの出力から，指定した区間内における投票を行った．出力では，各時刻ごとのクラスの識別結果の尤度が得られるが，本研究では，時系列データでは前後関係が重要であることから，指定した区間内で各クラスの尤度の平均を取り，区間内で最も尤度が高いクラスを最終的な出力として採用した．例えば，20 人の人物識別の場合，9 次元のデータの入力に対し，出力はデータごとの 20 人の Softmax 関数による尤度が得られる．区間を 512 とすると，前後 512 個の 20 人分の尤度から，各人物ごとの尤度の平均を求め，区間内における尤度の平均が最も高いクラスを，その区間内全てのデータにおける識別結果とした．

4.6.　実験結果

取得した 24 人，3 種類のイス，20 回分のデータを，3.1 で述べた Transformer を含むネットワークで訓練させた．また，マルチタスク学習により，人物とイスの同時識別を行い，4.5 の処理により最終的な出力を得た．訓練したネットワークによって，24 人を 0.9768，3 種類のイスを 0.9702 の正解率で識別することができた．結果から，3.4 の考察の通り，着席中のパワーの小さな信号と比較して，離着席動作のパワーの大きな信号には，人物ごとの特徴が多く現れ，大人数の識別に対して有効であることが示された．また，イス 1・3 から，同じ製品のイスの識別も可能となった．

5.　結論

イスを利用する人物と活動を同時に識別することを目指した．まず，着席中の活動の信号により，人物と活動の識別を行った．識別には，時系列に有効なモデルである Transformer を採用した．実験では，22 人の 3 種類の活動に対し，22 人の人物を 0.1585，3 種類の活動を 0.8142 の正解率で識別できた．さらに，人物の正解率が低いことを受けて，手法を改善した．着席中の小さな信号からは人物の識別は難しいと考え，離着席動作の大きな信号から人物の識別を行った．実験では，24 人の 3 種類のイスを 20 回離着席したデータを，Transformer を含むネットワークで訓練させた．結果，24 人の人物を 0.9768，3 種類の活動を 0.9796 の正解率で識別できた．

以上から，離着席動作と着席中の信号で 2 段階の識別を行うことで，イスを利用する人物と活動を同時に識別することが可能となった．

今後の課題として，活動やイスの種類を増やした場合の正解率への影響を調べることを挙げる．イスの種類については，3 種類でしか検証を行っておらず，提案手法がどのような種類のイスに対しても有効であるかは確かめられていない．特に，ガスチェア以外の回転しないイスに対して，どれほど提案手法が有効に働くかについては，興味がある．また，活動の種類についても，3 種類のみを識別するにとどまっている．これまでのデータの観測からは，静止時と会話時にはほとんど信号に差がないことが確かめられており，活動によっては識別が難しいものがあると考えている．実際の利用で必要とされる活動の種類を踏まえながら，識別ネットワークの精度を高め，手法の限界を検証したい．

謝辞

本研究は JST さきがけ JPMJPR213 の支援を受けた.

参考文献

[1] Mizumoto, T., Otoda, Y., Nakajima, C., Kohana, M., Uenishi, M., Yasumoto, K. and Arakawa, Y.: Design and implementation of sensor-embedded chair for continuous sitting posture recognition, *IEICE Transactions on Information and Systems*, Vol. 103, No. 5, pp. 1067–1077 (2020).

[2] Yoo, W.-g., Yi, C.-h. and Kim, M.-h.: Effects of a proximity-sensing feedback chair on head, shoulder, and trunk postures when working at a visual display terminal, *Journal of Occupational Rehabilitation*, Vol. 16, pp. 631–637 (2006).

[3] Sztyler, T., Stuckenschmidt, H. and Petrich, W.: Position-aware activity recognition with wearable devices, *Pervasive and mobile computing*, Vol. 38, pp. 281–295 (2017).

[4] Chen, Y. and Xue, Y.: A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer, *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1488–1492 (2015).

[5] Wang, L. and Liu, R.: Human activity recognition based on wearable sensor using hierarchical deep LSTM networks, *Circuits, Systems, and Signal Processing*, Vol. 39, pp. 837–856 (2020).

[6] Hu, W.-S., Li, H.-C., Pan, L., Li, W., Tao, R. and Du, Q.: Spatial–spectral feature extraction via deep ConvLSTM neural networks for hyperspectral image classification, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58, No. 6, pp. 4237–4250 (2020).

[7] Shinmoto Torres, R. L., Visvanathan, R., Abbott, D., Hill, K. D. and Ranasinghe, D. C.: A battery-less and wireless wearable sensor system for identifying bed and chair exits in a pilot trial in hospitalized older people, *PLOS ONE*, Vol. 12, No. 10, p. e0185670 (online), 10.1371/journal.pone.0185670 (2017).

[8] Ishac, K. and Suzuki, K.: Lifechair: A conductive fabric sensor-based smart cushion for actively shaping sitting posture, *Sensors*, Vol. 18, No. 7, p. 2261 (2018).

[9] Enokibori, Y., Ito, Y., Ikeda, K., Suzuki, A., Shimakami, Y., Kawabe, T. and Mase, K.: Respiratory Volume Estimation by a Stretchable Textile Sensor, *Advances in Science and Technology*, Vol. 80, Trans Tech Publ, pp. 136–141 (2013).

[10] Wang, A., Chen, G., Yang, J., Zhao, S. and Chang, C.-Y.: A comparative study on human activity recognition using inertial sensors in a smartphone, *IEEE Sensors Journal*, Vol. 16, No. 11, pp. 4566–4578 (2016).

[11] Brémond, F., Thonnat, M. and Zúniga, M.: Video-understanding framework for automatic behavior recognition, *Behavior Research Methods*, Vol. 38, No. 3, pp. 416–426 (2006).

[12] Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L. and Hoi, S. C.: Deep learning for person re-identification: A survey and outlook, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 6, pp. 2872–2893 (2021).

[13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I.: Attention is all you need, *Advances in neural information processing systems*, Vol. 30 (2017).