

生物模倣的自己注意機構を導入したスパイキングニューラル

ネットワークによる神経動態予測モデルの提案

6年B組 岡本 晃朋

指導教員 鵜飼 哲真

1. 要約

ニューロンモデルに生物学的な自己注意機構を導入することで、スパイキングニューラルネットワーク(SNN)の神経動態予測精度を向上させると共に、より生物学的な挙動を再現することを目的としている。SNNの時間的情報処理能力にTransformerの強力な注意機構とイオンチャネル等の神経論的計算科学の理論を統合した新しいモデルを提案することで、現在76%の精度でのスパイク予測を達成している。

キーワード

Computational Neuroscience, Bioinformatics, Machine Learning,

2. 研究の背景と目的

私は昨年脳波や五感情報を用いたマルチモーダルな感情推論モデルの作成を行っていた。その際に脳波の θ 帯について感情によって時系列による変化が大きく異なり、時間と共に減衰していくことを発見した。この原因としてシナプスの電位が時間と共に減衰していくことや、シナプス可塑性による影響だと仮説を立てた。そこで、スパイキングニューラルネットワークに自己注意メカニズムを導入することにより時空間パターンの学習がより容易になり、精度向上や生物学的なモデルの再現度の向上につながるかと考え、本研究を始めた。

3. 脳波及び生理的信号の解析

最初に仮説にて示した脳波の時間減衰について紹介する為、脳波が具体的にどのような形式で感情などに影響を与えるか解析を行った結果を示す。脳波の解析データに

はOpenBCIのヘッドセットを用いた測定データ、及びECSMP[1]データセットにおける映画鑑賞時の脳波を測定したものである。

3.1 感情による脳波の差異

10/20法にてFp1, Fpz, Fp2, AF3, AF4, F7, F8, A2の8箇所電極から250Hzでサンプリングを行っている。

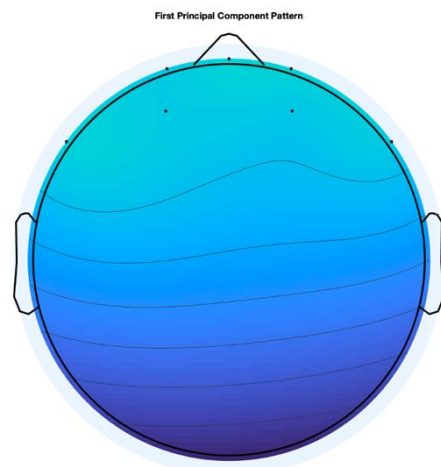


Figure 1: Topography of dataset

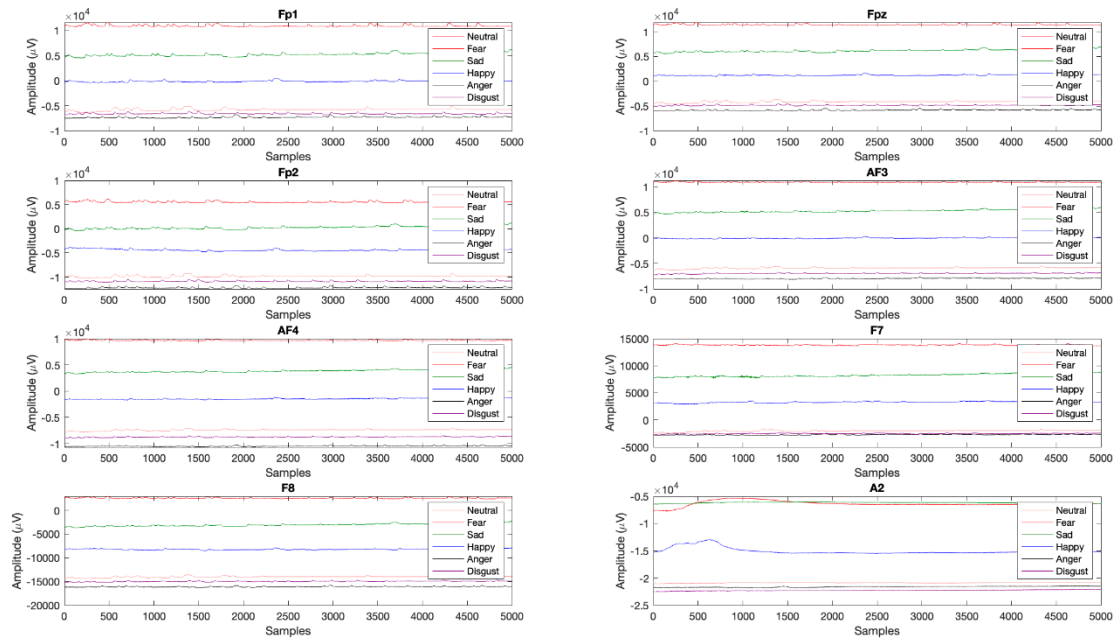


Figure 2: Electrode classified by emotion

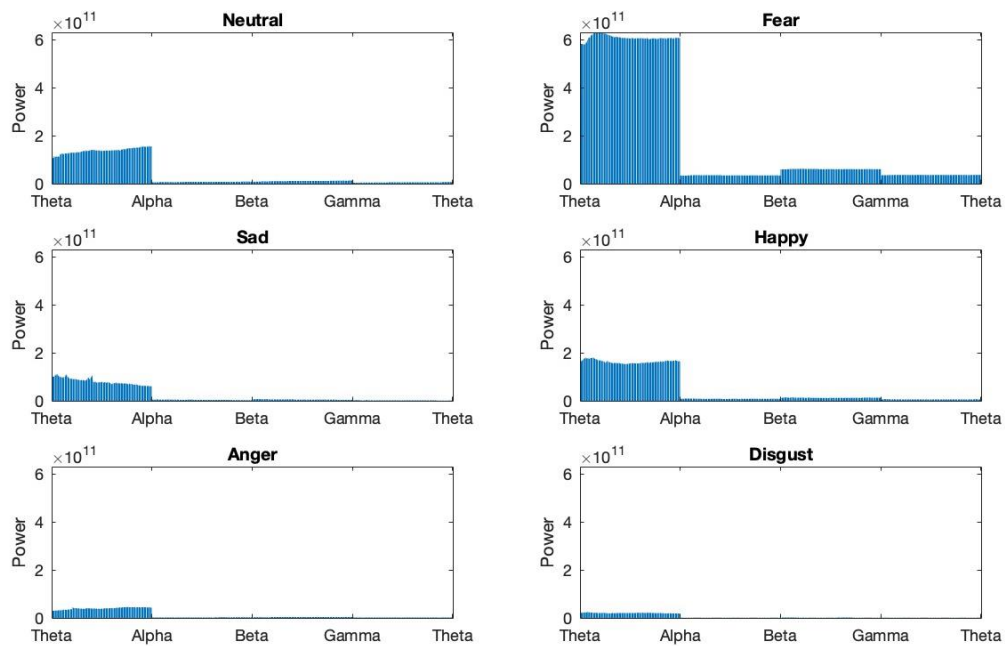


Figure 3: Band Power for Different Emotion

Figure 2, 3 より、感情によって Amplitude (Power)が異なることが確認できる。そこで、それぞれの感情についてスペクトラムを作成し、どのように変異していくのかに着目した。帯スペクトル Figure 4 から感情に

よってバンド周波数の持続性が異なることが読み取れる。これは、感情の種類によって時間変化の変位が異なることを示している。このことから、生命活動において時間減衰に関する関数が必要であると考えた。

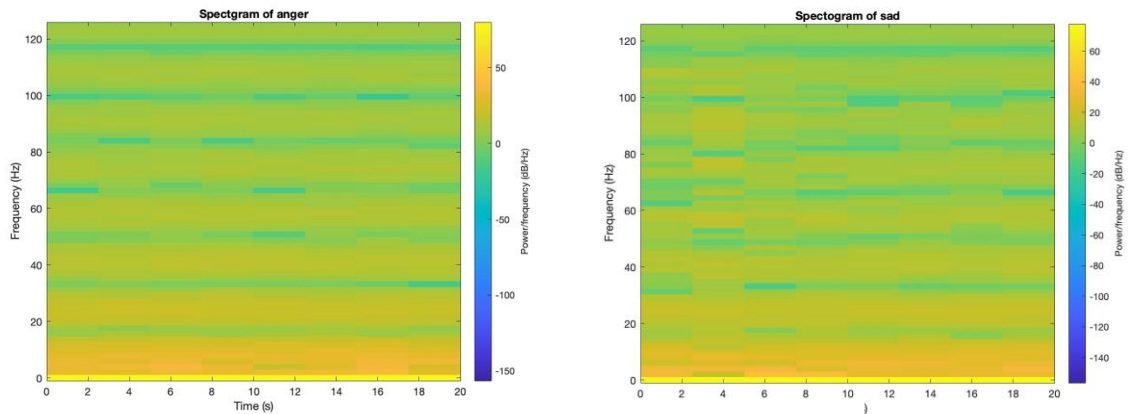


Figure 4: Spectrum of anger and sad

3.2 生理的信号の一致性

Figure 5~8 より生理的信号については、感情が変わる時などに反応がみられる場合が多いと結論づけた。インタラクティブでない場面(映画を見ている状態)で行う体の

移動や体温、皮膚電位などから、感情が変化する前後のタイミングを推論できると考えている。これらのデータについては発生時間が殆ど同じであり、感情のトリガーとしては一つに絞ることを考えている。

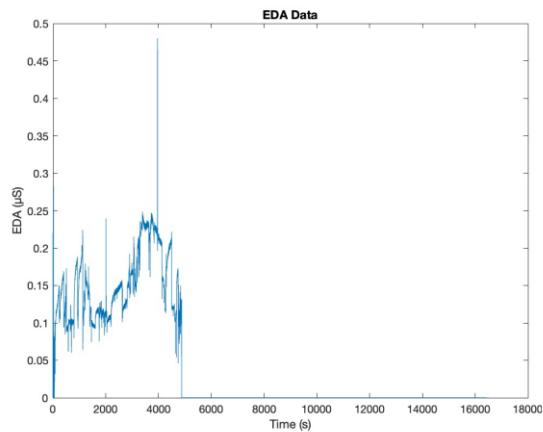


Figure 5: EDA

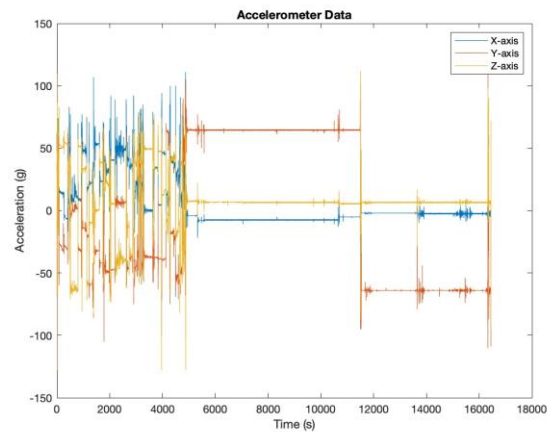


Figure 6: Accelerometer

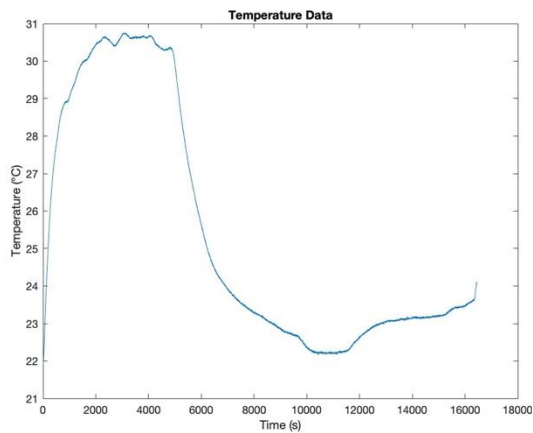


Figure 7: Temperature

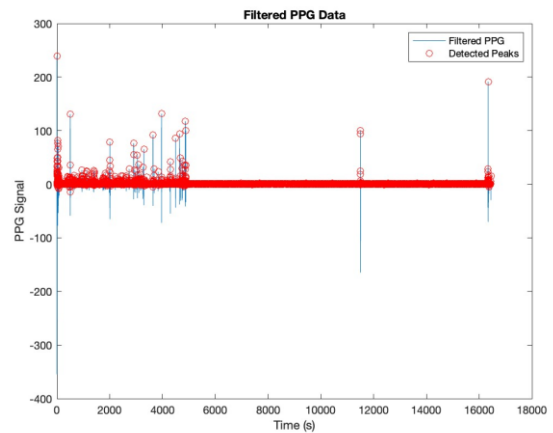


Figure 8: PPG(Filtered)

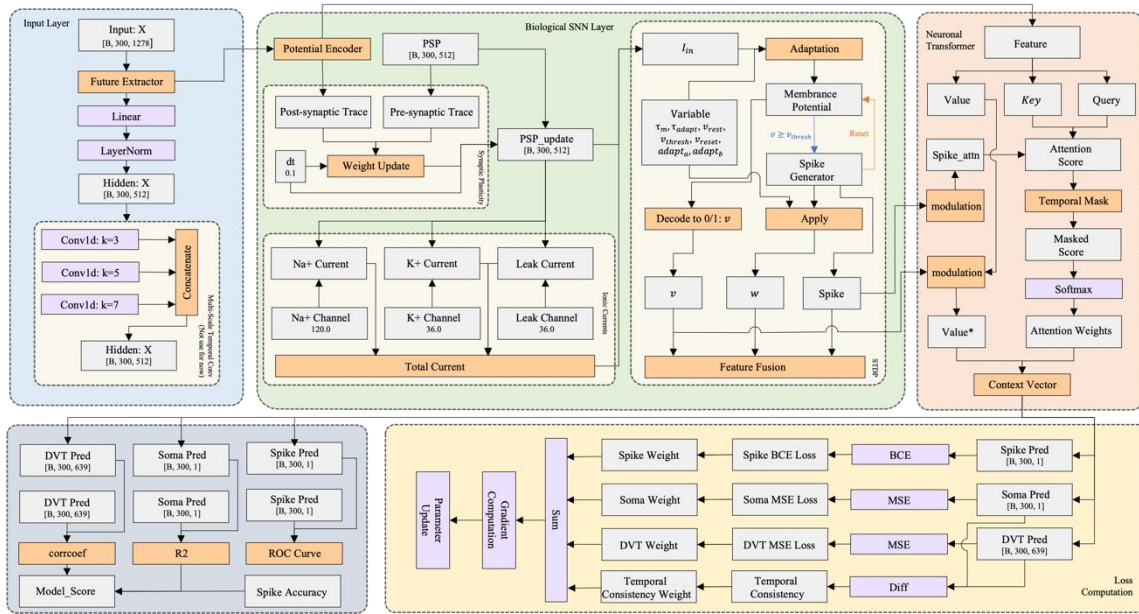


Figure 9: Architecture of CSnT

4 アーキテクチャの提案

脳波から見られる傾向を参考に、新たなアーキテクチャ CSnT: Continuous Spiking Neural Transformer (Figure 9)の提案を行った。以下にそれぞれのモジュールについて解説していく。

4.1 生物学的自己注意機構の提案

通常の Transformer[2]で用いられる自己注意機構では時間的相関を見出すことができない。時系列を扱える iTransformer などのアーキテクチャが提案されているが、生物学的な演算を行うにあたり妥当性に欠けていると考えた。そこで、新たに生物学的自己注意機構を提案した。

スパイクニューラルネットワーク (SNN)におけるスパイク S とその転置 (transpose) S^T 間での類似度を

$$A_{spike} = \frac{SS^T}{\sqrt{d}} \quad (1)$$

と表現する。ここで d は埋め込み次元(dim)であり、スケールリングすることで勾配の爆発や消失を防いでいる。

このモジュールに時間減衰を導入するために、実験的に SNN で用いられる指数減衰型の α 関数[3]を用いた。

$$K(\Delta t) = \exp\left(\frac{-|\Delta t|}{\tau}\right) \quad (2)$$

これらのアダマール積を主軸に、生物学的自己注意メカニズムを以下のように定義した。

$$A(Q, K) =$$

$$\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}} \odot K(\Delta t) \odot (1 + A_{spike})\right) \quad (3)$$

$$\text{Context} = A(Q, K)V \odot \varsigma(g(V_m)) \quad (4)$$

これらにより、Transformer にスパイクの変調 $\frac{SS^T}{\sqrt{d}}$ 、時間依存 $K(\Delta t)$ 、膜電位の相関 A_{spike} などを組み込み、シナプスの電位が時間と共に減衰していくことを再現した。Context では 4.3 にて説明する膜電位 V_m にシグモイド関数 ς を施し神経細胞の特性を抽出したものと積を用いることでモデルの挙動に生物性を持たせている。

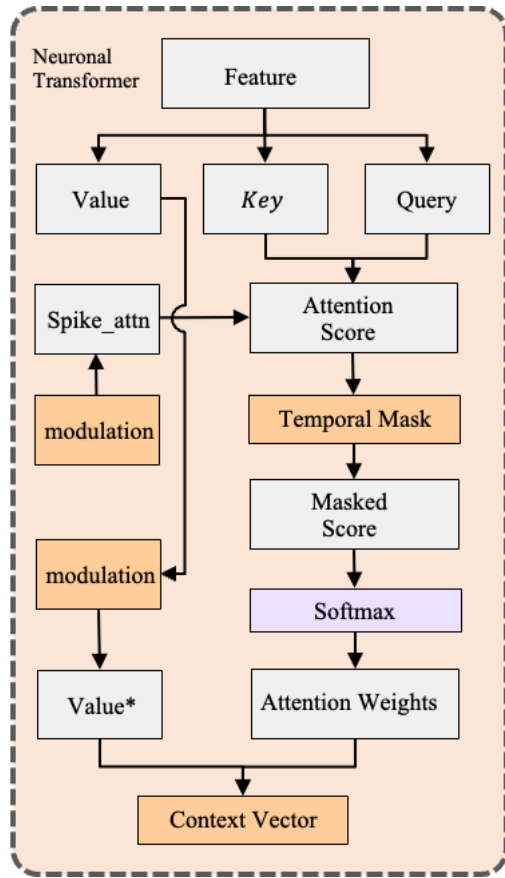


Figure 9: Architecture of Neuronal Transformer

4.2 SNN レイヤーの定義

次に、SNN における生物妥当性を保持するために、Synaptic Plasticity(STDP)[4]と Hodgkin-Huxley 型ニューロンモデル(HHModel)[5]を用いて電流 I と電圧 V を計算する。

4.2.1 STDP の導入

実際のシナプスの学習メカニズムを模倣することを目的に、ニューロン間の発火時間差に起因するシナプス可塑性を再現した STDP を導入した。プレシナプスとポストシナプスの動力学について、プレシナプス性トレース x_{pre} とポストシナプス性スパイク S_{pre} 間で LTP(Long-Term Potentiation)を、ポストシナプス性トレース x_{post} とプレシナプス性スパイク S_{post} 間で LTD(Long-

Term Depression)を以下の微分系で定める。

$$\frac{dw}{dt} = \eta(x_{post}(t)S_{pre}(t) - x_{pre}(t)S_{post}(t)) \quad (5)$$

また、トレースの動力学における時間発展については以下の微分方程式で表される。

$$\frac{dx_{pre}}{dt} = -\frac{x_{pre}}{\tau_{pre}} + \sum_f \delta(t - t_f^{pre}) \quad (6)$$

$$\frac{dx_{post}}{dt} = -\frac{x_{post}}{\tau_{pre}} + \sum_f \delta(t - t_f^{post}) \quad (7)$$

ここで、 τ_{pre} は時定数、 δ はディラックのデルタ関数である。 $\sum_f \delta(t - t_f^{pre})$ はスパイク発生時にトレース変数を増加させること、 $\frac{x}{\tau}$ はトレース変数の指数的な減衰を表すことで、直近のスパイク活動が保持されるようにした。そして、計算したプレシナプススパイクと重みの線形結合によって膜電位を計算した。

4.2.1 HHModel への接続

STDP にて計算した膜電位を元に HHModel による電流の計算を行った。

膜容量 C_m をコンデンサ、イオンチャンネルを可変抵抗と見做し、キルヒホッフの第二法則から

$$C_m \frac{dV}{dt} = -\sum I_{ion} + I_{ext} \\ = -(I_{Na} + I_K + I_L) + I_{ext} \quad (8)$$

と定められる。Hodgkin と Huxley がイカの神経軸索を膜電位固定法から実験的に求めた式 9~11 を用いて、STDP にて計算した膜電位を用いて電流を計算した。

$$I_{Na} = g_{Na} m^3 h (V - E_{Na}) \quad (9)$$

$$I_K = g_K n^4 (V - E_K) \quad (10)$$

$$I_L = g_L (V - E_L) \quad (11)$$

ここで、 g_{Na} , g_K は Na^+ , K^+ の最大コンダクタンス、 g_L はオームの法則から導かれる。また、 m , h , n はゲートの開閉確率を表している。

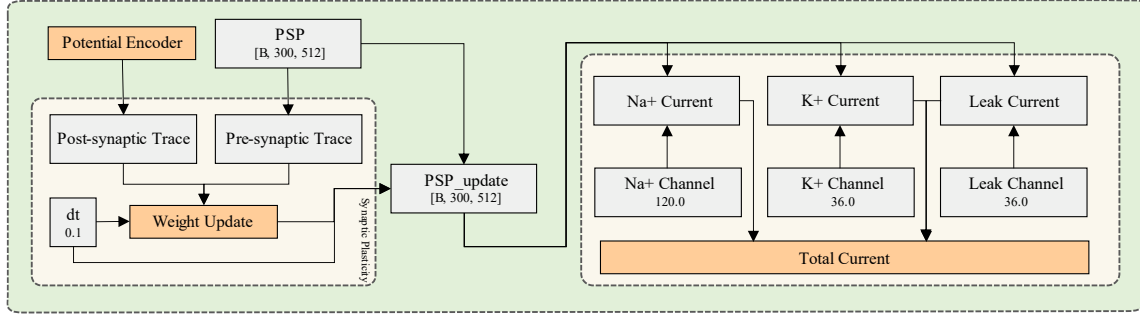


Figure 10: Structure of SNN Layer

$$\frac{dx}{dt} = \alpha_x(V)(1 - x) - \beta_x(V)x$$

$$(x = m, h, n) \quad (12)$$

また、 α_x, β_x は以下のように表される。

$$\alpha_m(V) = \frac{0.1(25 - V)}{\exp[(25 - V)/10]} - 1 \quad (13)$$

$$\beta_m(V) = 4\exp(-V/18) \quad (14)$$

$$\alpha_h(V) = 0.07\exp(-V/20) \quad (15)$$

$$\beta_h(V) = 4\exp(-V/18) \quad (16)$$

$$\alpha_n(V) = \frac{0.01(10 - V)}{\exp[(10 - V)/10]} - 1 \quad (17)$$

$$\beta_n(V) = 0.125 \exp(-V/80) \quad (18)$$

4.3 LIF Neuronの実装

STDPにて得られた電圧 V と Hodgkin-Huxley型ニューロンモデルにて得た電流 I を用いて Leaky integrate-and-fire(LIF) Neuron[5]に適応機構を追加した Adaptive exponential integrate-and-fire model(AdEx)[6]を実装した。膜電位の動力学を適応変数 w を用いて

$$\tau_m \frac{dV}{dt} = -(V - V_{rest}) - w + RI \quad (19)$$

と表す。ここで、 τ_m は膜時定数、 V は膜電位、 V_{rest} は静止膜電位、 RI は入力電流を表す。また、適応変数 w の動力学は

$$\tau_w \frac{dw}{dt} = -a(V - V_{rest}) - w + b \sum_i \delta(t - t_i) \quad (20)$$

となる。 τ_w は適応時定数、 a は電圧依存性適応の強さ、 b はスパイク依存性適応の強さ、 $\delta(t - t_i)$ はスパイクの発生を表している。メカニズムとして、1. 膜電位が上昇する程適応変数 w も増加する、2. 適応変数 w の増加により膜電位を下げる方向に働く、3. スパイクの発生により w が急激に増加、4. 連続発生を抑制し、スパイク発火頻度の適応を実現するようになっている。

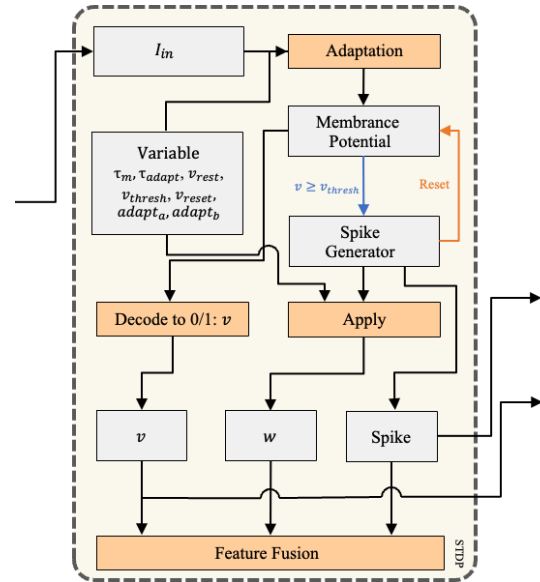


Figure 11: Structure of AdEX

そして、最終的なスパイクを

$$\text{if } V \geq V_{thresh}: \begin{cases} V \rightarrow V_{reset} \\ \text{emit spike} \end{cases} \quad (21)$$

として、発火の是非を判断した。

4.4 特徴抽出モジュールの実装

単純な特徴抽出モジュールを設計することで、神経細胞からの入力情報を失うことなくモデルに渡すことを目指した。データには David ら [7] が Neuron にてシミュレーションを行ったデータを用いた。

$$F = \phi(W_{in}X + b) \quad (22)$$

入力 $x \in \mathbb{R}^{\{\text{input_size}\}}$ に対して、線形変換を行う。

$$h^1 = W^1x + b^1 \quad (23)$$

ここで、 $W_1 \in \mathbb{R}^{\{\text{hidden_size} \times 2\} \times \text{input_size}}$, $b_1 \in \mathbb{R}^{\{\text{hidden_size} \times 2\}}$ の次元を持つ。次に、レイヤー正規化を行う。

$$\begin{aligned} h^2 &= \text{LayerNorm}(h^1) \\ &= \gamma \odot (h^1 - \mu) / \sqrt{(\sigma^2 + \varepsilon)} + \beta \end{aligned} \quad (24)$$

ここで、 μ は平均、 σ^2 は分散、 γ, β は学習パラメータである。次に、活性化関数 GELU を用いる。これは標準正規分布の累積分布関数を用いており、モデルが巨大であるほど効率的に正規化が可能である。

$$\begin{aligned} h^3 &= \text{GELU}(h^2) = h^2 \Phi(h^2) \\ &= h^2 \cdot \frac{1}{2} \left[1 + \frac{2}{\sqrt{\pi}} \int_0^{\frac{x}{\sqrt{2}}} e^{-t^2} dt \right] \end{aligned} \quad (25)$$

その後 Dropout, 線形変換, Layer Normalization を再度行ったものをモデルの入力とした。

$$\begin{aligned} h_4 &= \text{Dropout}(h_3, p = 0.1) \\ &= h^3 \odot \frac{m}{(1-p)}, m \sim \text{Bernoulli}(1-p) \end{aligned} \quad (26)$$

$$h^5 = W^2h^4 + b^2 \quad (27)$$

$$\begin{aligned} y &= \text{LayerNorm}(h_5) \\ &= \gamma' \odot (h^5 - \mu') / \sqrt{(\sigma'^2 + \varepsilon)} + \beta' \end{aligned} \quad (28)$$

4.5 損失関数の定義

総損失関数を以下のように定義した。

$$\begin{aligned} L &= \lambda^1 L_{\text{spike}} + \lambda^2 L_{\text{soma}} + \lambda^3 L_{\text{dvt}} \\ &\quad + \lambda^4 L_{\text{temporal}} \end{aligned} \quad (29)$$

ここで、 λ は各損失関数の重みを表し、Epoch が進むごとに配分を変えることで各予測に対する適応性の向上を狙う。以降、正解を y , 予測を \hat{y} で表す。

L_{spike} にはバイナリクロスエントロピーロス (BCE Loss) を採用した。2 値分類が可能なため、Spike に適すると考える。

$$\begin{aligned} L_{\text{spike}}(y_{\text{spike}}, \hat{y}_{\text{spike}}) \\ = - \sum (y \log(\sigma(\hat{y})) + (1-y) \log(1 - \sigma(\hat{y}))) \end{aligned} \quad (30)$$

ここで、 σ はシグモイド関数である。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (31)$$

また、 $L_{\text{soma}}, L_{\text{dvt}}$ には平均二乗誤差 (MSE) (24) を採用した。

$$L_{\text{soma}} = \frac{1}{n} \sum (y_{\text{soma}} - \hat{y}_{\text{soma}})^2 \quad (31)$$

$$L_{\text{dvt}} = \frac{1}{n} \sum (y_{\text{dvt}} - \hat{y}_{\text{dvt}})^2 \quad (32)$$

L_{temporal} には Temporal consistency Loss を用いた。

$$\begin{aligned} L_{\text{temporal}} &= \frac{1}{T} \sum |\hat{y}_{\text{soma}}^t - \hat{y}_{\text{soma}}^{t-1}| \\ &\quad + \frac{1}{T} \sum |\hat{y}_{\text{dvt}}^t - \hat{y}_{\text{dvt}}^{t-1}| \end{aligned} \quad (33)$$

このように、時間変化による差分を計算しており、膜電位 y_{soma} や Dendritic Voltage Trace (樹状突起電圧) y_{dvt} は物理的制約により急激な変化が起こりにくい連続信号であるため、この生物学的な特徴を再現する為に Temporal consistency Loss を用いた。

4.5 評価メトリクスの定義

損失関数と同様に、それぞれの予測対象別に評価メトリクスを別途定義した。

スパイク予測の性能評価は AUC-ROC を用いた。TPR (True Positive Rate) と

FPR(False Positive Rate)の関係を描くROC 曲線(Receiver Operating Characteristic Curve)を用い、予測スコアから正答率を算出する。また、Soma 電位には説明分散 R を用いた。

$$R_{Soma}^2 = 1 - \frac{MSE(y_{soma}, \hat{y}_{soma})}{U_{y_{soma}}^2} \quad (34)$$

ここで U は不偏分散である。

DVT には相関係数を用いた。

$$\rho_{dvt} = \text{Corr}(y_{dvt}, \hat{y}_{dvt}) \quad (35)$$

5. 研究結果及び考察

研究結果を Figure.12~16 に示す。

Figure12 は、実際にモデルをシミュレーションして得られた神経活動である。上側が 600ms, 下側に拡大した 60ms の活動を示す。拡大されたグラフから分かるように、興奮性シナプス後電位に見られる特徴である数ミリ程度の電位変動が見られること、不規則な振動パターンを再現できていることから神経活動の特性をつかめていると考えられる。

Figure13~16 では Validation における Loss 及び Accuracy の推移を示す。

Figure14, 16 より Spike の Loss は 0.0062, DVT の Loss は 0.0008 とかなり収束が良いと考えられる。一方で, Spike の Accuracy は 76.69%と高くないことから, 評価メトリクス及び損失関数の最高が必要である。

6. 今後の展望

6.1 時間減衰の再考

現在, 生物学的自己注意メカニズムにおいて時間減衰を表すのに α 関数(式 2)を使用しているが, これは二重指数関数型モデル [8](式 36)において立ち上がり時定数 τ_r (シ

ナプス後電流の立ち上がりにかかる時間)と減衰時定数 τ_s (神経伝達物質の減少量)が等しい時に成り立つものであり, 推論上神経伝達物質のシミュレーションデータがない為用いた。今後神経動態シミュレーションにてこれらの要素も考慮したモデルを作成し, モデルに帰属させたい。

$$A = \frac{\tau_d}{\tau_d - \tau_r} \cdot \left(\frac{\tau_r}{\tau_d}\right)^{\frac{\tau_r}{\tau_r - \tau_d}} \quad (36)$$

6.2 偏微分系の導入

モデルの多くの部分で微分方程式を用いており計算オーダーが大きくなっている。

$$\tau_{ion} \sim \mathcal{O}(1)$$

$$\tau_{stdp} \sim \mathcal{O}(B * T * N * M)$$

$$\tau_{neuron} \sim \mathcal{O}(1)$$

$$\tau_{SNN} \sim \mathcal{O}(B * T * N * M + H)$$

$$\tau_{attn} \sim \mathcal{O}(B * H * T * T)$$

$$\tau_{trans} \sim \mathcal{O}(B * H * T * T)$$

これより, モデル全体での \mathcal{O} は

$$\tau_{csnt} \sim \mathcal{O}(B * T * (N * M + H * T))$$

と, 微分方程式を用いている部分と Transformer 関連で特に計算オーダーが大きくなっていることがわかる。これらの式をまとめる上で偏微分系を用いることで計算コストの削減が可能であるので, 実用に向けて最適化を行なっていきたい。

6.3 Spike 発火の改良及び生理活性の利用

Spike 発火について実際の生理活性及び化学反応にて取りうるエンタルピー等を用いることで, 人間の脳で表現できるレベルを拡張することを目指す。これにより細胞活動がより豊かな表現が可能になると考えられ, 人間の活動との紐付けが可能になると考えられる。

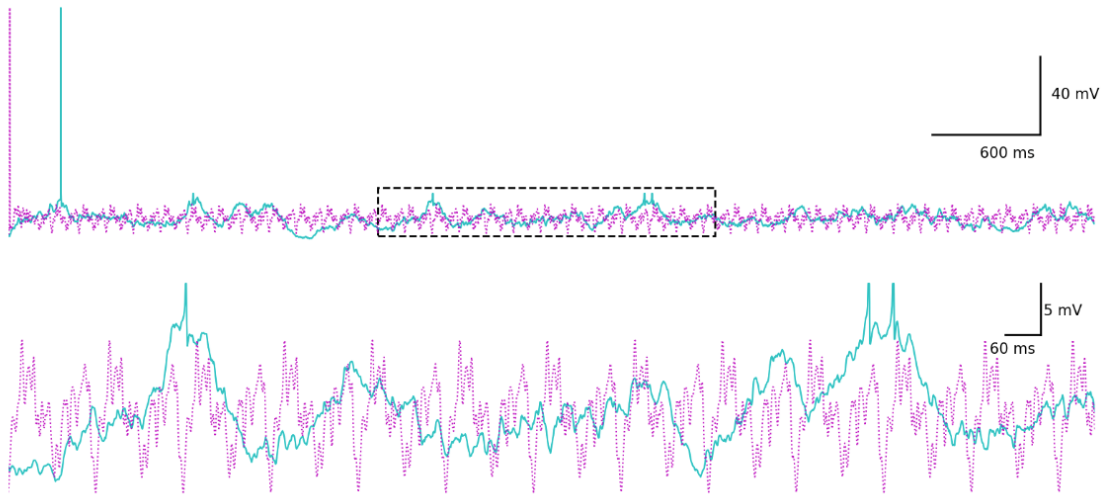


Figure 12: The Simulation of Neural Activity

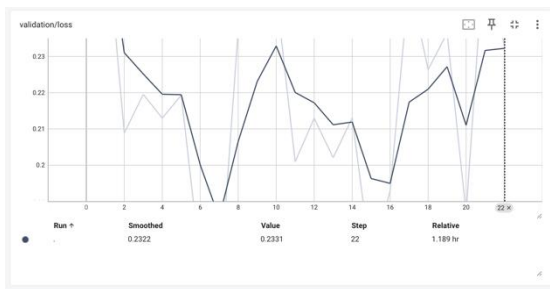


Figure 13: Validation/All Loss

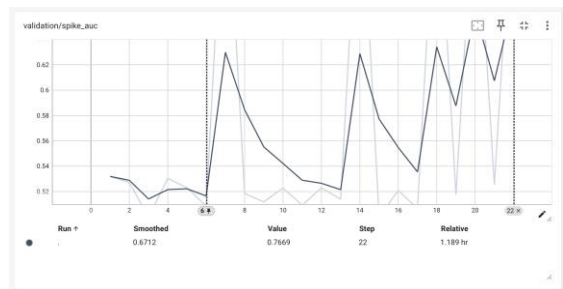


Figure 15: Validation/Spike Accuracy

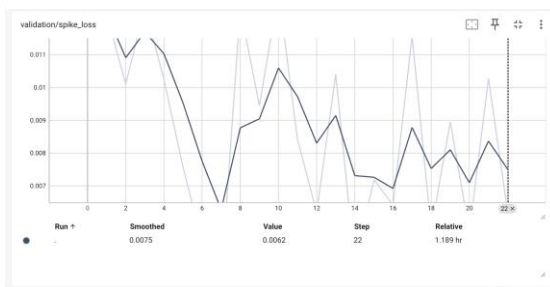


Figure 14: Validation/Spike Loss

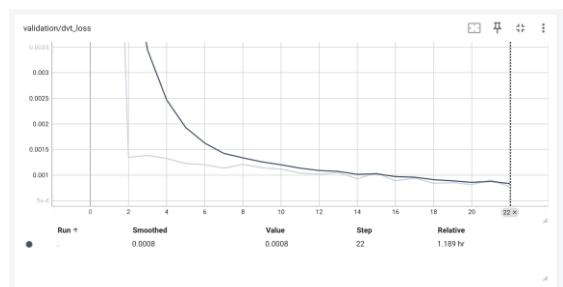


Figure 16: Validation/DVT Loss

6.4 外部入力の利用

昨年の研究などで用いた脳活動を特徴量として用いることで、実際の神経細胞の活動の推論を目指す。現在 HCI 分野にて脳波と行動の関連付けを行なっている。第一歩として Eye Tracking による人間の注視機能と脳活動の関連付けを行なっている。

また、人間の脳活動が人によらず一定であるという仮説を立て、個人によらず脳活動を定量化することにより、昨年の感情推論技術への統合を目指す。

7. 謝辞

本研究において、情報科学の達人プログラムより計算資源 ABCI の提供、MATLAB

の購入など、様々なサポートをしていただきました。前期メンターの大阪大学の浦西友樹さんと情報学研究所の杉山麿人さん、後期メンターの筑波大学の平木剛史さんと慶應義塾大学の鳴海紘也さんにこの場を借りてお礼を申し上げます。

8. プロジェクトの公開及び質問

本プロジェクトのソースコード等のアセットを学習目的で公開している。また、本論文で不明瞭な点やソースコードに関する質問は GitHub アカウント(@Tps-F)上に記載しているメールアドレスに送信して頂けると幸いです。



<https://github.com/Tps-F/CSnT>

9. 参考文献

[1] Gao, Z., Cui, X., Wan, W., Zheng, W., & Gu, Z. (2021). ECSMP: A dataset on emotion, cognition, sleep, and multi-model physiological signals. *Data in Brief*, 39, 107660. <https://doi.org/10.1016/j.dib.2021.107660>

[2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *arXiv preprint arXiv:1706.03762*. <https://doi.org/10.48550/arXiv.1706.03762>

[3] Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions

of synaptic input. *Journal of Neurophysiology*, 30(5), 1138–1168. <https://doi.org/10.1152/jn.1967.30.5.1138>

[4] Bi, G.-q., & Poo, M.-m. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, 18(24), 10464–10472. <https://doi.org/10.1523/JNEUROSCI.18-24-10464.1998>

[5] HODGKIN AL, HUXLEY AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*. 1952 Aug;117(4):500-44. doi: 10.1113/jphysiol.1952.sp004764. PMID: 12991237; PMCID: PMC1392413.

[6] Brette, R., & Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5), 3637–3642. <https://doi.org/10.1152/jn.00686.2005>

[7] Beniaguev, D., Segev, I., & London, M. (2021). Single cortical neurons as deep artificial neural networks. *Neuron*, 109(17), 2727–2739. [10.1016/j.neuron.2021.07.002](https://doi.org/10.1016/j.neuron.2021.07.002)

[8] A. Destexhe, Z. F. Mainen and T. J. Sejnowski, "An Efficient Method for Computing Synaptic Conductances Based on a Kinetic Model of Receptor Binding," in *Neural Computation*, vol. 6, no. 1, pp. 14-18, Jan. 1994, doi: 10.1162/neco.1994.6.1.14.