

電力効率化に向けたスパイキングニューラルネットワークの符号化手法に関する研究

令和5年度 修士論文

横浜市立大学大学院

生命ナノシステム科学研究科
物質システム科学専攻

白川雄也

目次

要旨

1. 序論

1-1: 研究背景

1-2: 目的

2. 関連研究

2-1: ニューロン

2-2: シナプス

2-3: スパイキングニューラルネットワーク

3. 研究手法

3-1: Diehl&Cook モデル

3-2: 提案手法

3-3: 評価指標

4. 実験概要

4-1: 実験方法

4-2: 結果

4-3: 考察

5. 結論

5-1: 結論

5-2: 今後の展望

6. 参考文献

7. 謝辞

電力効率化に向けたスパイキングニューラルネットワークの符号化手法に関する研究

白川 雄也

主指導教員

Ruggero Micheletto 先生

副指導教員

吉本和生先生

山田重樹先生

電力効率化に向けたスパイクニューラルネットワークの符号化手法に関する研究
A Study on Coding Methods of Spiking Neural Networks for Power Efficiency

物質システム科学専攻 白川雄也
指導教員 Micheletto Ruggero

スパイク:ニューロンが出力する電気信号
スパイクニューラルネットワーク (Spiking Neural Networks,SNN): 脳内のニューロン間のスパイク伝達の挙動を模したニューラルネットワークモデル。ニューロンがスパイクを介して情報伝達を行う。スパイクの数と消費電力は比例関係にある。
スパイクタイミング依存可塑性 (Spike-Timing-Dependent synaptic Plasticity,STDP):ニューロン間のスパイク発射のタイミングによって、結合強度の増加や減少が決定される学習則。
MNIST:手書き数字画像からなるデータセット。

【研究背景と目的】

脳は少ない電力で知覚し驚異的な力を発揮する。現在使用されるニューラルネットワークや深層学習は脳の特徴を模倣し作成されたが、これらネットワークと実際の脳の間にはわずかな類似性しか認められない。一方、より生物学的妥当性を持ち、専用ハードウェアで動作させることで実際の神経回路がもつ低消費電力の特性を示すとして SNN が注目されている [1]。

SNN は実際の脳が 1 ビットに近いスパイクを使用しているのと同様に 2 値信号であるスパイク列で情報伝達を行うためエネルギー消費が少ない [1]。また学習則は実際に存在すると示唆される STDP が活用されている。自動運転や IoT (Internet of Things) への深層学習の実装が増加しているが、性能の向上はモデルの複雑性や学習推論時間、パラメータ数と比例している。そのため開発や搭載の際にハードウェアのサイズや電力の制限を考慮する必要がある。この背景から比較的良好な性能を維持しながら消費電力を削減できる可能性がある SNN への関心が高まっている。SNN は現在主流の DNN を効率的に実装するための候補の一つであり、生物学的に妥当なメカニズムを用いて複雑な計算や認識を行う SNN が着目されている [2]。

しかし SNN の実世界への応用はまだ不十分である。その理由の一つとして、入力データを離散的なスパイクにどのように符号化するかという問題が存在する [1]。SNN はスパイクを使用し情報伝達を行うため学習対象のデータが画像の場合、学習・推論を行うためには画素値を離散的なスパイク列に変換する符号化が必要だ。最も用いられる手法は、各画素から実数値である画素値を取り出し、画素値に比例した発火率をもつポアソン過程から引き出されるスパイク列(ポアソンスパイク)に変換する手法だ [1] [2]。変換されたスパイクはポアソン分布に従うという手法であるが、符号化手法の研究は不十分である。SNN の特性を生かし、より電力効率の良い SNN を構築するためにも、新たな符号化手法の考案が求められている。

本研究では、SNN の符号化手法に着目し、電力効率化を目的とした符号化手法を考案。従来手法との分類精度や電力効率を評価・比較することで符号化手法が影響を及ぼすことを示し、従来手法より電力の効率化を達成することを目的とする。少ないスパイク数で高い精度を出力する電力効率に優れた SNN を目指す。

【方法】

SNN のネットワークとして Diehl&Cook モデル [2] を使用し、符号化手法を考案・変更することで、分類精度や学習に使用したスパイクの数、電力効率に対する影響を調べる。このモデルは STDP を学習則として使用する。入力には MNIST データセットを使用する。同数の興奮性ニューロンと抑制性ニューロンの集団が存在し、符号化された入力スパイクはそれぞれ全ての興奮性ニューロンに対して入力される。この間の重みが STDP で学習される。興奮性ニューロンは抑制性ニューロンと 1 対 1 で接続する。抑制性ニューロンの出力はペア以外全ての興

興奮性ニューロンに接続され、ペア以外の周りの興奮性ニューロンの発火を抑制する。これにより興奮性ニューロン間の競争を引き起こし、それぞれ異なる特徴を学習する。その後、興奮性ニューロンのスパイク数を取得し、設定した閾値よりスパイクの数が下回っていた場合は、発火率のレートを一定数増加させ、再度同じ画像で学習を行う。スパイク数が閾値より上回った場合、1つの画像の学習が終了する。再入力のプロセスは閾値を上回るまで繰り返される。全ての学習後、興奮性ニューロンの応答により画像分類を行う。ニューロンの数によって精度が比例し、最大 95%の精度を達成する。本研究では 100 個のニューロンを使用し、再入力のプロセスを行わない場合の結果も算出する。

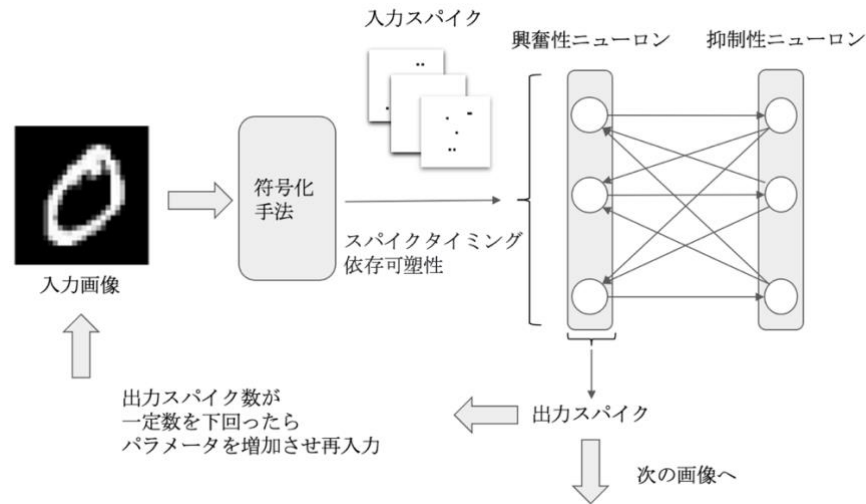


図 1 使用するネットワークのアーキテクチャ

符号化手法を考案するにあたって、実際の生物学的メカニズムを参考にする。また画像からスパイクへの変換手法だけでなく、どのように画像をスパイクへの変換手法に入力するかに着目する。なお画像への加工処理だけでは画像からスパイクの変換は行われなため、画像加工処理をした後、それぞれの分布による変換を行いスパイクの抽出をする。

【画像からスパイクの変換手法の提案】

SNN における符号化手法は主にポアソン過程によって生成されるポアソン分布に従うスパイク(ポアソンスパイク)を出力することで行われているが、これは脳の自発的発火が不規則であり、ポアソン分布はその不規則性を適切に記述することができる、すなわちスパイクの間隔がポアソン分布に従うとの仮定が存在するためだ。一方、近年は異なるのではないかと考えられており、脳の発火がガンマ分布に従っているのではないかと主張が存在する[3]。そこで本研究ではガンマ分布を用いた符号化手法を考案する。従来使用されるポアソン分布による符号化手法と同様、入力するデータセット画像の各画素値を取り出し、倍率(レート)を掛け合わせパラメータとし、ガンマ分布に従うスパイク(ガンマスパイク)を発生させる。

【入力前処理の提案】

〈エッジフィルター〉

人の網膜は物体の境界を抽出する働きをもつ。そこで使用する画像を加工し、輪郭を抽出した画像を作成した後、入力とすることでスパイクを抽出する手法を提案する。

〈抽出範囲の変更①〉

人の網膜では視細胞から視神経に向かうまでに横方向の情報を取り入れ統合され脳へ視覚情報が送られている。しかし従来の符号化手法では画素値ごとにスパイクが出力されているので画素値同士横方向の情報は統合されていない。また視細胞には大きいサイズの視細胞と小さいサイズの視細胞が交互に配置されている構造が確認できる。そこで大きい範囲(9つの画

素値)から画素値を合計し平均を算出し、小さい範囲(1つの画素値)からそのまま画素値を交互に取得することで、視細胞の構造を倣いつつ横方向の情報を取り入れ、スパイクを出力する手法を提案する。

〈抽出範囲の変更②〉

網膜では中心に小さな視細胞が、中心から離れるにつれ大きな視細胞が配置されているのを確認できる。そこで画像の外端に近づくにつれ大きい範囲から画素値を取得し、中心ではそのまま画素値を取得することで視細胞の構造を倣いつつ横方向の情報を取り入れ、スパイクを出力する手法を提案する。外端の部分は 3×3 の画素値の平均を取得する。中間は 2×2 の画素値の平均を取得する。中心はそのままの画素値を使用する。

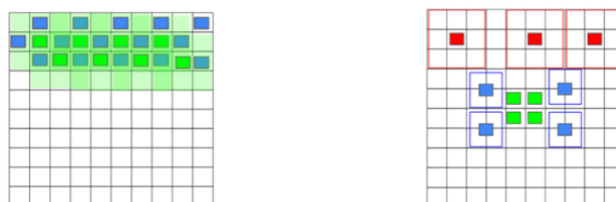


図 2 (左)は抽出範囲の変更①を表し、(右)は抽出範囲の変更②を表す

実験は処理を行わない場合を含む入力前処理 4 つと、画素値からスパイクへの変換手法 3 つを組み合わせることで 12 種類の符号化手法の精度や効率を調査する。使用データは MNIST である。手法ごとに異なるスパイク数を出力するため、画素値に掛ける倍率を変更し、手法ごと 6 つの結果を算出した。

【評価指標】

SNN においてはネットワークの学習で使用されるスパイクの数と消費電力は比例する。そのため学習に使用した全てのスパイク数を評価する。また電力効率を測定するため、学習に使用した全てのスパイク数とテストデータの精度から 1 スパイクあたりの精度を作成した。1 スパイクあたりの精度が高いほど高効率であると考えられる。

【結果と考察】

テストデータの精度が 0.5 以上を記録した手法を抽出し図 3、4 を作成した。入力前処理なしのガンマスパイクの精度は 0.5 以上を示しており、画像からスパイクへの符号化に活用できることが示された。抽出範囲の変更 1 とガンマスパイク、抽出範囲の変更 2 とガンマスパイクの組み合わせは精度が 0.5 を上回ることがなかった。入力前処理が精度に影響を与えることがわかる。精度が最も高い手法は、抽出範囲の変更 1 と再入力があるポアソンスパイクの組み合わせであった。この手法は、学習に使用した全てのスパイク数も最も多い結果となった。一方、抽出範囲の変更 1 と再入力がないポアソンスパイクの組み合わせは最も精度が低い結果となった。抽出範囲の変更 1 とポアソンスパイクの組み合わせは十分なスパイクの出力ができず、再入力機構の存在が精度に影響を与えたことが考えられる。また抽出範囲の変更 2 の手法も全ての場合において、スパイク数は低下しているものの精度も同様に低下しており、精度や効率性の向上に寄与しないことが考えられる。

スパイク数が最も少なかったのはエッジフィルターとガンマスパイクの組み合わせであった。電力効率を評価するために 1 スパイクあたりの精度を確認するとどの手法よりも高い値を示した。エッジフィルターにより入力画像の情報が削減され、精度や効率性が向上したと考えられるが、ポアソンスパイクにおいては向上が確認されなかったため、スパイクの種類によって適切な入力前処理が異なることが考えられる。引き続きの調査が必要だ。

- ▲入力前処理なし×ポアソンスパイク(再入力機構あり) ×エッジフィルター×ポアソンスパイク(再入力機構あり)
- ✱抽出範囲変更1×ポアソンスパイク(再入力機構あり) ●抽出範囲変更2×ポアソンスパイク(再入力機構あり)
- +入力前処理なし×ポアソンスパイク(再入力機構なし) ■エッジフィルター×ポアソンスパイク(再入力機構なし)
- 抽出範囲変更1×ポアソンスパイク(再入力機構なし) ◆抽出範囲変更2×ポアソンスパイク(再入力機構なし)
- 入力前処理なし×ガンマスパイク ●エッジフィルター×ガンマスパイク

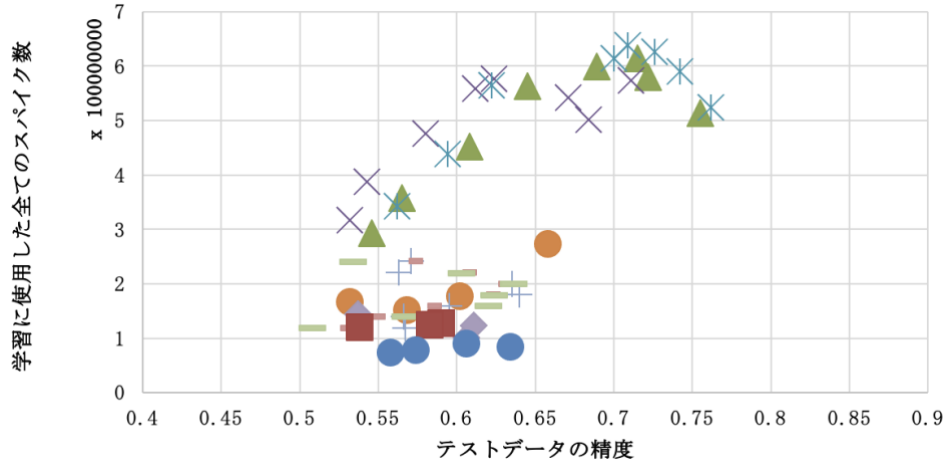


図3 学習に使用した全てのスパイク数とテストデータの精度の関係

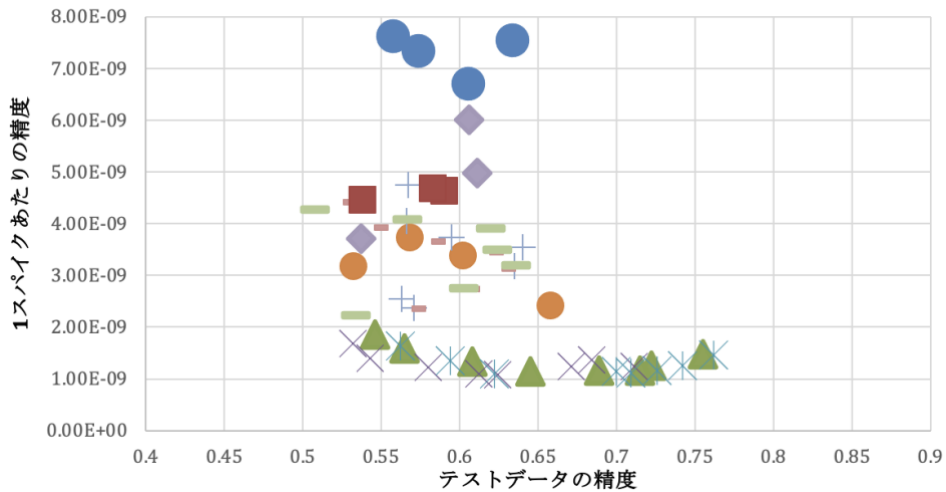


図4 1スパイクあたりの精度とテストデータの精度の関係

〈参考文献〉

- [1] Auge, Daniel & Hille, Julian & Mueller, Etienne & Knoll, Alois. (2021). A Survey of Encoding Techniques for Signal Processing in Spiking Neural Networks. *Neural Processing Letters*. 53.10.1007/s11063-021-10562-2.
- [2] P. U. Diehl, M. Cook, (2015). Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in computational neuroscience*, 9, 99.
- [3] Maimon, G. and Assad, J. (2009). Beyond poisson: Increased spike-time regularity across primate parietal cortex *Neuron*. 62, 426-440

1 章 序論

1-1 研究背景

人の脳はわずか 20W の消費電力で驚異的な力を発揮し知覚を行う。現在社会に広く普及し、多くの革新を起こしている人工ニューラルネットワーク (Artificial Neural Network : ANN) やディープラーニング (Deep Neural Network : DNN) は生物に見られる特徴を模倣されることで作成されてきた。近年、新たなニューラルネットワークの一つであるスパイクニューラルネットワーク (Spiking Neural Networks : SNN) が注目されている。専用のハードウェアで動作させることで脳のような実際の神経回路がもつ低消費電力、高速推論などの特性を示しつつ複雑な問題を解決できるからだ。[1][2] 現在の人工知能の成功に貢献している従来のニューラルネットワークと実際の脳の間には、わずかな類似性しか認められない。ANN や DNN は 32 ビットまたは 64 ビットで情報をやり取りしており、学習や推論における学習則として誤差逆伝播法が標準手法として使用されるが、このような現象は脳に実装されている可能性は低い。

SNN も同様に生物学における情報処理に着想を得ているがより生物学的妥当性を持つ。SNN は実際の脳が 1 ビットに近いスパイクを使用しているのと同様に 2 値信号であるスパイク列で情報伝達を行う。情報が処理されるときにのみスパイクが交換されるため、エネルギー消費が少ない[1]。またネットワークの学習の際に使用される学習則もより生物学に近いものである。SNN の学習則としては実際の脳に存在すると示唆されるスパイクタイミング依存可塑性が活用されている。生物学的に妥当なメカニズムを用いて複雑な計算を実行したり、パターン認識の課題を解決することは、人の脳がどのように働いているか理解するためにも必要とされており、これからの研究が不可欠である。なお ANN の学習結果を SNN のネットワークが活用できるように変換し精度を保つ手法も存在するが、誤差逆伝播法に依存することは変わらないため、生物学的に妥当ではない。性能の良さだけを追求できても、実際の脳の計算と比較できないため、理解にはつながらない。このような面からも、より生物学的に妥当なメカニズムを用いたパターン認識タスクのための SNN モデルが数多く開発され着目されている [3]。

今日、自動運転や IoT(Internet of Things)への DNN の実装が増加しているが、性能の向上はモデルの複雑性や学習推論時間、パラメータ数と比例している。そのためハードウェアのサイズや電力の制限のために、開発や搭載の際に考慮する必要がある。この背景から比較的良好な性能を維持しながら消費電力を削減できる可能性がある SNN への関心が高まっている。また SNN の専用ハードウェアであるニューロモルフィックハードウェアも開発が盛んとなっている。SNN は現在主流の DNN を効率的に実装するための候補の一つであり、研究がされている。

1-2 課題と目的

SNN は注目されているものの、実世界への応用はまだ不十分である。その理由の一部は、入力データを離散的なスパイクにどのように符号化するかという問題が存在するからだ[1]。SNN はスパイクを使用し情報伝達を行う。そのため学習対象のデータが画像の場合、学習・推論を行うためには画素値を離散的なスパイク列に変換する符号化が必要である。現在、SNN においてポアソン分布を利用した符号化が多く使用されるが、符号化に関する研究は不十分である。SNN の特性を生かし、より電力効率の良い SNN を構築するためにも、新たな符号化手法の考案が求められている[4]。

そこで本研究の目的としては、スパイクニューラルネットワークの符号化手法に着目し、電力効率化を目的とした符号化手法を考案。従来手法との分類精度や電力効率を評価・比較することで符号化手法が影響を及ぼすことを示し、従来手法より電力の効率化を達成することである。SNN においてネットワーク内のスパイク数が少なければ消費電力は少ないと考えられている[5]。そのため新たな符号化手法を考案し、少ないスパイク数で高い精度を出力する SNN を目指す。

符号化手法を考案するにあたっては、SNN と同様に実際の生物学的メカニズムを参考にし、網膜の機能や構造を模倣した手法を考案する。SNN において生物学にヒントを得ることは有力であると考えられている。脳は視覚、聴覚など、さまざまな感覚からのデータを表現するために、さまざまな符号化方式を用いる。とりわけ網膜は生物が外の世界を認知する場合、最初に視覚情報を電位として脳に送る機関であり、模倣するに最適であると考えた。

まとめると本研究では、網膜の機能や構造を模倣した符号化手法を考案し、各手法におけるスパイク数と精度の関係や、新たな指標を作成し各手法ごとの効率性を算出し評価を行う。またネットワークの学習に使用されるスパイク数より消費電力の概算を求め評価を行う。

2 章 関連研究

2-1:ニューロンとシナプス

2-1-1 ニューロン

脳はニューロンとよばれる神経細胞で構成され、ヒトの脳には約 860 億個のニューロンと約 100 兆個のシナプスが存在するとされる。

ニューロン自体の構造は大きく分けると樹状突起・細胞体・軸索である。樹状突起は他の神経細胞から信号を受け取る。細胞体は入力された信号を加算し、新たな出力信号を生成する。軸索は他のニューロンに信号出力を行う。ニューロンは電気的な素子として振る舞い膜電圧とよばれる電圧の値をパラメータとして持つ。ニューロンは他のニューロンからの入力や外部からの電流入力によって自分の膜電位の値を変化させ、膜電位がある一定の値を越えると、スパイクと呼ばれる短い電気パルスを発射する。ニューロン同士はシナプスを介して結合しており、ネットワークを構築している。スパイクがシナプスに達すると神経伝達物質とよばれる化学物質の放出を引き起こし、それが接続しているニューロンに到達すると電流が流れる。スパイクはビットのように 0 または 1 に対応するとされており、ビットを用いて情報通信をすることが脳の情報処理の原理であるとされている。

スパイクは Na⁺イオンと K⁺イオンがニューロンの細胞膜の表面にあるイオンチャネルと呼ばれる穴を通ることで生じる電流によって生成される。まずニューロンの細胞体の細胞膜内には K⁺イオンが多く存在し、膜外には Na⁺イオンが多く存在することで細胞膜内外の電位差が生まれる。この電位差は膜電位と呼ばれる。外部電流がなくニューロンが静止した状態では、膜電位の値は-65mv 付近の値をとる。この電位は静止膜電位と呼ばれる。細胞に刺激が加わると Na⁺イオンのチャネルが開き膜電位が上昇する。この上昇はほかに刺激がなく時間が経過すると落ち着き、静止膜電位になるが、さらに刺激が加えられ、膜電位が閾値を超えた場合、ナトリウムイオンがさらに加わることで、膜電位も上昇し正の電位に達する。膜電位の値が上がることを脱分極または発火と呼ばれ、この時の膜電位は活動電位またはスパイクと呼ばれる。脱分極後はイオンが戻り静止膜電位に戻る。膜電位の値が下がることを過分極とよぶ。この際、膜電位は静止電位よりも過分極されるが、この期間は不応期と呼ばれ、次のスパイクは発射ができない。またニューロンには大きく分けて興奮性ニューロンと抑制性ニューロンの 2 種類があり、異なる神経伝達物質を放出する。[6]

このようにニューロンはスパイクを作りスパイクで情報伝達を行っているが、スパイクの波形自体は情報を持たず、スパイク有無や発火頻度、発火時刻が重要であるとされる。なお発火を示す 0 と 1 の 2 値信号を時間的に連続したスパイク信号はスパイク列またはスパイクトレインと呼ばれる。

2-1-2 ニューロンモデルについて

脳がネットワークとしてどのように複雑な機能を生み出しているか、機能や原理は明らかではない。しかしネットワークを構成する単一ニューロンについては、具体的にその挙動を微分方程式で記述することができる[6]。さまざまなニューロンモデルが、膜電位の挙動とスパイク発生を異なるレベルで詳細に記述しており、生物学的な情報処理のモデルとして多くのモデルが研究されている。現在 DNN に多く使用されている McCulloch & Pitts モデル。単一ニューロンの世界で最初の数理モデルである Hodgkin Huxley モデル。Hodgkin Huxley モデルはイオンチャネルの挙動を細かく定義することでニューロンの挙動、スパイク発射の挙動を再現しており、神経科学の解析や再現に使われる。最も多く使用されるメジャーなモデルである積分発火(Leaky Integrate and Fire)モデル[4]。単一ニューロンの挙動を数学的に解析することで作成された、実際のようなニューロンの挙動を再現可能な Izhikevich モデルなど、それぞれのモデルが膜電位やスパイクの発生を異なるレベルで記述している。一般に共通しているのは膜電位が到着したスパイクからの電流を統合し、ある閾値を越えるたびに新しいスパイクを発生させる構造である。スパイクが生成されると膜電位はリセットされる[5]。膜電位の基本的なダイナミクスは以下のように記述される[6]。

$$C \frac{dV}{dt} = -\bar{g}_{leak}(V(t) - E_{leak}) + I_{ext}(t) \quad (1)$$

$V(t)$ は時刻 t における膜電位[mV]、 $I_{ext}(t)$ は外部から流入する電流[$\mu A/cm^2$]。 C 、 \bar{g}_{leak} 、 E_{leak} は定数でありそれぞれキャパシタンス[$\mu F/cm^2$]、コンダクタンス(抵抗の逆数)[mS/cm^2]、反転電位[mV]とよばれる。この定式化はニューロンを電気回路とみなすことと同じである。この定式化において外部電流がない状態では膜電位 $V(t)$ の値は E_{leak} となり、一定の電流 I_{ext} を流すと時定数 $\tau_m = C/\bar{g}_{leak}$ で、 $V(t) = E_{leak} + I_{ext}/\bar{g}_{leak}$ へと収束する。外部電流を止めると $V(t)$ は再び E_{leak} に戻る。この状態では膜電位の上限が起きるのみでスパイクは発射されない。スパイクは膜の表面に存在するイオンチャネルとよばれる穴を Na^+ イオンと K^+ イオンが通ることによって生じる電流により生成される。それぞれの電流 $I_{Na}(t)$ 、 $I_K(t)$ は以下の式で表すことができる。

$$I_{Na}(t) = -g_{Na}(V, t)(V(t) - E_{Na}) \quad (2)$$

$$I_K(t) = -g_K(V, t)(V(t) - E_K) \quad (3)$$

ここで $g_{Na}(V, t)$ 、 $g_K(V, t)$ はそれぞれ Na^+ チャネルと K^+ チャネルのコンダクタンスであり時間と膜電位の関数。 E_{Na} 、 E_K はそれぞれのイオンチャネルの反転電位である。膜電位と反転電位の差が外力となり、コンダクタンスとの積でチャネルを流れる電流となる。コンダクタンスは抵抗の逆数であるので上記の2式はオームの法則と等しい。コンダクタンス値の変化によって膜電位の値が時間的に急速に上下することでスパイクが生成される。 Na^+ イオンと K^+ イオン、加えて存在する Cl^- イオンの電流を全て考慮すると外部電流がなくニューロンが静止状態の場合では、膜電位は-65mVの値をとる。なおコンダクタンスの式は動物の種類やニューロンの種類ごとに異なる。このようにキャパシタンスやコンダクタンスが記述されるモデルをコンダクタンスベースモデルとよぶ。存在しない場合はカレントベースモデルとよばれる。

2-1-4 積分発火型モデル

積分発火型(Leaky integrate-and-fire:LIF)モデルは1変数のみをパラメータにもつシンプルなニューロンモデルである。また McCulloch & Pitts モデルより本来の神経細胞の挙動に近いニューロンモデルであり、SNN や神経回路シミュレーションで多く使用される。

多くの生理実験がスパイクの個数や発射のタイミングが情報処理を担うことを示している。そのためスパイクの波形を考慮せず、スパイクを発射した時刻のみを提供する目的で提案されたのがこのモデルである。スパイク生成の計算を省略し変数を減らしたモデルである。

積分発火型モデルは、スパイクは閾値とリセット電位により発生する。他のニューロンからスパイクが入力され、重みに応じて膜電位が変化し、発火閾値を超えた際に発火する。出力はスパイク列となる。LIF モデルは次の式で記述される[6]。

$$\tau \frac{dv}{dt} = -(v(t) - V_{rest}) + RI_{ext}(t) \quad (4)$$

$$v(t) > \theta \Rightarrow S(t) = 1, v(t) \leftarrow V_{reset} \quad (5)$$

$$v(0) = V_{init} \quad (6)$$

ここで、 $v(t)$ [mV]は時刻 t での膜電位、 τ [ms]はどの程度漏れが生じるかを制御する膜時定数、 V_{rest} [mv]は静止電位、 R [M Ω]は膜の抵抗、 $I_{ext}(t)$ [nA]は時刻 t での外部電流、 θ [mv]はスパイク発射のための閾値、 V_{reset} [mv]はリセット電位、 V_{init} [mv]は膜電位の初期値である。(4)が基本的な膜電位のダイナミクスを記述する。膜電位は V_{rest} を平衡点とし、外部入力 $RI_{ext}(t)$ を加えた $V_{rest} + RI_{ext}(t)$ に時定数 τ で漸近する挙動を示す。(5)はスパイク発射の条件である。膜電位が閾値を越えると、その時刻でスパイクを発射したものとし $S(t) = 1$ かつ膜電位をリセットする。(6)は膜電位の初期値を与える。

閾値とリセット電位は定数なので変数は膜電位のみとなる。積分発火型モデルは入力がない場合、膜電位は指数関数的に静止状態まで減衰する。入力スパイクにより、各入力膜電位 $v(t)$ に膜電位をもたらし、膜電位が閾値に向かって上昇し、膜電位が閾値に達するたびに、膜電位 $v(t)$ は V_{reset} にリセットされ、スパイクが発火するという挙動を繰り返す[14]。ニューロンにおける不応期は、1回の出力スパイクの後に生じる。積分発火型モデルはスパイクの具体的な波形は計算せず、スパイクを発射した時刻のみを提供する[6]。

2-2 シナプス

2-2-1 人のシナプス

人の脳には約 100 兆個のシナプスが存在するといわれている。シナプスはニューロン同士の結合ポイントである。軸索と他の細胞の樹状突起とが結合し、結合の強度が調整される部位である。ニューロンはシナプスと呼ばれる構造を作って互いに結合しており、ネットワークをなしている。

送り手側のニューロンがスパイクを発射すると、信号が軸索末端に伝播し、結合場所であるシナプス間隙に神経伝達物質が放出される。送り手側のニューロンが興奮性ニューロンの場合(興奮性シナプス)はグルタミン酸が、抑制性ニューロンの場合(抑制性シナプス)は GABA (γ アミノ酸)が放出される。受け手側のニューロンには、樹状突起の表面に神経伝達物質に対する受容体が存在しており、神経伝達物質を受け取るとイオンチャンネルに働きかけて電流を発生させる。興奮性シナプスからは脱分極させる電流が、抑制性シナプスからは脱分極させる電流が発生する。それによりポスト側のニューロンの膜電位が変化する。これらをそれぞれ興奮性後シナプス電位、抑制性後シナプス電位とよぶ[6]。

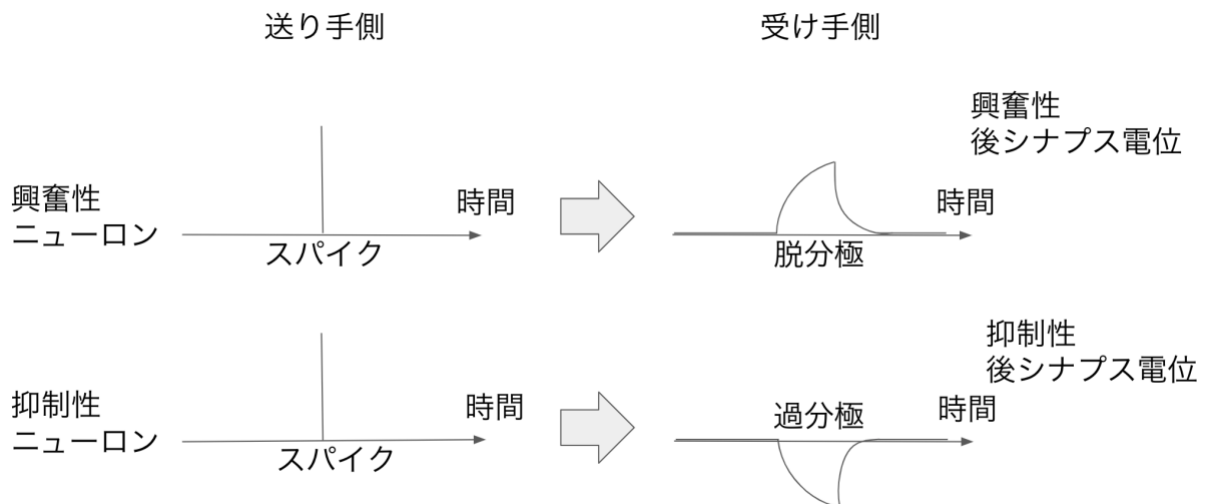


図1 シナプスにおける電位発生

2-2-2 シナプスのモデル

スパイクの伝播による膜電位の変化は以下で示すシナプス電流を加えることで表現される。プレ側のニューロンが活性化すると、ポスト側のニューロンにシナプス電流という形で、シナプスの重さ、つまりコンダクタンスに比例する信号を送る。一般に、シナプス電流がポスト側のニューロンに到達すると、その膜電位を変化させる。膜電位があらかじめ定義された閾値に達すると、ポスト側のニューロンはスパイクを放出し、その膜電位を静止電位にリセットする[4]。

シナプス電流 $I_{syn}(t)$ もイオン電流と同様にシナプスコンダクタンス $g_{syn}(t)$ 、反転電位 E_{syn} を用いて計算される。反転電位 E_{syn} は静止電位が -65mv の場合、興奮性シナプスでは 0mv 、抑制性シナプスでは -65 から -80mv となる。

$$I_{syn}(t) = -g_{syn}(V(t) - E_{syn}) \quad (7)$$

(7)で表される電流をポスト側のニューロンに加える。シナプスコンダクタンス $g_{syn}(t)$ の値はプレ側のニューロンからスパイクを受け取ると変化し、何もなければゼロに減衰する。シナプスの計算はコンダクタンスの値を変化させることである。値の変化の時定数はシナプスの種類で決まる。興奮性シナプスの場合は大きく AMPA 型と NMDA 型に分かれ、AMPA 型は数 ms 程度の短い時定数。NMDA 型は数十~百 ms 程度の長い時定数をもつ。抑制性シナプスの場合は GABAa, GABAb に分かれ前者は短い時定数、後者は長い時定数をそれぞれ持つ。1 個の受け手側のニューロンに 1 個の送り手側のニューロンが結合していることを考える。一般的にシナプスの基本的なモデルは指数減衰シナプスとよばれ、コンダクタンス値 $g_{syn}(t)$ は一般的に以下の式で計算される。

$$g_{syn}(t) = \bar{g}_{syn} \sum_{f \in S(t)} \exp\left(-\frac{t-t^{(f)}}{\tau}\right) \Theta(t-t^{(f)}) \quad (8)$$

\bar{g}_{syn} はスパイク 1 つあたりの最大コンダクタンス値で、神経伝達物質の放出量や密度などで決まる定数。送り手側のニューロンはスパイク列 $S(t)$ を出力するとして、スパイク f の出力時刻を $t^{(f)}$ とする。 τ は時定数でスパイク 1 つあたりのコンダクタンス値の減衰を表す。コンダクタンス値が $1/e$ 倍になるまでの値となる。

$\Theta(x)$ はヘビサイドのステップ関数で $\Theta(x) = 1 (x \geq 0)$ or $0 (x < 0)$ である。

送り手側のニューロンが単発のスパイクを出力するとコンダクタンスは瞬時に上昇し、その後時定数 τ に従って減衰し 0 に戻る。その際コンダクタンスと膜電位の反転電位からの差の積によってシナプス電流が発生する。複数のスパイクが同時または短い時間間隔で到達するとシナプス電流の値はその重ね合わせになる。

なおカレントベースのシナプスにおいては、 $g_{syn}(t)$ を直接シナプス電流の値として考え、以下の通りにする。

$$I_{syn}(t) = w \times g_{syn}(t) \quad (9)$$

ここで w は結合強度を表す定数であり、興奮性結合では正、抑制性結合では負の値をとる。送り手側のニューロンからのスパイクの個数がそのまま電流の強度になる。送り手側の膜電位の値を必要としないため、計算が簡単になる。カレントベースの LIF ニューロンで利用されることが多い。 [6]

2-3 スパイキングニューラルネットワーク

2-3-1 スパイキングニューラルネットワークの概要

スパイキングニューラルネットワーク (SNN) は、脳内のニューロン間のスパイク伝達の挙動を模した低消費電力な機械学習モデルである。ニューロンがスパイクを介して情報を交換し、生物学的な情報処理のモデルとしても研究される。SNN は低消費電力、高速推論などの実際の脳のような神経回路で示される有利な特性を示すため、機器の消費電力削減、物理的制約の解決に向け有望とされている。またエネルギー効率の高いニューロモーフィックハードウェアデバイスを利用するのに適した計算モデルであり、組み合わせることでより従来の ANN より効率的になることが期待されている [7]。SNN では入力信号を時系列のバイナリ信号 {0/1} に変換して処理を行う。バイナリ信号であるスパイクの有無・発生頻度を通じて情報伝達を行う。この情報伝達において、スパイクは同じものであり、スパイク単体では何の情報も持たず、スパイクの数とタイミングが重要である [1]。

高効率が着目される一方で実用化に向け改善するために多くの研究が必要とされている。その一つが符号化手法である。SNN を学習させるためには、入力データをバイナリ信号であるスパイク列に変換する必要がある。スパイク列で構築されたデータセットは少なく、また既存のデータセットを活用するためにも符号化手法の研究は必要である。本研究では符号化手法に着目する。

2-3-2 学習則

現在使用される ANN は主に誤差逆伝播法という学習則が使用されているが SNN には適用できない。バイナリであるスパイク列を活用するため、勾配を求めることができないからだ。誤差逆伝播法を SNN に適用するための研究も存在する。Wu らはスパイクニューロンを使用したディープラーニングにおいて、従来の誤差逆伝播法のアルゴリズムを拡張して適用することで高い精度が得られることを示した[8]。また Yang らは畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を学習後、Spiking Convolutional Neural Networks (SCNN) 変換するための一般的な変換技術を提案した[9]。

一方誤差逆伝播法のメカニズムは実際の脳には搭載されていないと考えられており、SNN においては生物学的に妥当な学習側のスパイクタイミング依存可塑性 (Spike Timing Dependent Plasticity : STDP) が多く使用される。STDP は、1949 年にドナルド・ヘブがシナプス可塑性の基本原則を提唱し、脳内でどのように学習が達成されるかを説明したことから始まった一連の神経生物学的知見から生まれた。STDP 学習則では学習のため 2 つのニューロンの発火時刻差に依存したシナプス結合強化が行われる[4]。

2-3-3 スパイクタイミング依存可塑性 (STDP)

脳は学習によって自分自身の挙動を変化させている。神経における学習とは、ニューロン間の結合の強さを変化させることである。つまり記憶はシナプ스에蓄積されており、シナプス結合の強さを変えることで記憶を書き換える。この仕組みをシナプス可塑性とよぶ。

シナプス可塑性の基本的なメカニズムとして **Hebb** が提唱したヘブ則がある。シナプス結合したニューロンが同時に発火したらその間のシナプス結合を強くするというものである。実際の神経回路で同様の現象が発見されたので解明され確立された。実際にはヘブ則のみでは結合は一方的に強くなるだけなので、安定的に学習を行うために結合を弱める仕組みも必要である。ヘブ則はニューロンの発火頻度に基づいたコンセプトであるが、それをスパイク 1 発 1 発に発展させたものがスパイクタイミング依存可塑性 (STDP) である。

STDP では、プレ側ニューロンとポスト側ニューロンの発火のタイミングが重要となり、タイミングによって結合強度の増加・減少が決定される。プレ側の発火が受けて側の発火に関係が認められるか否かで結合の強弱が変化する。プレ側の発火がポスト側の発火より少しだけ先行して起これば結合を強くする。これはプレ側からのスパイク入力によってポスト側の発火が引き起こされる発火順序の関係が生じた時に結合が強くなることを意味する。発火時間差が小さいほど、結合は大きくなる。逆にポスト側の発火の方がプレ側より少しだけ発火して起これば結合を弱くする。これはポスト側の発火はプレ側によって引き起こされていない発火順序の関係が生じた時に、結合が弱くなることを意味する。発火時間差が大きいほど結合は小さくなり、抑制性になる[6]。

SNN では、離散的なスパイク列を使用しているため学習則として誤差逆伝播法は使用できない。そのため STDP による学習側が多く使用されており、最も有名である。[10]

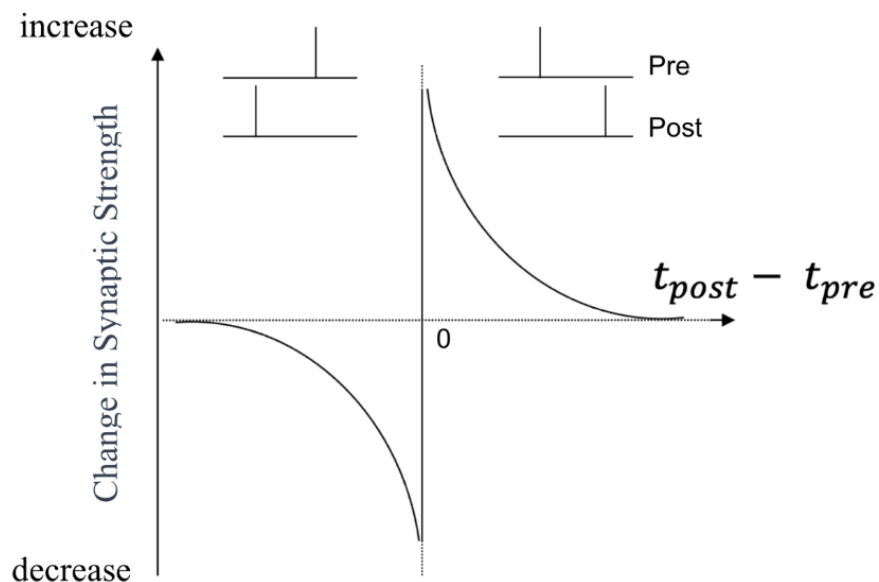


図2 STDPにおける結合強度の変化について
横軸は結合ニューロンの発火の時間差、縦軸は結合強度

STDPの実装においては、トレースという値を用意する。ここで i, j はそれぞれ受け手側、送り手側のニューロンの添字とし、そのニューロン間のシナプスの結合強度 w_{ij} を考える。送り手側のトレース x_j を以下のように定義する。

$$x_j(t) = \exp\left(-\frac{\Delta t}{\tau_{pre}}\right) x_j(t - \Delta t) + S_j(t) \quad (10)$$

τ_{pre} は減衰の時定数、 $S_j(t)$ は時刻 t でニューロンがスパイクを発射したら 1、発射しなければ 0。スパイクを発射するとトレース値は上昇し、その後は指数的に減衰する。受け手側のトレース y_i も考える。

$$y_i(t) = \exp\left(-\frac{\Delta t}{\tau_{post}}\right) y_i(t - \Delta t) + S_i(t) \quad (11)$$

τ_{post} は減衰の時定数である。トレースが STDP の指数関数の窓を表す。つまり時定数が STDP の窓関数の形状を規定する。これらの値を用いて結合強度を以下の通りに更新する。 $-A_-$ 、 A_+ は定数である。

$$w_{ij}(t + \Delta t) = w_{ij}(t) + \Delta w_{ij}(t) \quad (12)$$

$$\Delta w_{ij}(t) = -A_- y_i(t) S_j(t) + A_+ x_j(t) S_i(t) \quad (13)$$

送り手側のスパイク 1 つに対して、その時の受け手側のトレース値を引き、ポスト側のスパイク 1 つに対して、その時の送り手側のトレースの値を足す[6]。

2-3-4 SNN における符号化手法

SNN はスパイクで情報伝達、すなわち学習や推論を行うため、入力データからスパイクへの何らかの変換が必要となる。スパイクを出力する専用のセンサーならば変換は必要としないが、多くの場合は従来のデータセットを使用することが多いため符号化手法は必須となる。

画像に関連するタスクを実行する際には、従来の ANN ではそのまま画素値を用いて特徴を抽出するが、SNN においては特徴抽出の前に画像をスパイク列に変換することが最初の手順となる。

最も広く用いられている手法は、各画素から実数値である画素値(強度)を取り出し、比例した発火率をもつポアソン過程から引き出されるスパイク列(ポアソンスパイク)に変換することである[1][3][5]。変換されたスパイクはポアソン分布に従う。放出されたスパイクの数が多いと、関連する画素値が高いことに対応する。

実際の脳内、とりわけ大脳皮質で起きている活動電位のタイミング、スパイク発火のタイミングは不規則であるとされている。ポアソン分布はスパイクを生成するための不規則性を適切に記述することができると考えられていることもありこの手法が使用されることが多い。SNN の特徴でもある生物学的妥当性を保持するためにも、画像をスパイク列に変換する際に使用される[7]。

スパイクが従うポアソン分布の確率質量関数は以下のように表される[8]。

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!} \quad (14)$$

x はカウントデータなどの確率変数で、 λ はポアソン分布の強度である。ポアソン分布は 1 個 2 個や 1 回 2 回といったカウントデータが従う確率分布である。ある事象が一定の時間内に発生する回数を表す。具体的な使用例としては、釣り道具を変えることで釣獲尾数が変わるのか、天気によって商品の販売個数がどれほど変わるのかということを調べる際にポアソン分布が活用される。[9]

現在多くの研究で使用されるのは画素値が大きいほどスパイク数が多く、密なスパイク列が出力され、画素値が低いほどスパイク数が少なく疎なスパイク列が出力されるという手法である。その場合画素値をからパラメータとして扱い、パラメータの大きさを調節するレート(倍率)と掛け合わせ、 λ として扱いスパイクを出力する。生成されたスパイク列はポアソン分布に従い、大脳皮質の不規則なスパイクをシミュレートするのに適しているとされる[7]。

SNN はまだ発展途上であるため、符号化方式とニューロンモデルおよびネットワークアーキテクチャを組み合わせる最も効果的な方法は、まだ発見されていない。しかし、適切に符号化されなければ、ネットワーク層を伝播する際に情報が失われる可能性があるため、これは非常に重要である。[4]

2-3-5 ニューロモルフィックと消費電力

機械学習を実装した製品を開発する企業にとって、エネルギー消費は大きなコスト要因である。そのため消費電力を低減することはこれからの将来に必須であり、着目されるのが SNN であり、SNN は専用のハードウェアであるニューロモルフィックチップで作動させることで脳

のような実際の神経回路がもつ低消費電力、高速推論などの特性を示す。またニューロモルフィックチップに実装されたスパイクニューロンを使用したディープラーニングも報告されており、高速かつ省電力な稼働が可能であることが示されている[1]。

現在の使用される CPU はクロック数が数 GHz というスピードで動いているが、ニューロンはせいぜい 1000spikes/s でしかスパイクを発火できない。そのためオーダーが大幅に異なる。CPU はクロックごとにコンデンサの充放電で主に電力を消費するため、クロック数に比例して電力が決定される。一方ニューロンの場合は発火時にイオン電流が流れて電力を消費する。そのためニューロンの 10^6 以上も低い動作速度は低消費電力の要素の 1 つになっている。また現在の計算機アーキテクチャは演算器とメモリが分離しバスを介してメモリを読み書きしているが、バス上のメモリ転送に多くの電力を消費する。バスは複数の回路や装置、機器間を結びデータなどのやりとりのために共有される伝送路のことである。一方ニューロンの場合はメモリに相当するのはシナプス結合の強度であり、演算器に相当するニューロンと一体となっている。したがってメモリというよりは大量のレジスタでありいわゆる外部メモリは存在していない。外部メモリアクセスに相当するのはニューロン間のスパイクの伝播であり、ここで転送されるデータはスパイクなのでわずか 1bit となる。膨大な数のスパイクを伝播させてしまえばメモリアクセスが増えてしまうが、発火頻度を低く抑えることで結果的に消費電力を抑えることができる[1]。

ニューロモルフィックハードウェア上の SNN の現在の実装は、スパイクを送信するために数 nJ または pJ さえ使用するだけとされる。いくつかのセットアップはスパイクあたり 0.02pJ、およびシナプスあたり数 pW という小さなエネルギーを使用するのみである[3]。この低消費電力は電力が重要課題となる現代のエレクトロニクス分野や、エッジコンピューティングに有用であるとされ、人の脳が示すような高度な情報処理を人の脳と同程度の消費電力で実現することを目指しニューロモルフィックチップが開発され続けている。

2-3-6 SNN の利用や展望

現在、ニューラルネットワークは自動運転や Iot、ロボット工学など幅広い分野で活躍しているが、普及率の増加や消費電力への関心が高まるにつれ、エネルギー効率が重視され始めている。スパイクニューラルネットワークはニューロモルフィックチップ上で作動させ

ると低消費電力や高速推論といった特性を示すため、従来のニューラルネットワークに代わる可能性を秘めており、多岐にわたる利用が考えられることは先に述べた[5]。高速推論、低消費電力はリソースが限られた環境での高速な画像認識や音声認識を可能とする。

高速推論という特性はリアルタイム性を向上させることができる。例えば自動運転技術の場合、特に物体の検出や追跡などの画像解析タスクに役立つ。センサー情報をリアルタイムで処理し、厳しい光条件下で突然現れる歩行者に対処する緊急ブレーキアシストシステムが考えられる。刻々と変化する環境に対応しなければならないシステム、すなわち、一度学習した重みを使って推論するだけでは不十分な場合も有効である。事前に学習させた上でユーザーのアクセントに適応させる音声認識システムが考えられる。

低消費電力という特性は電力制約を受けるスマートフォンなどのモバイル機器や時計や眼鏡などのウェアラブル機器、ドローンなどの移動・空中ロボット、24時間稼働し続けるIoT機器への応用が考えられる。医療分野にも応用が考えられる。例えば脳波解析や脳内のインプラント、ブレイン・マシン・インターフェースである[5]。

低消費電力は近年求められる脱炭素にも効果的である。火力発電に依存する日本であれば一層効果的だ。またスパイクニューラルネットワークは生物学的情報処理に着想を得ており、生物学的忠実性が高いため、脳科学や神経科学の研究にも役立つ。

視覚情報をスパイクに符号化するビデオカメラや音声情報を符号化するオーディオセンサーの開発も盛んであり、積分発火モデルを使用し、センサーで取得したスパイクから情報を再構築する研究も存在する[7]。

このようにスパイクニューラルネットワークが活躍できる分野は数多く存在する。しかし現在のスパイクニューラルネットワークは多くの乗り越えるべき障壁が存在するため、実世界に応用するために多くの研究が必要とされている。従来のニューラルネットワークに並ぶ精度と低消費電力の両方において高い性能を実現するためにも引き続きの研究が必要だ。

3 章 研究手法

3-1 Diehl&Cook モデル[3]

このモデルは STDP のような生物学的に妥当なメカニズムの組み合わせを使用し、ネットワークの重みをラベルなしで入力例の構造を学習する。学習には MNIST データセット [11] を使用しており、教師なし学習である。また使用されるニューロンの数によって精度も比例し、6400 個のニューロンを使用した際には MNIST ベンチマークで最高 95% の精度を達成した。このモデルではニューロンの間に競合状態を作ることによってデータの分類を学習している。このモデルの特徴としては積分発火型 (LIF) ニューロン、STDP、横方向抑制、固有可塑性を用いていることである。

Diehl&Cook モデルで使用されるニューロンモデルとシナプスについて説明する。ニューロンには積分発火型ニューロンが使用されており、膜電位 V は以下の式で表すことができる。

$$\tau \frac{dV}{dt} = (E_{rest} - V) + g_e(E_{exc} - V) + g_i(E_{inh} - V) \quad (15)$$

$rest$ は静止膜電位、 E_{exc} と E_{inh} は興奮性シナプスと抑制性シナプスの平衡電位、 g_e と g_i はそれぞれ興奮性シナプスと抑制性シナプスのコンダクタンスである。時定数 τ を用い、これは興奮性ニューロンの方が抑制性ニューロンよりも大きい。ニューロンの膜電位がその膜閾値 v_{thres} を越えると、ニューロンは発火する。発火後は膜電位は v_{reset} にリセットされる。リセット後の次の数ミリ秒以内に、ニューロンは不応期となり、再びスパイクすることができなくなる。シナプスはコンダクタンスの変化によってモデル化され、送り手側からのスパイクがシナプスに到達すると、シナプスの重み w だけ瞬間的にコンダクタンスを増加させるが、そうでなければコンダクタンスは指数関数的に減衰する。送り手側のニューロンが興奮性であれば、コンダクタンス g_e のダイナミクスは次のようになる。

$$\tau_{g_e} \frac{dg_e}{dt} = -g_e \quad (16)$$

τ_{g_e} は興奮性シナプス後電位の時定数である。また受け手側のニューロンが抑制性である場合、コンダクタンス g_i は同じ式を用いるが、抑制性シナプス後電位の時定数 τ_{g_i} を用いる。Diehl&Cook モデルでは膜の時定数、シナプスの時定数など、パラメータのほとんど全てに生物学的に妥当な範囲を使用しているが、分類精度向上のため興奮性ニューロンの膜電圧の時定数は例外で通常観察される値よりも 5 倍程度増加させ利用している。

またニューロンの発火閾値は発火数と比例しながら増加し、時間と共に減衰するように設定されている。発火すればするほど増加し、スパイクを発火させるためにより多くの入力を必要とする。なお発火をしない時には、時間と共に指数関数的に減衰するため、一定時間立つと発火頻度は元に戻る。これは、ネットワーク全体の発火頻度を一定にし、一つのニューロンだけが学習をし続けないようにするためである。この機構が各ニューロンに公平な競争の機会を与える。

Diehl&Cook モデルのネットワーク構造を説明する。ネットワークは 2 つに分かれている。画像からスパイクへの符号化が行われる入力層とスパイクから情報を取得し処理する処理層である。さらに処理層には同数の興奮性ニューロンと抑制性ニューロンの集団が存在する。まず符号化が行われる入力層では、画像中の画素ごとに画素値に比例する発火率を作成しポア

ソン分布に従うスパイク列を出力する。この発火率は画素値に、大きさを調節するレート(倍率)を掛け合わせて作成される。そしてそのスパイク列は 350ms の間、処理層の興奮性ニューロンに提示される。この入力層と興奮性ニューロンの間の重みが STDP で学習され、入力データに応じて興奮性ニューロンの発火のしやすさが定まることになる。次にスパイク列はそれぞれ全ての興奮性ニューロンに対して接続し興奮性ニューロンに供給される。供給された興奮性ニューロンから抑制性ニューロンへ 1 対 1 で接続を行い、ペアを作成する。興奮性ニューロンから出力された各スパイクは、対応する抑制性ニューロンのスパイクを誘発する。抑制性ニューロンの出力は接続されたペア以外全ての興奮性ニューロンに接続されている。この接続が横方向の抑制をもたらし、興奮性ニューロン同士の競争を引き起こす。抑制性ニューロンは負の重みである結合をもつため、スパイクが残りの興奮性ニューロンに送られる際、興奮性ニューロンの膜電位が低下し発火の抑制を行う。つまり 1 つのニューロンが発火した際、それ以外の周りのニューロンの発火を抑制する。これにより 1 つの興奮性ニューロンだけが発火しやすくなるため、この接続性が横方向の抑制をもたらすことで、ニューロン間の競争を引き起こし、それぞれのニューロンが異なる特徴を学習することができる。抑制性ニューロンが興奮性ニューロンに出力した後、興奮性ニューロンのスパイクの数を取得し、入力のスパイク列が提示される 350ms の間、十分に発火が行われているか確認する。もし設定した閾値より下回っていた場合は、発火率の倍率を一定数増加させ、再度これまでの学習を行う。再度学習を行い、興奮性ニューロンのスパイク数が設定した閾値より上回った場合、一つの画像の学習が終了する。この閾値は 5 と設定されている。なおこのプロセスは閾値を上回るまで繰り返される。その後次の新しい画像を提示する前に、すべてのニューロンの発火閾値を除く変数とその静止値に減衰するように、何も入力しない段階が 150ms 秒存在する。

データセット全体の学習が終わると、学習率を 0 にし、各興奮性ニューロンの発火閾値を固定し、学習用データセットの 1 回の提示における 10 クラスの数字に対しての最も高い応答すなわち発火率に基づき、各ニューロンにクラスラベルを割り当てる。ここがクラスラベルを使用する唯一のステップであり、シナプスの重みの学習にはクラスラベルを使用しないため教示なし学習とされる。クラスラベルが割り当てられたニューロンの反応が、テスト用データセットにおける分類精度を測定するために使用される。予測されるクラスごとに各ニューロンの反応を平均し、平均発火率が最も高いクラスが選択されることで決定される。精度はネットワーク内のニューロン数に比例し、6400 個のニューロンにおいては 95%の精度を実現する。ネットワーク内のニューロン数が多いほどピークの性能に達するまで長い学習時間を必要とする。

Diehl&Cook モデルは、現在も使用される深層学習より精度は低いもののネットワークのシンプルさや STDP での学習が使用されていることから大きな注目を集め、STDP を使用したスパイクニューラルネットワークの代表として取り上げられることが多い。

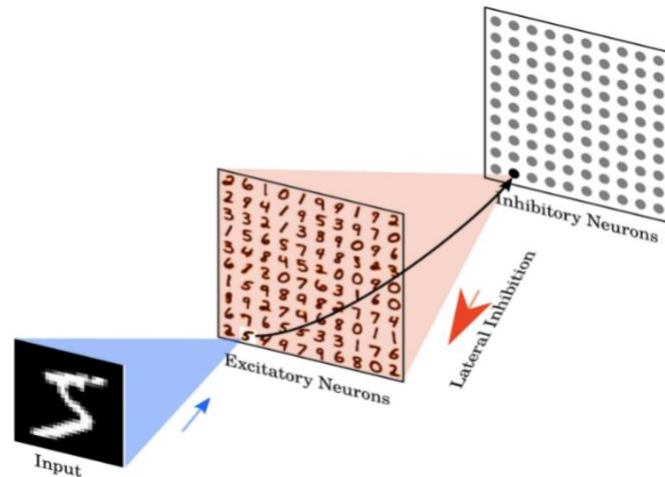


図3 ネットワークアーキテクチャ。[3]より出典

28×28画素のMNIST画像の画素値は、対応する画素の強度に比例した発火率を持つポアソンスパイクに変換される。ポアソンスパイク列は興奮性ニューロンへの入力として網羅的に供給される。図の青の領域は、ある特定の興奮性ニューロンへの入力接続を示す。興奮性ニューロンは1対1の接続を介して抑制性ニューロンに接続される。図の赤色の部分は、1つの抑制性ニューロンから興奮性ニューロンへのすべての接続を示す。各抑制性ニューロンはすでに接続されたニューロンを除く全ての興奮性ニューロンと接続されている。興奮性ニューロンは、学習後に学習データセット中の数字クラスに対する最も高い応答に基づいて、クラスが割り当てられる。

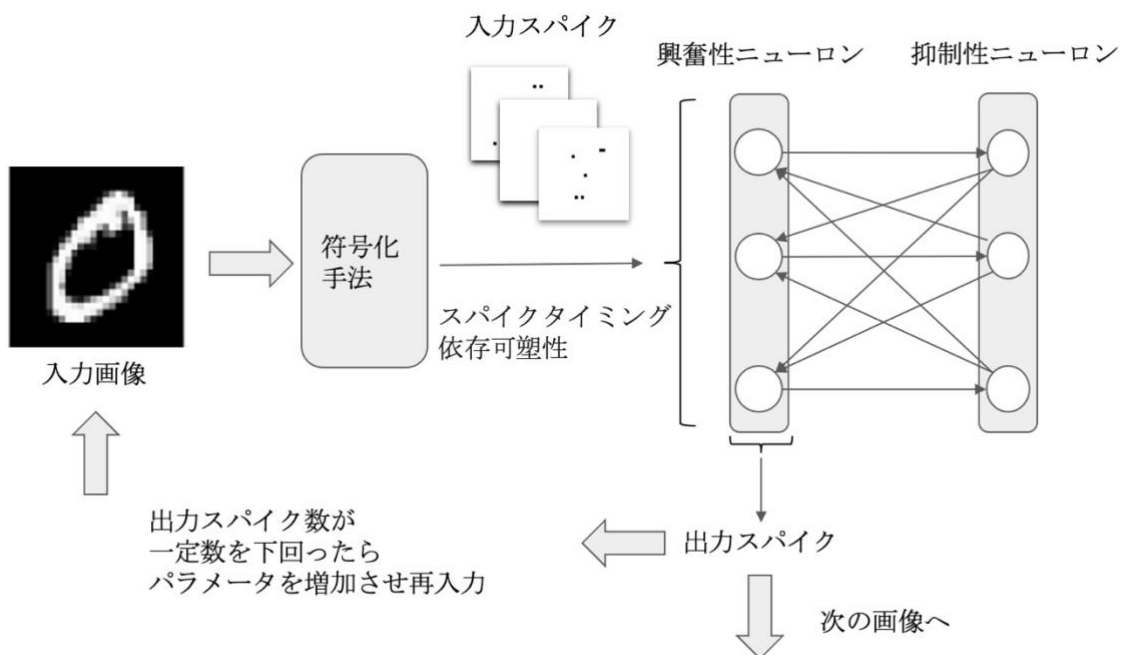


図4 ネットワークアーキテクチャと学習について

2-3 人の網膜機能と構造

人の視覚系は眼から始まり、眼の奥には網膜が存在する。網膜はカメラのフィルム以上の機能を持つ脳の一部である。厚さが 0.2mm 前後の薄い膜であり、光を最初に受容する器官である。目に映し出された画像を処理し、視神経から脳に伝達する機能をもつ。外界の像は眼球の角膜やレンズなどの光学系を介して網膜面上に投射される。投射された像は、網膜に配置された視細胞の細胞膜電位に変換される。電気信号に変化された外界像は、網膜神経回路網によって、高次の視覚認識に重要な情報が抽出され、網膜の出力細胞である神経節細胞の活動電位にコードされ網膜ニューロンの軸索は視神経の束にまとめられ、神経節細胞の軸索すなわち視神経を介して脳へ送られる。視神経はいくつかの脳部位に活動電位という形で視覚情報を送る。

網膜の神経回路は、大きく分類して視細胞、水平細胞、双極細胞、アマクリン細胞および神経節細胞の 5 種類の細胞から成る。視細胞から双極細胞を経て神経節細胞に至るまでが、視覚信号の流れの最も直接的な経路である。神経節細胞が光に反応して活動電位を発すると、活動電位は視神経を経て脳に伝えられる。それぞれのシナプス中継地点で水平細胞とアマクリン細胞の側方からのシナプス結合による修飾を受けている。

それぞれの細胞を説明する。視細胞は光刺激を神経活動に変換する役割をもつ。水平細胞や双極細胞とシナプスし、脊椎動物の視覚系で最初の神経回路を形成する。網膜には 2 種類の視細胞が存在する。桿体と錐体である。桿体は暗所視に用いられ、錐体は明所視に用いられ、錐体は色の識別能をもつ。視細胞において中心窩近傍では細胞が小さく、視細胞密度も高いが、中心窩から離れるに従って細胞は大きくなり密度も低下する。[12]

桿体と錐体による違いも存在する。錐体は網膜の中心窩に多くあり、網膜周辺部へ遠ざかるに従って錐体の割合が低くなる。一方中心窩の真ん中には桿体は存在せず、網膜周辺部へ遠ざかるに従って桿体の割合が錐体よりも高くなる。双極細胞は視細胞から情報を受け取り、神経節細胞へ出力を送り出す役割を持つ。双極細胞は複数の視細胞からの入力を集めて出力を神経節細胞に伝えるが、網膜中心部では視細胞と一対一対応している。この構造が網膜中心部での高い空間分解能に寄与すると考えられている。またアマクリン細胞ともシナプス結合をもつ。ヒトの網膜における双極細胞の数ははっきりと特定されていない。神経節細胞は双極細胞やアマクリン細胞を介して視細胞からの情報を受け取り、活動電位を視神経を介して脳に視覚情報として伝達する。ヒトの網膜には約 70 万～150 万個の網膜神経節細胞があり、約 12 種類存在するとされている。神経節細胞に関しても中心窩部が最も密度が高く配置されている。水平細胞は視細胞からの入力を受けると、神経突起を横に伸ばして周囲の双極細胞や視細胞に伝達する。水平方向に情報を伝達し、隣接する視細胞からの情報を統合する役割があるとされている。アマクリン細胞は双極細胞からの入力を受け、周囲の神経節細胞、双極細胞、別のアマクリン細胞に信号を横方向に伝達する。水平方向に情報を伝達することで、隣り合う視細胞からの情報を統合する役割があるとされる。[13]

これら 5 種類の神経細胞が整然とした層構造を保ちながら配置されている。

網膜で実行されている光受容、情報抽出、情報伝送機能は工学的に見ても巧妙かつ効率的である。しかし外界像が神経節細胞の活動電位の中にもどのように符号化されているかは未知である。視覚生理学・心理物理学の重要な課題であるとされ、工学的にも全く新しい知能視覚センサー人工視覚の開発として大きな意義を持つ。またスパイクニューラルネットワークにおいても網膜と神経節細胞の振る舞いをモデル化し取り込むことは、高性能で効率的な SNN を開発するための有望な手段とされている。[4]

3-2 提案手法

3-2-1 ガンマ分布を使用した符号化手法

現在 SNN における画像からスパイク列への変換は、ポアソン過程によって生成されるポアソン分布に従うスパイクを出力することで行われているが、これは脳の自発的発火が不規則であり、ポアソン分布はその不規則性を適切に記述することができる、すなわちスパイクの間隔がポアソン分布に従うとの仮定が存在するためだ。

一方、脳の自発的発火には特別な意味がないと考えられてきたが、近年は異なるのではないかと考えられている [7]。Maimon らは脳の大脳皮質の発火がガンマ分布に従っているのではないかと主張している [14]。

そこで本研究ではガンマ分布を用いた符号化手法を考案する。

ガンマ分布の確率密度関数は以下の通りに表すことができる。 [8]

$$f(x) = \frac{\beta}{\Gamma(\alpha)} (\beta x)^{\alpha-1} e^{-\beta x} \quad (17)$$

x: 確率変数、 Γ : ガンマ関数、 α : 形状母数、 β : 尺度母数

従来使用されるポアソン分布による符号化手法と同様、入力するデータセット画像の各画素値を取り出し、倍率(レート)を掛け合わせパラメータとし、ガンマ分布に従うスパイクを発生させる。

3-2-2 エッジフィルター

人の網膜は物体の境界を抽出する働きをもつ。そこで使用する画像を加工し、輪郭を抽出した画像を作成した後、入力とすることでスパイクを抽出する手法を提案する。本研究では pillow とよばれる画像処理ライブラリの画像フィルターを使用する。画像の輪郭を検出するフィルターを適用し画像加工を行う。



図5 画像の輪郭を抽出するためのフィルターを適用した後の画像
(輪郭抽出は画像から物体の境界線を検出する処理のこと)

3-2-3 抽出範囲変更 1

人の網膜では視細胞から視神経に向かうまでに横方向の情報を取り入れ統合され脳へ視覚情報が送られている。しかし従来の符号化手法では画素値ごとにスパイクが出力されているので画素値同士横方向の情報は統合されていない。

また視細胞には大きいサイズの視細胞と小さいサイズの視細胞が交互に配置されている構造が確認できる。そこで大きい範囲(9つの画素値)で画素値を統合し、小さい範囲(1つの画素値)からそのまま画素値を交互に取得することで、視細胞の構造を倣いつつ横方向の情報を取り入れ、スパイクを出力する手法を提案する。

まず画素の最も外端の部分においては1画素飛ばしで画素値を抽出し、スパイクを抽出する。その後、他の画像内部分においては 3×3 の9つの画像値を合計しスパイクを抽出する。この処理に関しても1画素飛ばしで行い、飛ばされた部分に関してはそのまま画素値を抽出し、スパイクを出力する。スパイクを9つの画素値を合計し出力し、横方向の情報を統合する。(図)は本手法を説明するものである。

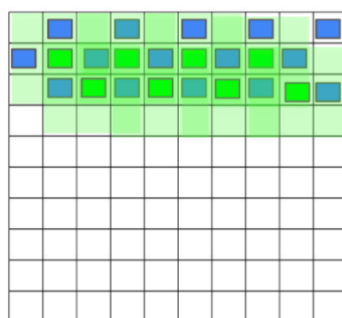


図 6 抽出範囲変更 1

青四角はそのまま画素値を取得することを、
緑四角は 3×3 の範囲から画素値を抽出することを表す。

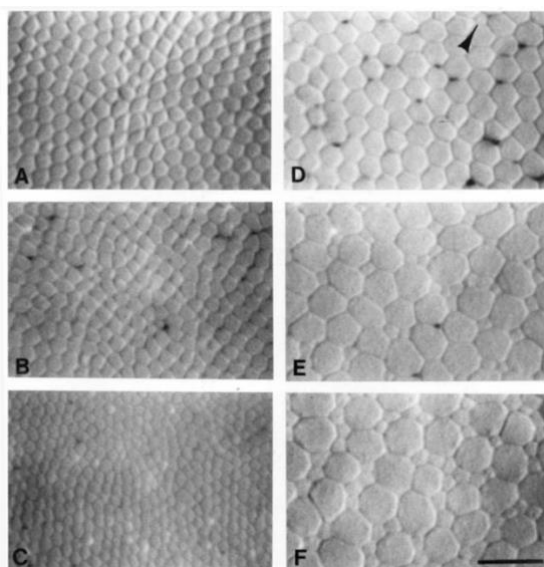


図 7 視細胞の表面 [16]より出典

3-2-4 抽出範囲変更 2

網膜では中心に小さな視細胞が、中心から離れるにつれ大きな視細胞が配置されているのを確認できる。また中心に近い視細胞ほど少ない数の神経節細胞に結合し、中心から離れるにつれ多い数の神経節細胞と結合する構造が確認できる。

そこで画像の外端に近づくにつれ大きい範囲から画素値を取得し、中心ではそのまま画素値を取得することで視細胞の構造を倣いつつ横方向の情報を取り入れ、スパイクを出力する手法を提案する。

(図) は本手法を説明するものである。

外端側は 3×3 の範囲の画素値の平均を取得する。中間は 2×2 の範囲の画素値の平均を取得する。中心はそのまま画素値を使用する。

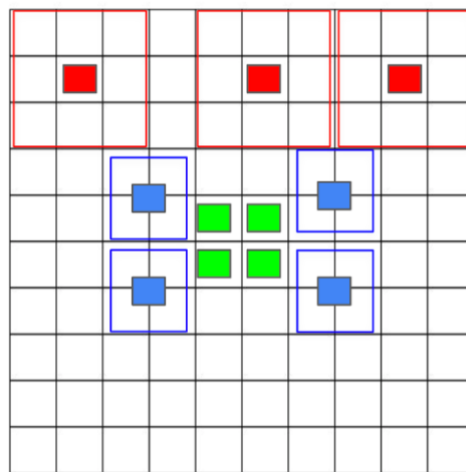


図 8 抽出範囲変更 2

赤四角は 3×3 の範囲の画素値の平均を、
青四角は 2×2 の範囲の画素値の平均を、
緑四角はそのままの画素値を取得する。

4 章 実験の概要

4-1 実験方法

SNN の基本ネットワークとして Diehl&Cook モデルを使用し、符号化手法を変更することで、符号化手法の精度や効率に対する影響を調べる。従来 Diehl&Cook モデルの符号化手法は、符号化されたスパイクがポアソン分布に従うポアソン過程を使用した手法である。3 章で提案した入力画像への加工処理手法だけでは画像からスパイクの変換は行われなため、それぞれの画像加工処理をした後、それぞれの分布による変換を行いスパイクの抽出をする。従来手法と同様、入力するデータセット画像の各画素値を取り出し、倍率(レート)を掛け合わせパラメータとし、分布に従うスパイクを発生させる。本研究では倍率(レート)を調節することで、異なる値のパラメータを作成し、スパイク数と精度の関係を調べる。

また Diehl&Cook モデルにおいては出力のスパイクが閾値を下回る場合、発火率の倍率を一定数増加させ、再度同じ画像より新たなスパイクを作成し入力・学習を行う機構が存在する。分布単体の影響を調べるためにも、再入力機構が実行されない場合も調査する。なお再入力機構は著者実装においては、興奮性ニューロンのスパイク数が 5 つ以下であったら倍率を +1 とし再度入力を行っているが、本研究においてはスパイク数が 5 つ以下であれば倍率を +16 している。また再入力の画像ごとの上限回数を 10 回に定めた。そのほかのパラメータは従来手法と同様である。そのため検証する手法としては画素値からスパイクへの変換手法 3 つと処理を行わない場合を含む入力前処理 4 つを組み合わせ 12 種類存在することになる。

Diehl&Cook モデルは Python と BRIAN シミュレータ [17] によって作成されているが、本研究では Python とそのライブラリである Pytorch [18] を使用し作成したネットワークを使用した。

使用データセットは MNIST データセットを用いる []。MNIST データセットは 0-9 までの数字の画像である。サイズは 28×28 である。

本研究において訓練データ (train) は 10000 枚, 学習 epoch は 10 回とした。この epoch 回数は事前に行った実験より設定した。また検証データ (test) に関しては訓練データの 10% 程度から 20% 程度の量が使われることが多いことから 1000 枚に設定した。手法ごとに異なるスパイク数を出力するための倍率は 20・24・28・32・36・40・44・48 とした。

画像からスパイクへの変換手法			
ポアソンスパイク(再入力機構あり)	ポアソンスパイク(再入力機構なし)	ガンマスパイク	

入力前処理の有無と種類			
入力前処理なし	エッジフィルタ	抽出範囲変更 1	抽出範囲変更 2

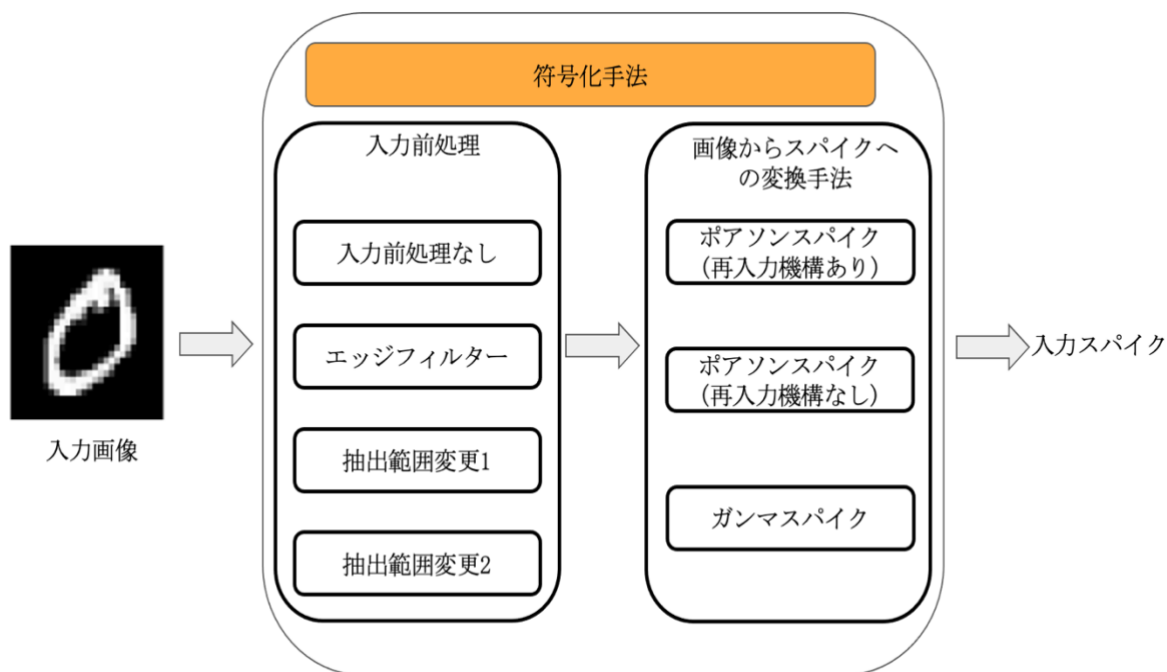


図9 本実験の符号化手法の組み合わせ

4-2 評価指標

本研究の評価指標としてスパイク数・検証データの正解率・1スパイクあたりの精度・スパイクの間隔・再入力回数を測定した。

スパイク数は入力やそれぞれのニューロンで発生したスパイク数である。符号化手法により作成された入力スパイク、興奮性ニューロンから出力された興奮性スパイク、抑制性ニューロンから作成された抑制性スパイクの数を取得し、訓練全体でのスパイク数を算出した。

検証データの正解率は検証に使用したデータのどれくらい正解したかである。

1スパイクあたりの精度は、検証データの正解率を訓練で使用したネットワーク内での発火数(スパイクの数)で割り算出する。その精度がどれくらいのスパイクを使用し学習され達成されたかを把握し、効率さを測定するためである。

またニューロンごとに発生したスパイクの間隔を測定する。入力画像の各画素や対応する各ニューロンにおいて2つ以上のスパイクが発生した場合、そのスパイク同士の間隔を取得。画像ごとの平均を算出、さらに訓練全体の間隔を算出して指標とした。この指標は符号化手法により作成された入力スパイク、興奮性ニューロンから出力された興奮性スパイク、抑制性ニューロンから作成された抑制性スパイクに適用される。その後それぞれのスパイク間隔より手法全体のスパイク間隔を算出した。

再入力の回数は再入力機構が存在する手法の場合のみ測定され、全体で再入力為何回行われたのかを表す。

4-3 結果

入力前処理なし×ポアソンスパイク (再入力機構あり)の結果

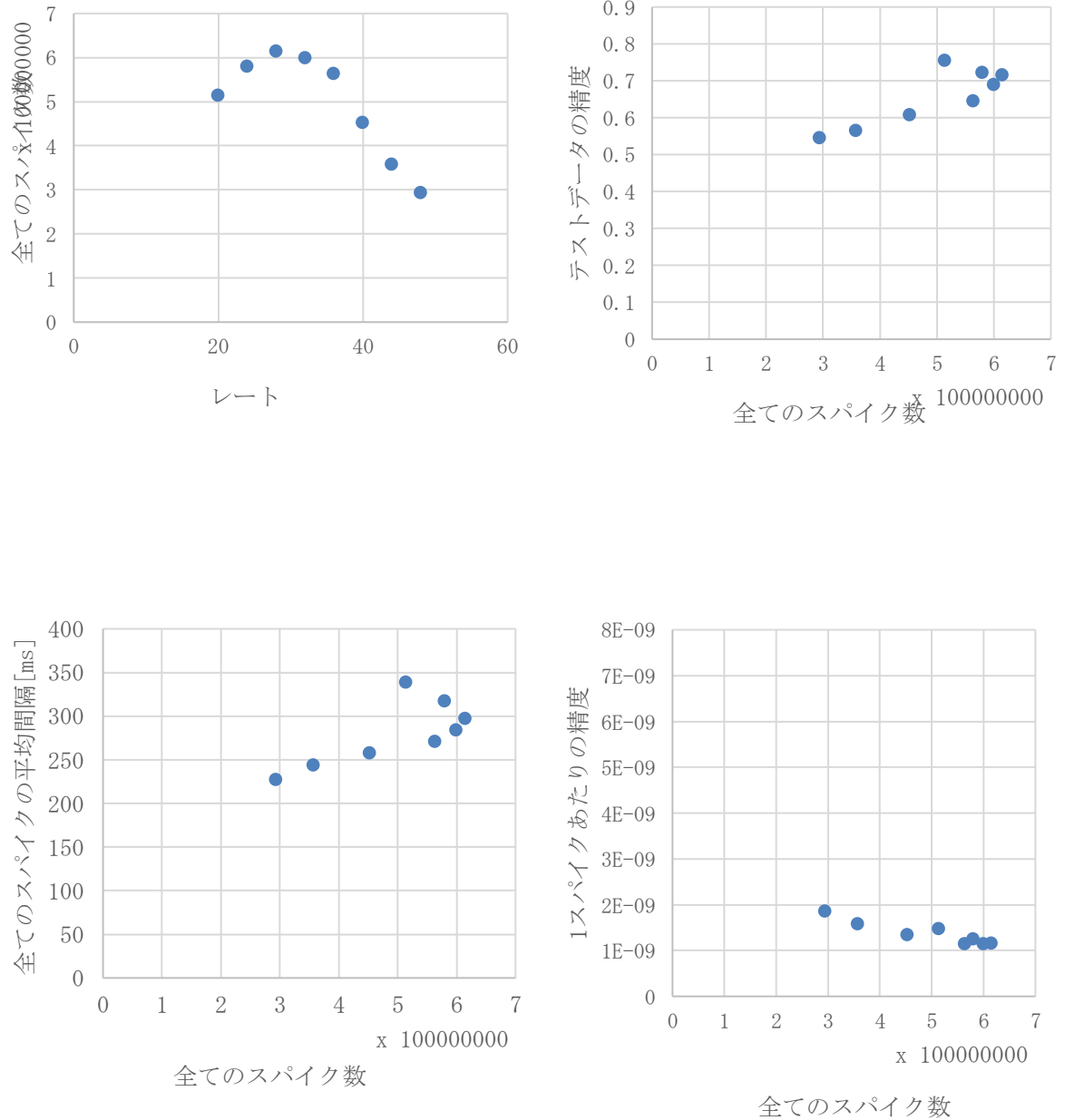
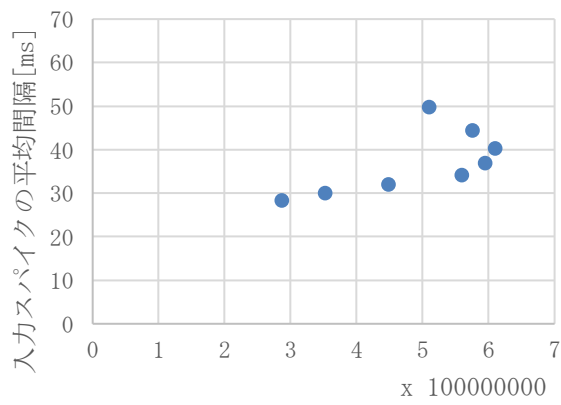
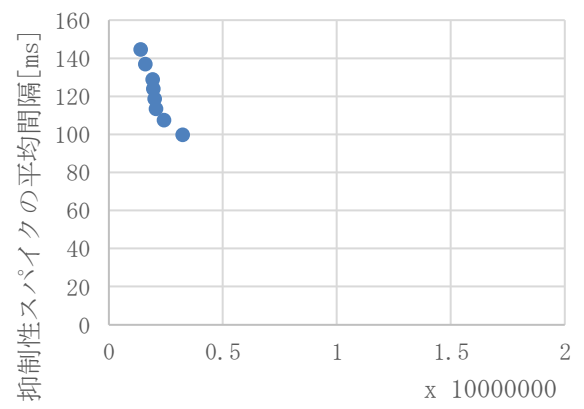


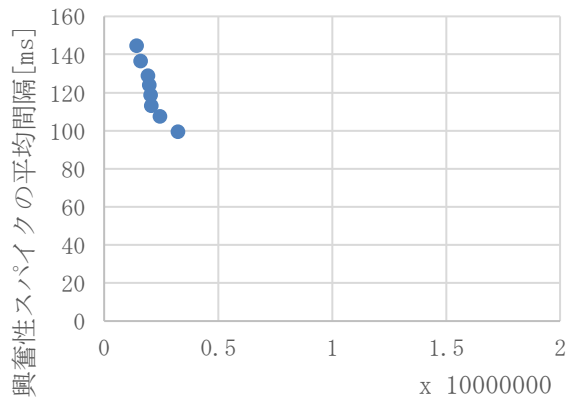
図10 入力前処理なし×ポアソンスパイク (再入力機構あり)の結果 1



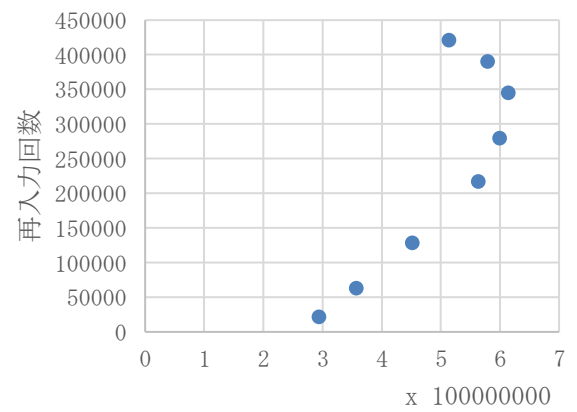
入力スパイク数



抑制性スパイク数



興奮性スパイク数



全てのスパイク数

図 11 入力前処理なし×ポアソンスパイク(再入力機構あり)の結果 2

エッジフィルター×ポアソンスパイク (再入力機構あり)の結果

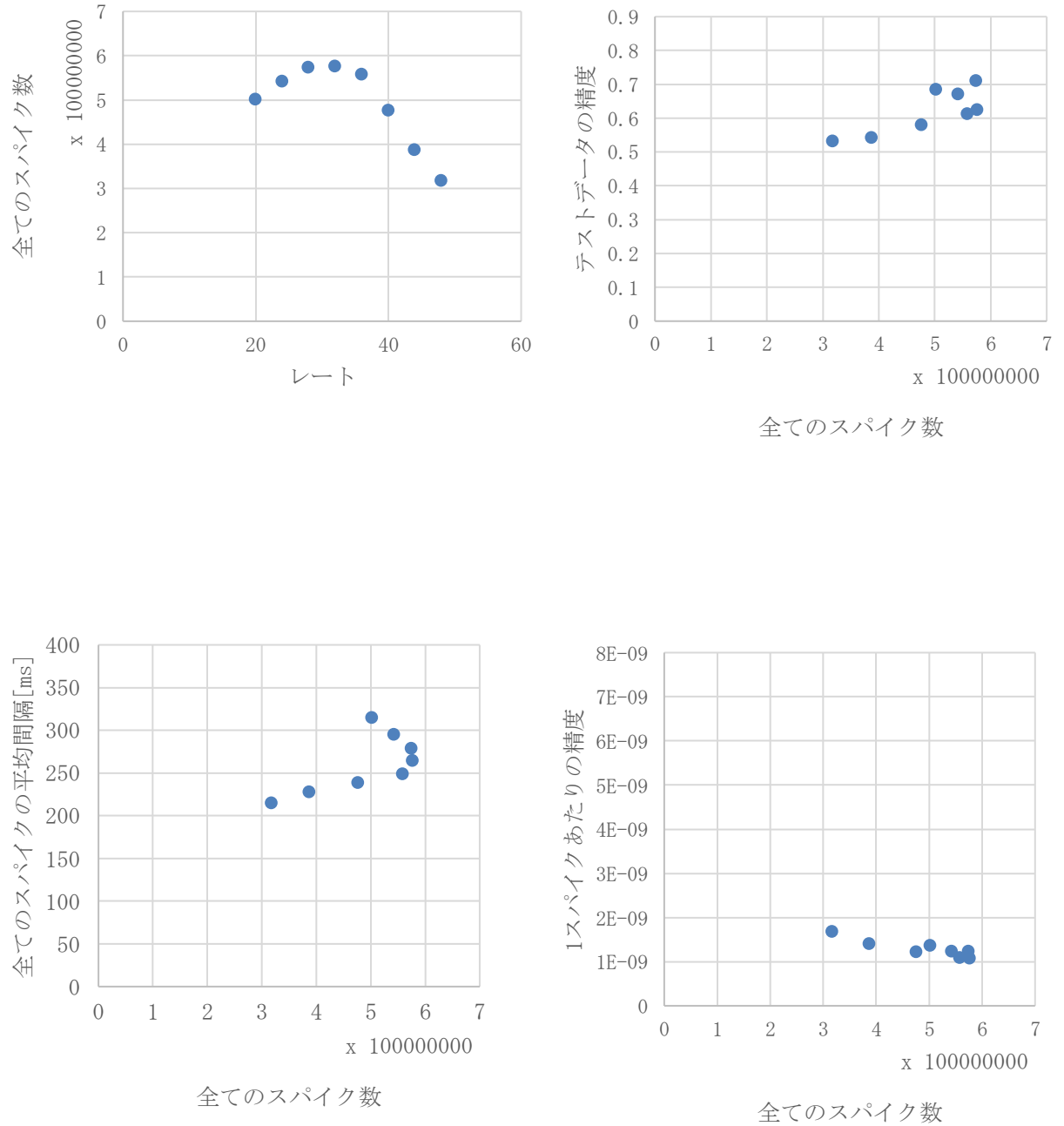
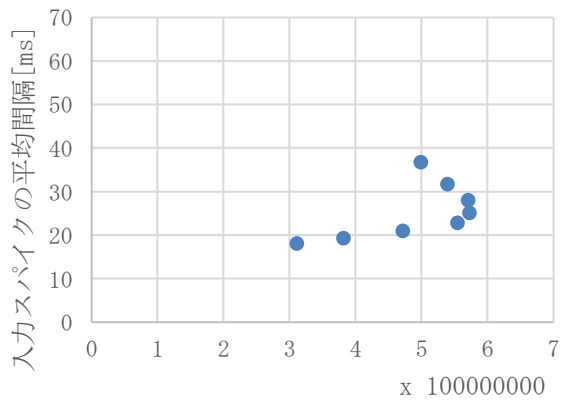
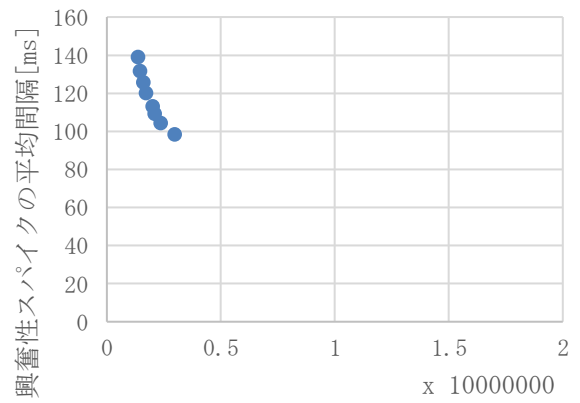


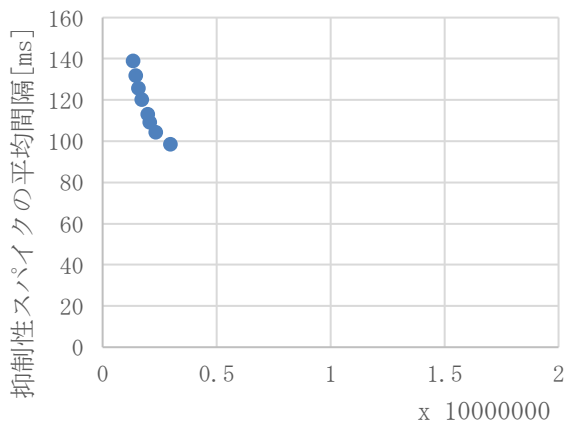
図 12 エッジフィルター×ポアソンスパイク (再入力機構あり)の結果 1



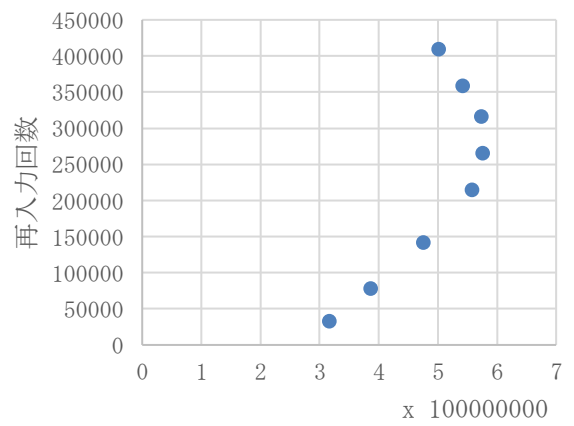
入力スパイク数



興奮性スパイク数



抑制性スパイク数



全てのスパイク数

図 13 エッジフィルター×ポアソンスパイク (再入力機構あり)の結果 2

抽出範囲変更 1×ポアソンスパイク (再入力機構あり) の結果

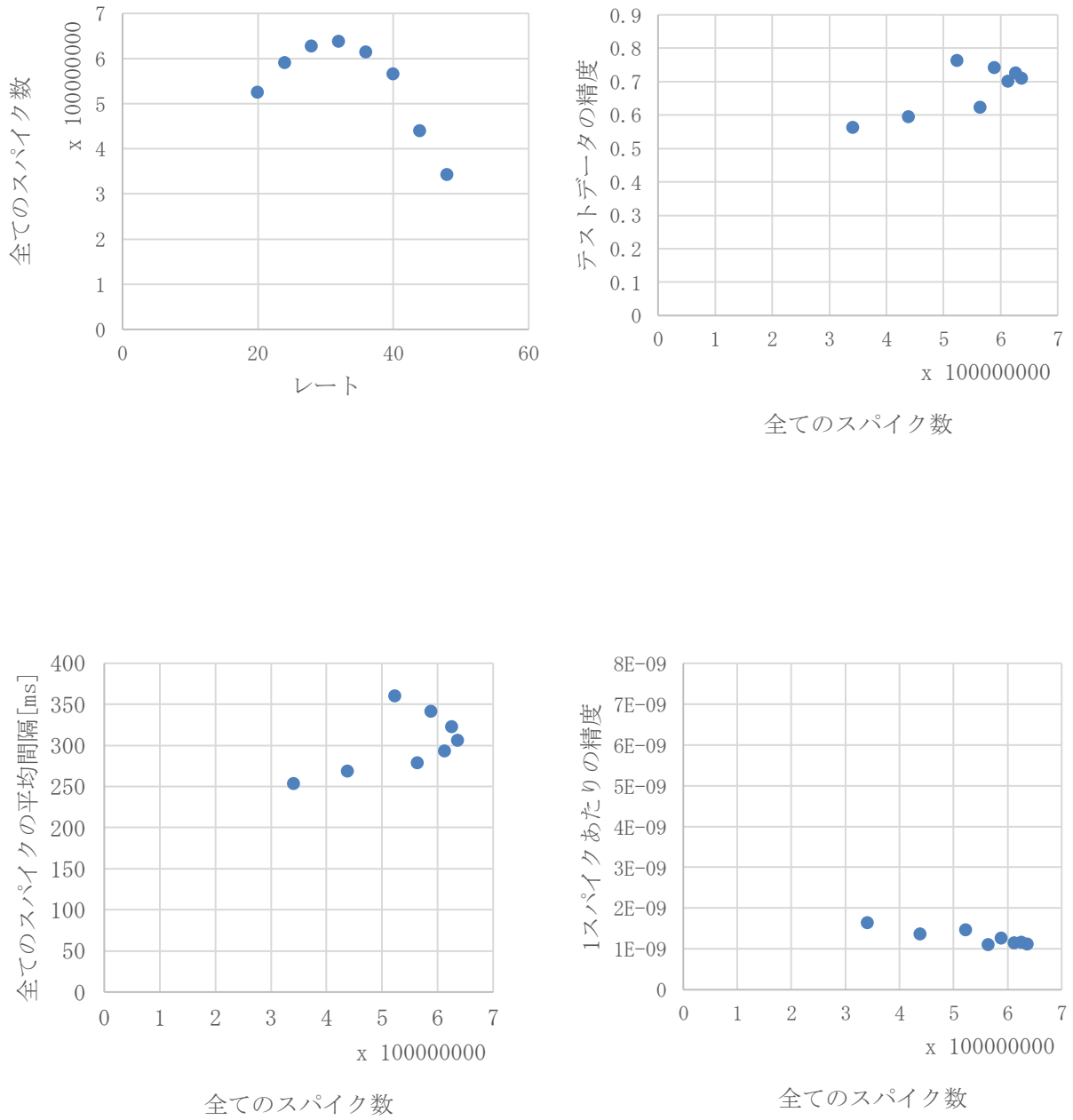
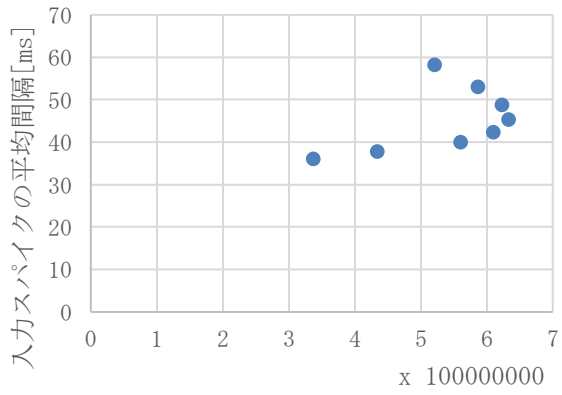
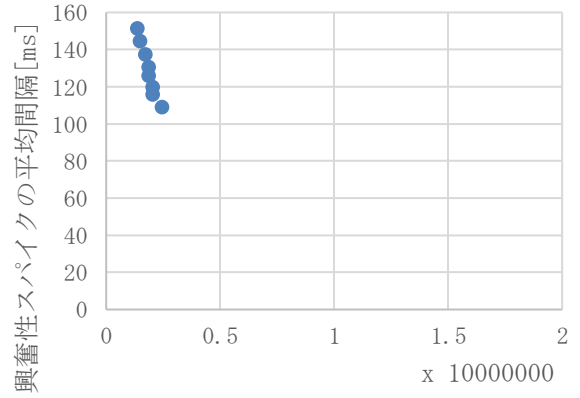


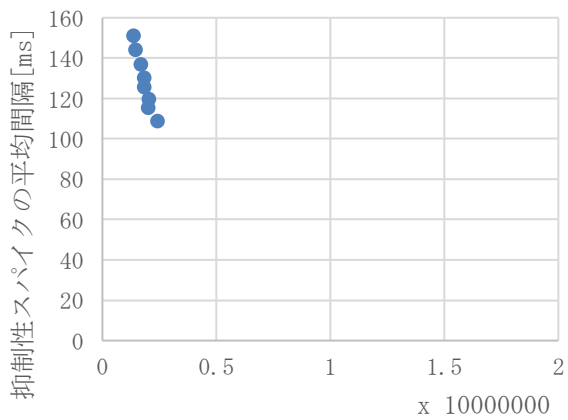
図 14 抽出範囲変更 1×ポアソンスパイク (再入力機構あり) の結果 1



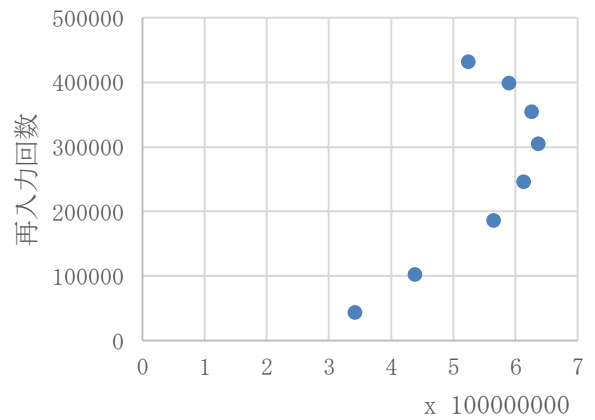
入力スパイク数



興奮性スパイク数



抑制性スパイク数



全てのスパイク数

図 15 抽出範囲変更 1×ポアソンスパイク (再入力機構あり)の結果 2

抽出範囲変更 2×ポアソンスパイク (再入力機構あり) の結果

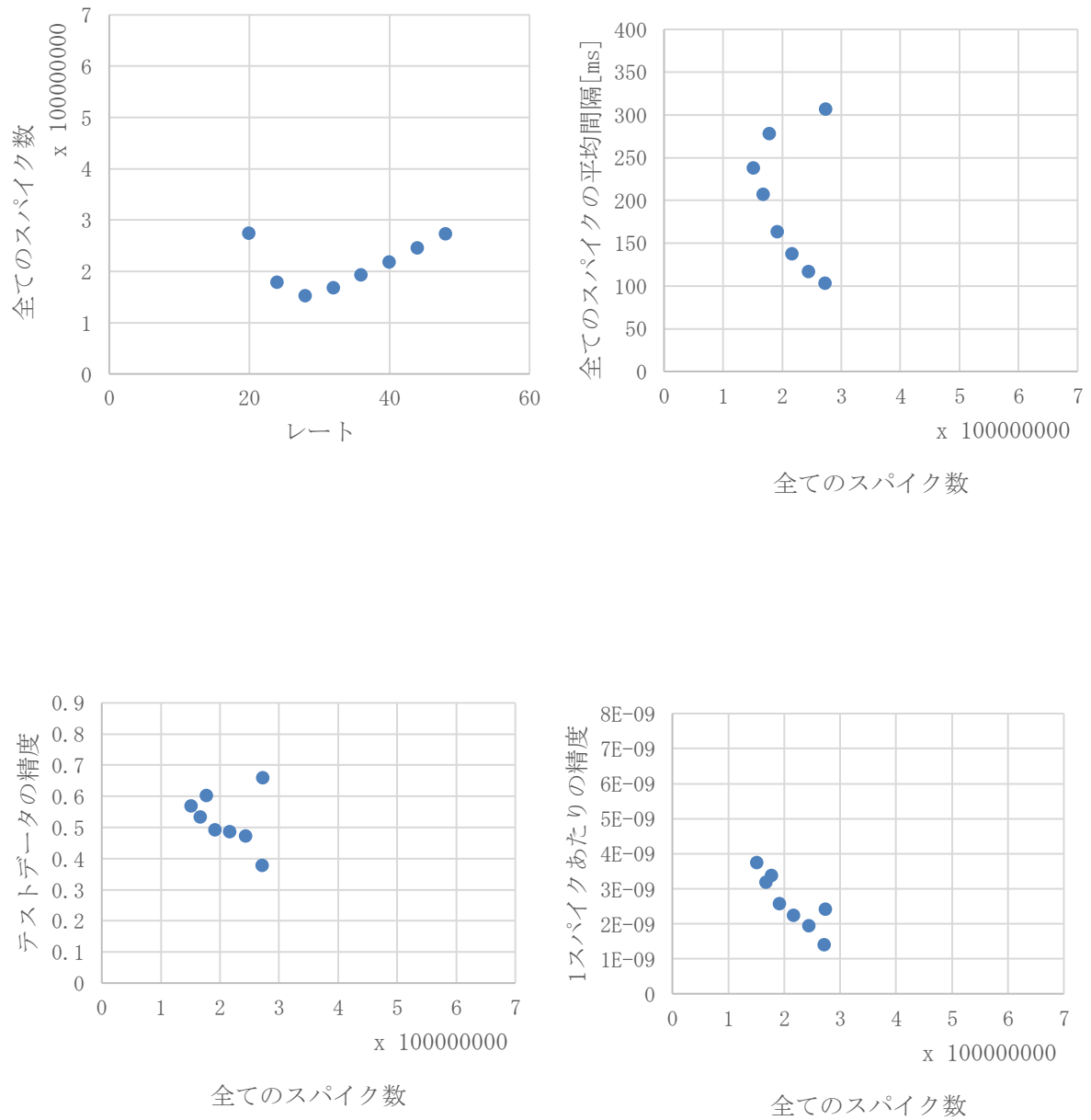
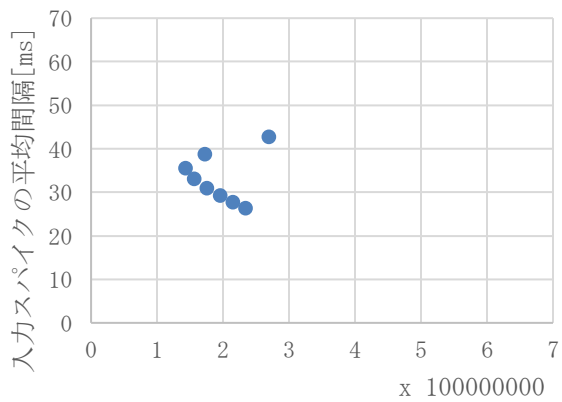
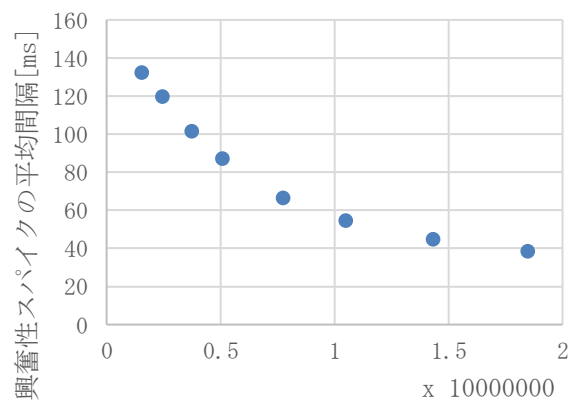


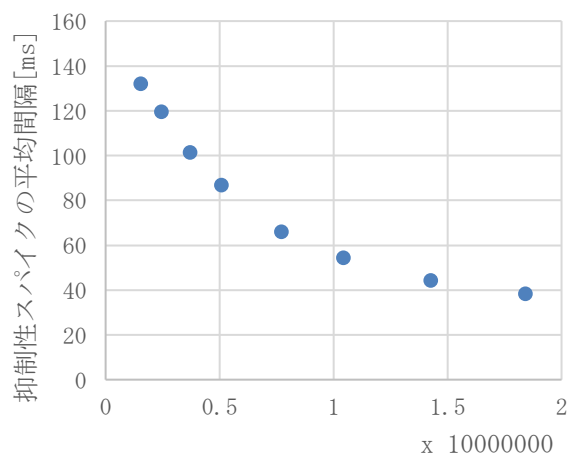
図 16 抽出範囲変更 2×ポアソンスパイク (再入力機構あり) の結果 1



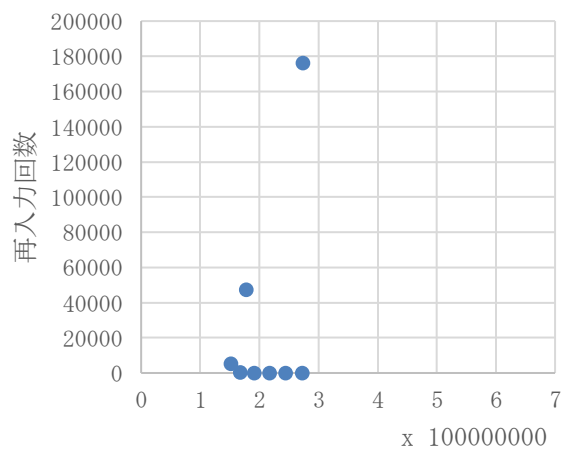
入力スパイク数



興奮性スパイク数



抑制性スパイク数



全てのスパイク数

図 17 抽出範囲変更 2×ポアソンスパイク (再入力機構あり) の結果 2

入力前処理なし×ポアソンスパイク (再入力機構なし)の結果

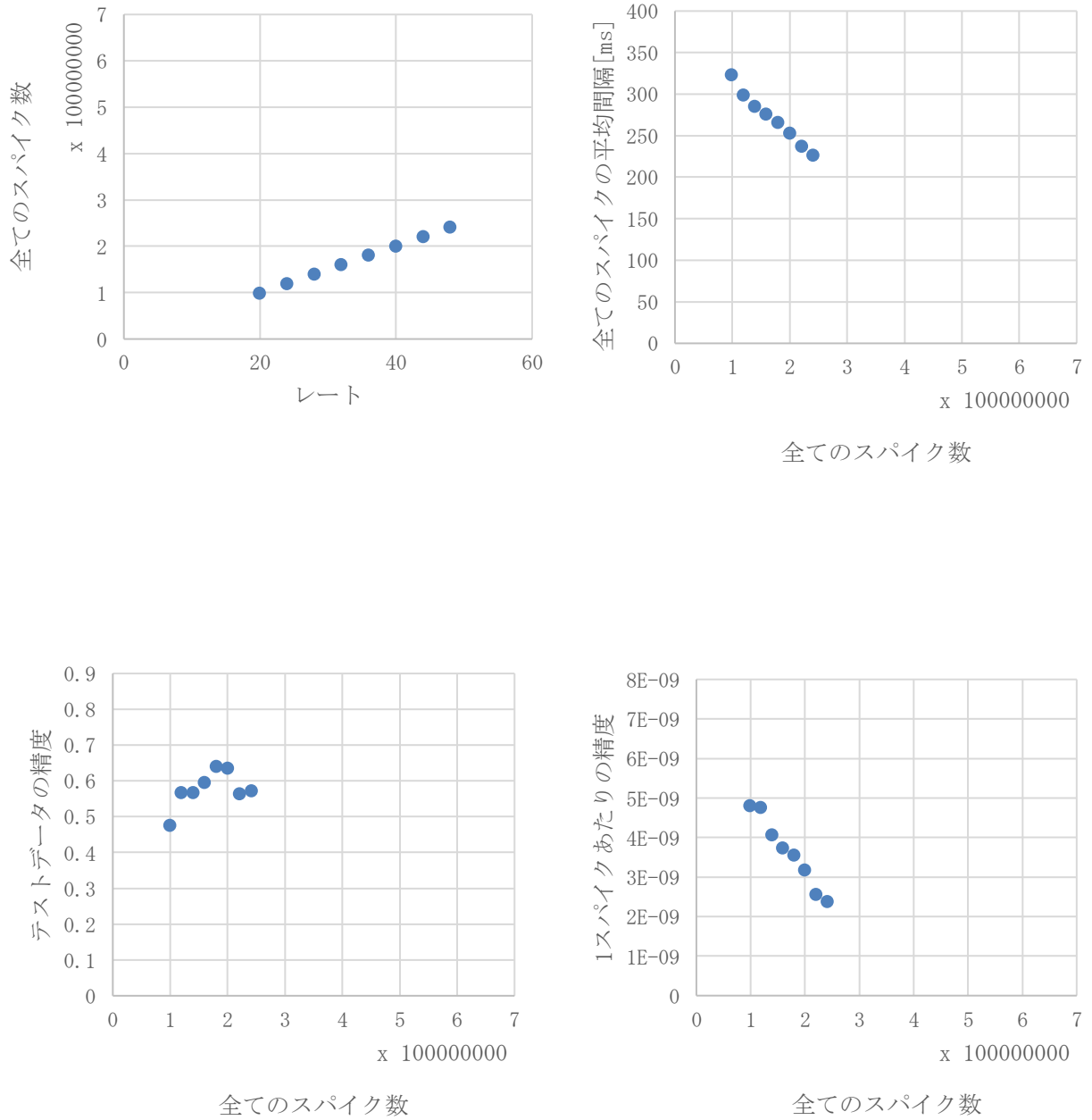


図 18 入力前処理なし×ポアソンスパイク (再入力機構なし)の結果 1

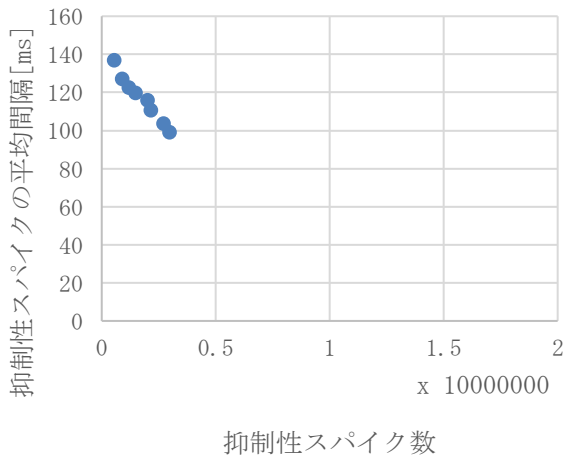
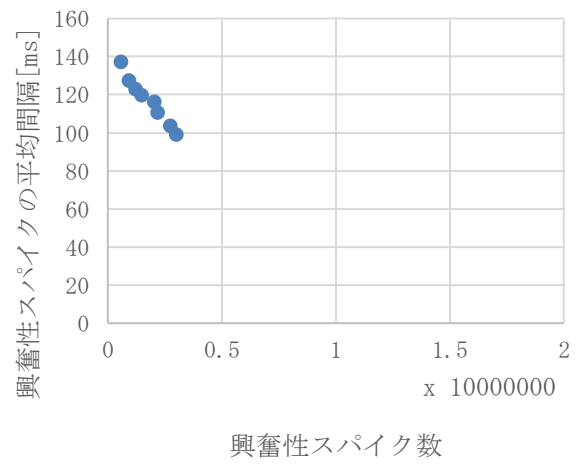
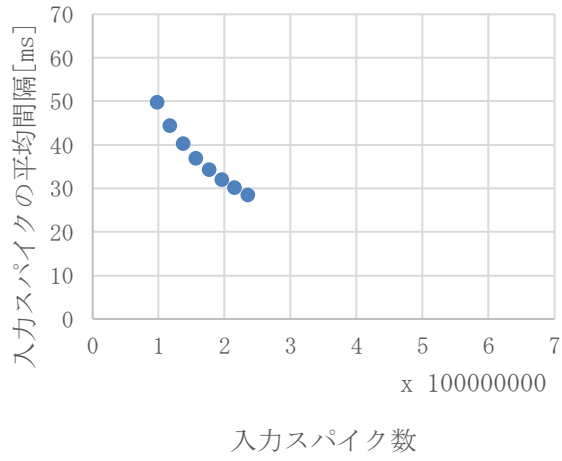


図 19 入力前処理なし×ポアソンスパイク (再入力機構なし)の結果 2

エッジフィルター×ポアソンスパイク (再入力機構なし)の結果

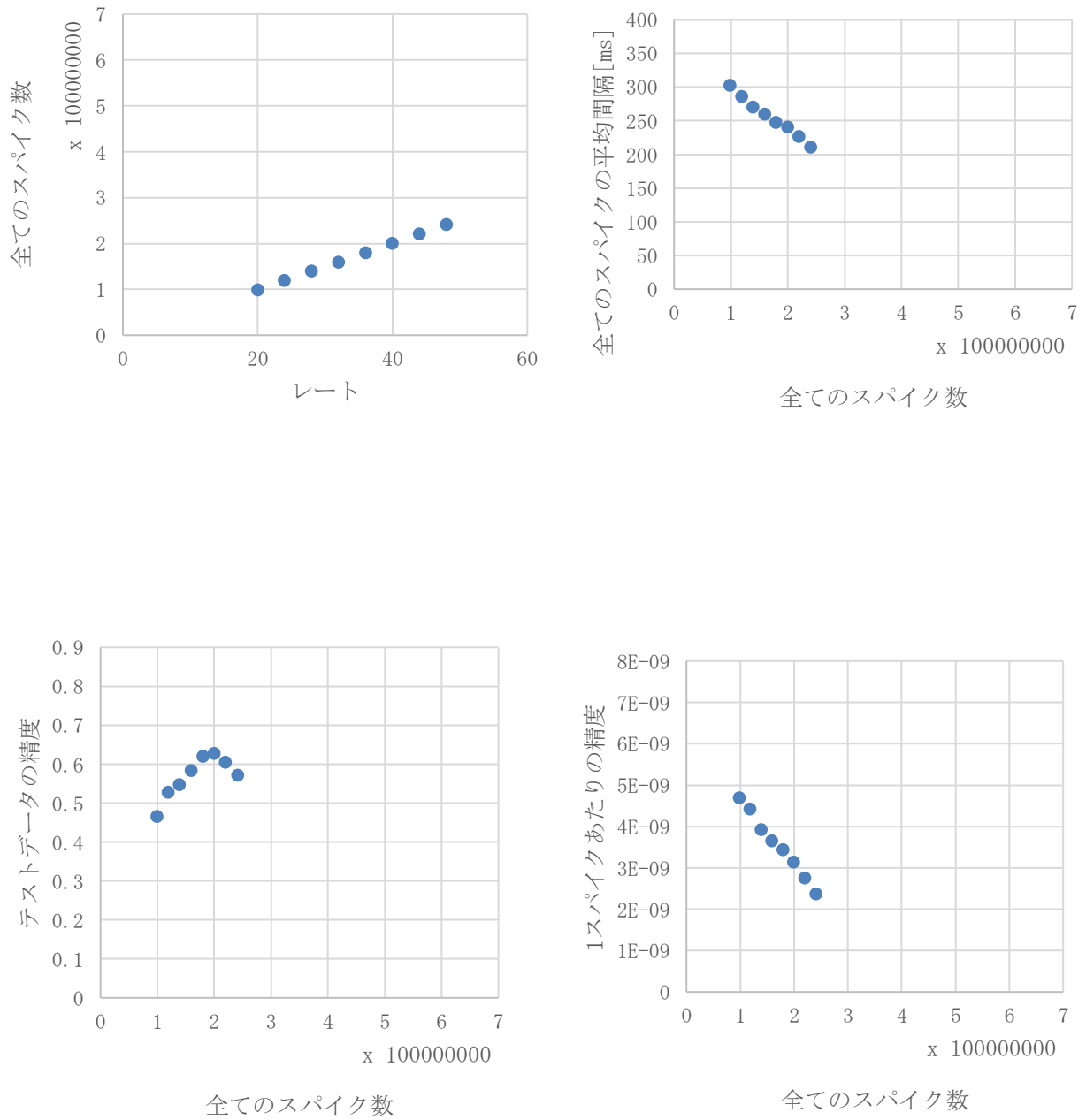
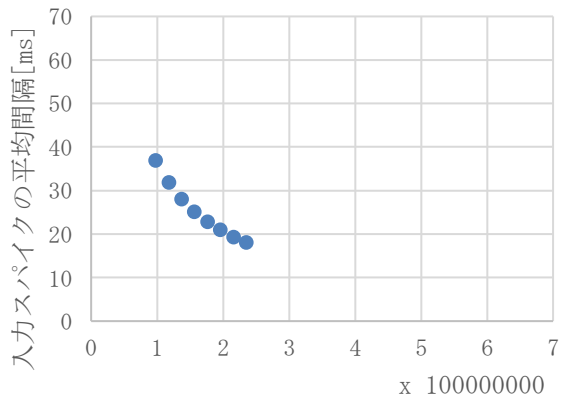
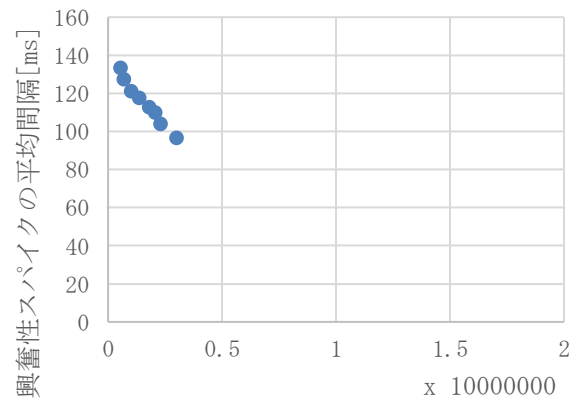


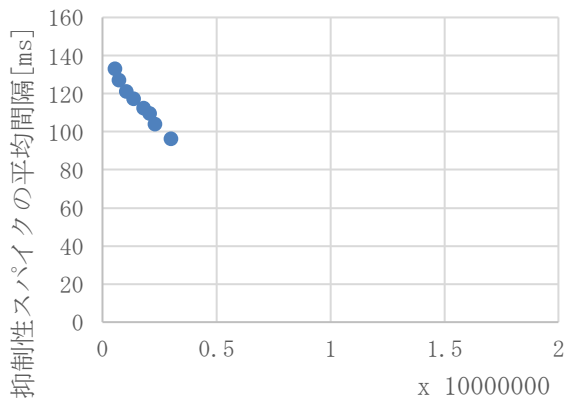
図 20 エッジフィルター×ポアソンスパイク (再入力機構なし)の結果 1



入力スパイク数



興奮性スパイク数



抑制性スパイク数

図 21 エッジフィルタ×ポアソンスパイク (再入力機構なし) の結果 2

抽出範囲変更 1×ポアソンスパイク(再入力機構なし)の結果

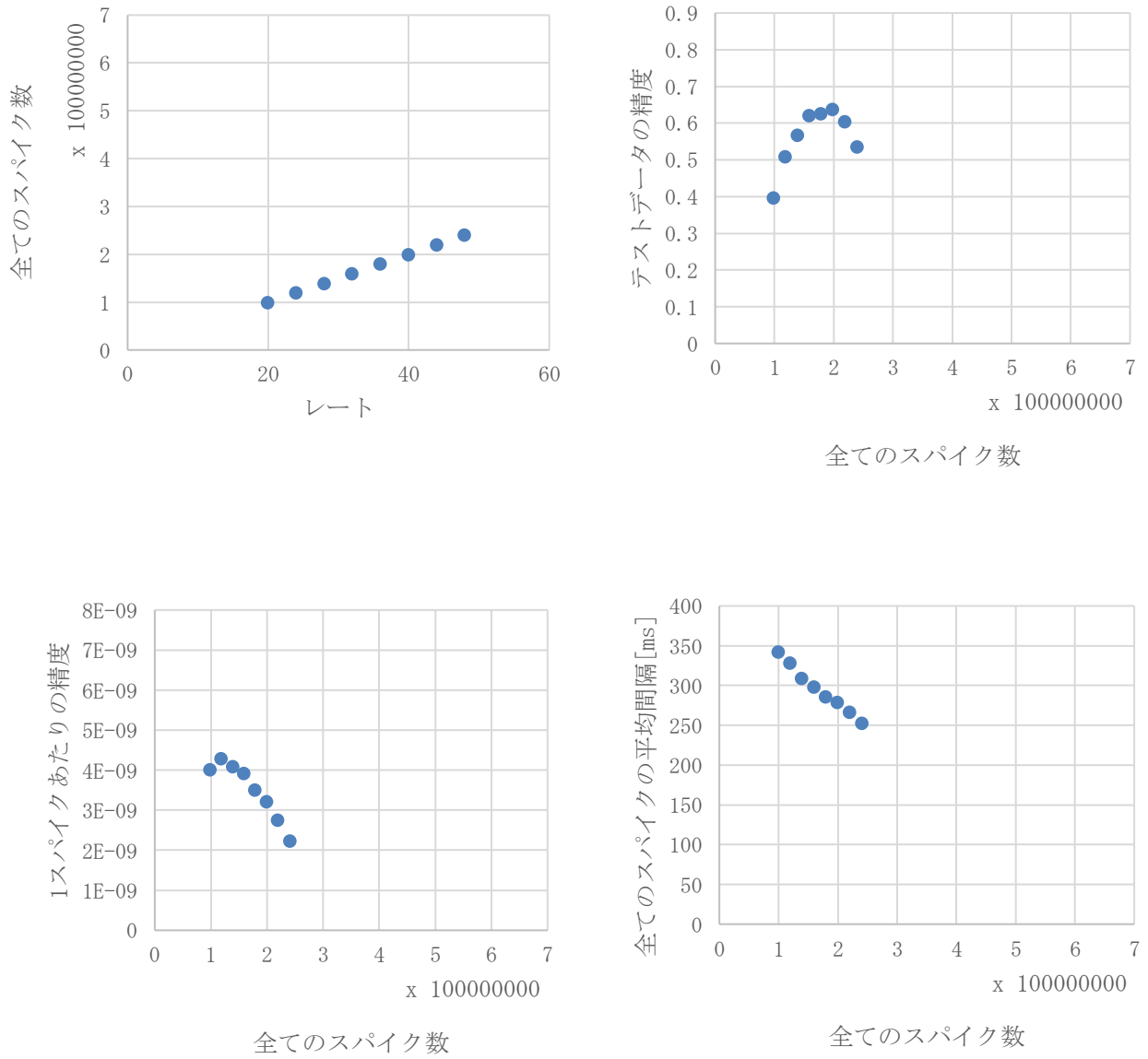
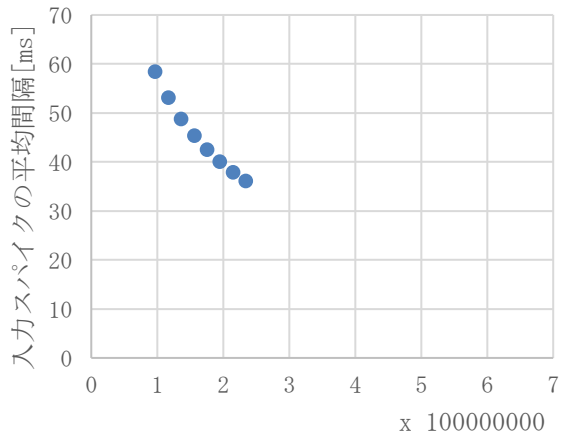
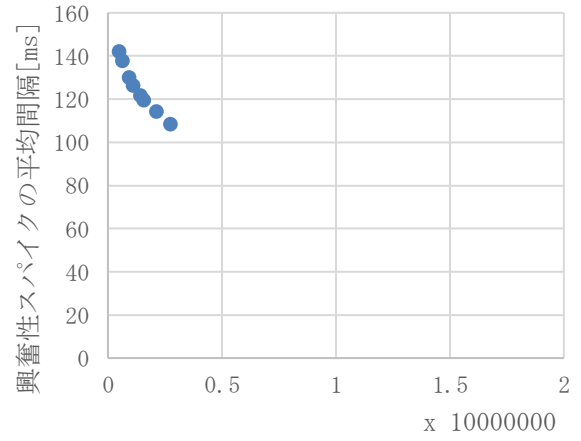


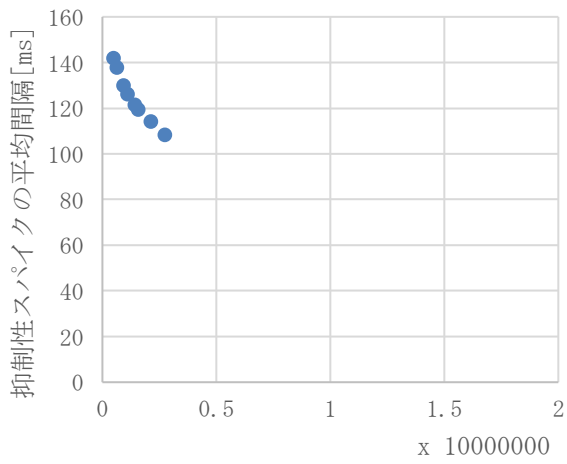
図 22 抽出範囲変更 1×ポアソンスパイク(再入力機構なし)の結果 1



入力スパイク数



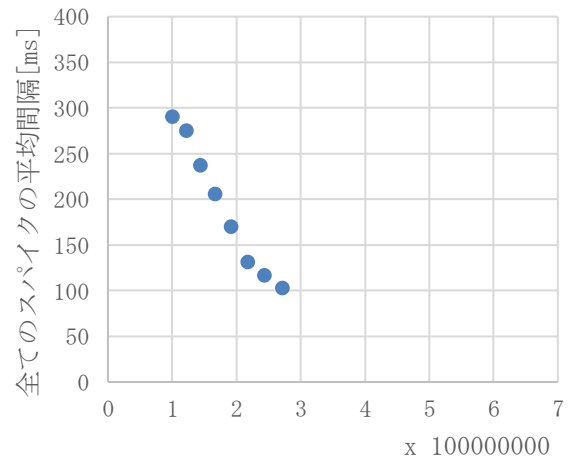
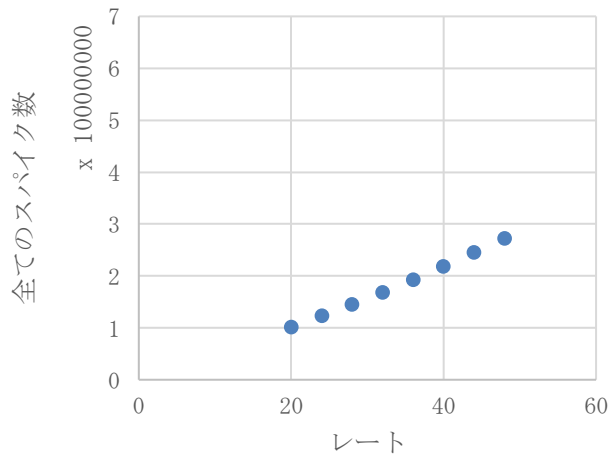
興奮性スパイク数



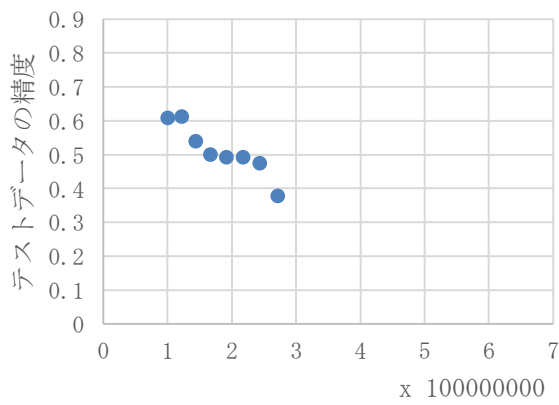
抑制性スパイク数

図 23 抽出範囲変更 1×ポアソンスパイク (再入力機構なし)の結果 2

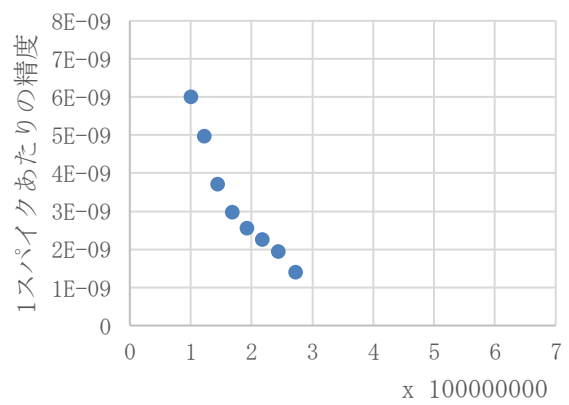
抽出範囲変更 2×ポアソンスパイク (再入力機構なし) の結果



全てのスパイク数

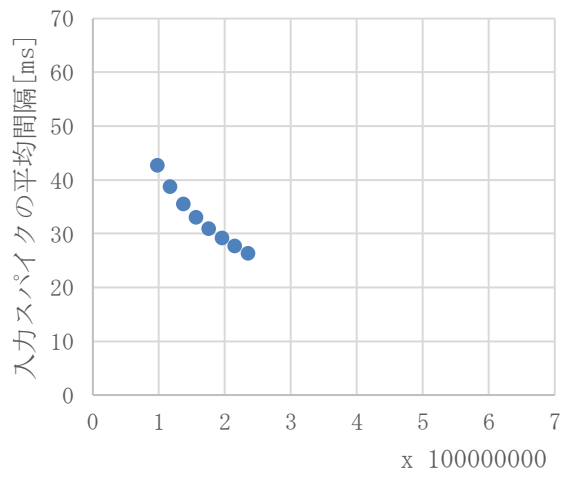


全てのスパイク数

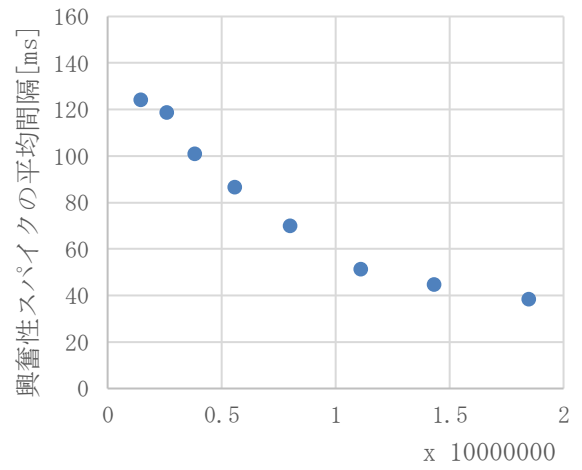


全てのスパイク数

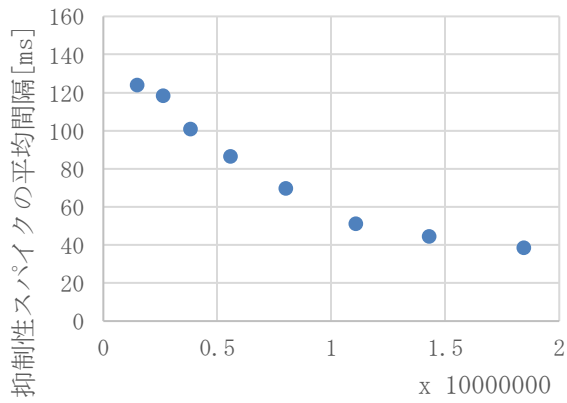
図 24 抽出範囲変更 2×ポアソンスパイク (再入力機構なし) の結果 1



入力スパイク数



興奮性スパイク数



抑制性スパイク数

図 25 抽出範囲変更 2×ポアソンスパイク (再入力機構なし) の結果 2

入力前処理なし×ガンマスパイクの結果

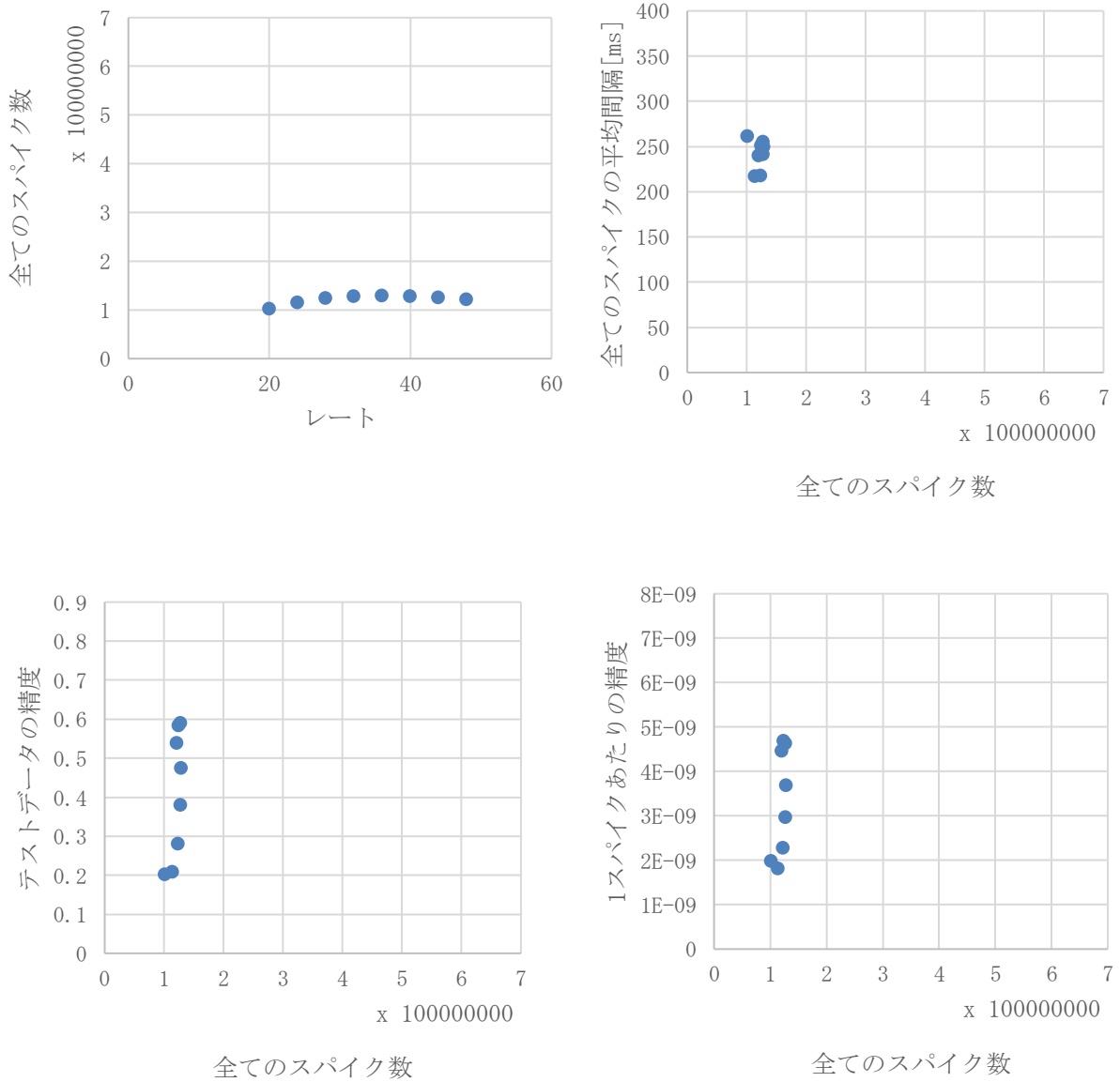
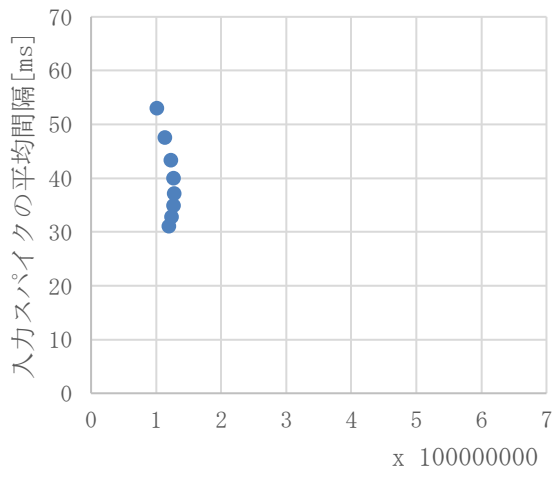
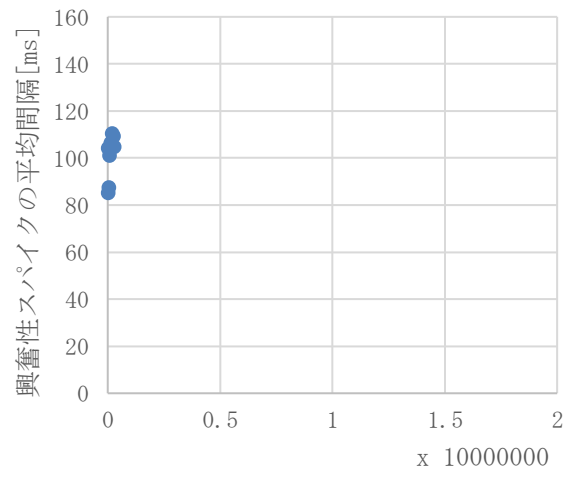


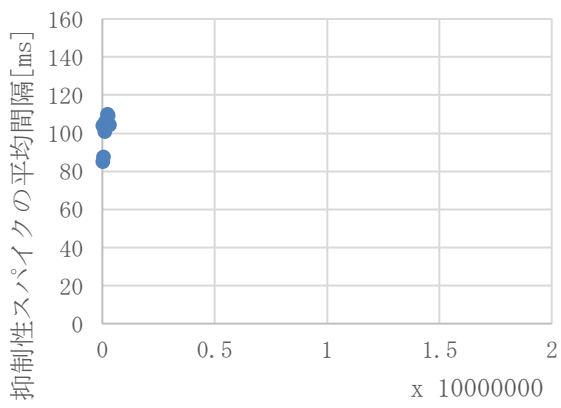
図 26 入力前処理なし×ガンマスパイクの結果 1



入力スパイク数



興奮性スパイク数



抑制性スパイク数

図 27 入力前処理なし×ガンマスパイクの結果 1

エッジフィルター×ガンマスパイクの結果

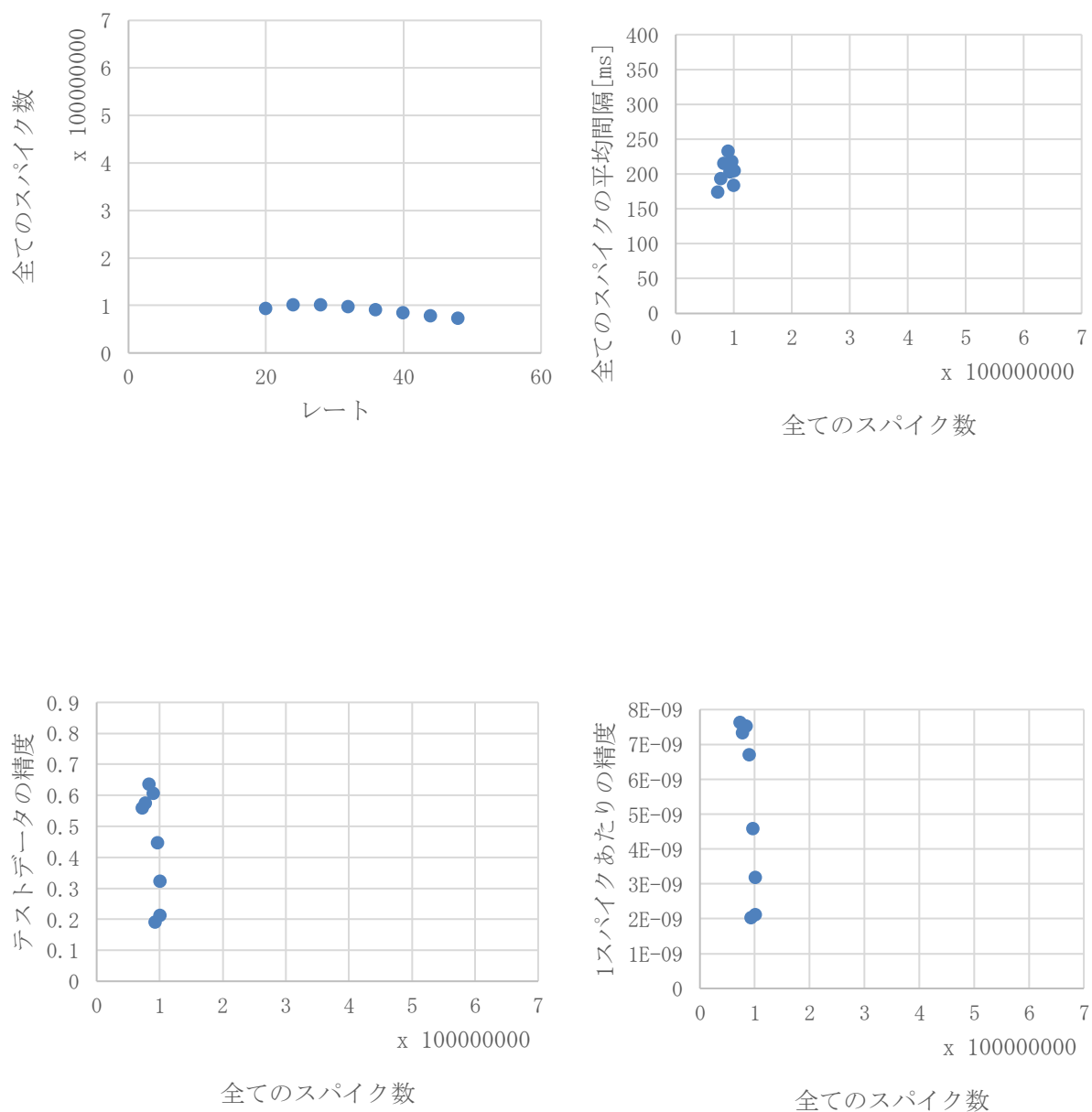
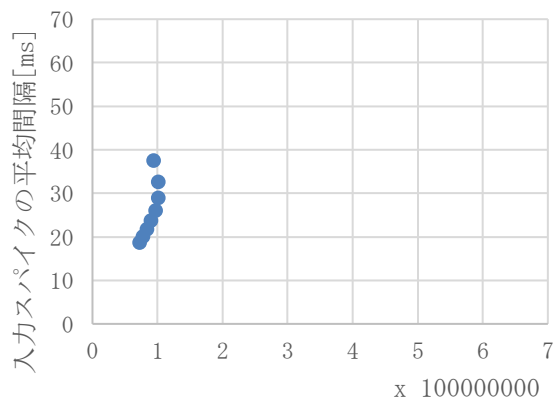
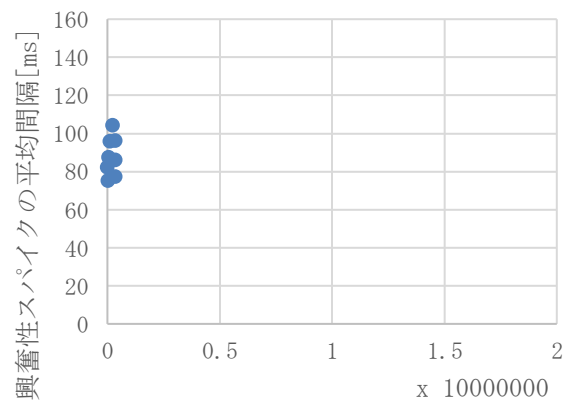


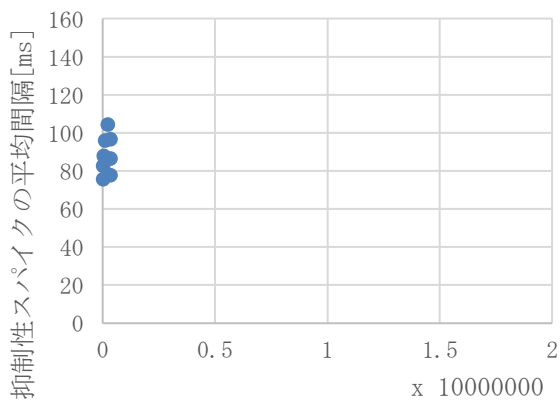
図 28 エッジフィルター×ガンマスパイクの結果 1



入力スパイク数



興奮性スパイク数



抑制性スパイク数

図 29 エッジフィルター×ガンマスパイクの結果 2

抽出範囲変更 1×ガンマスパイクの結果

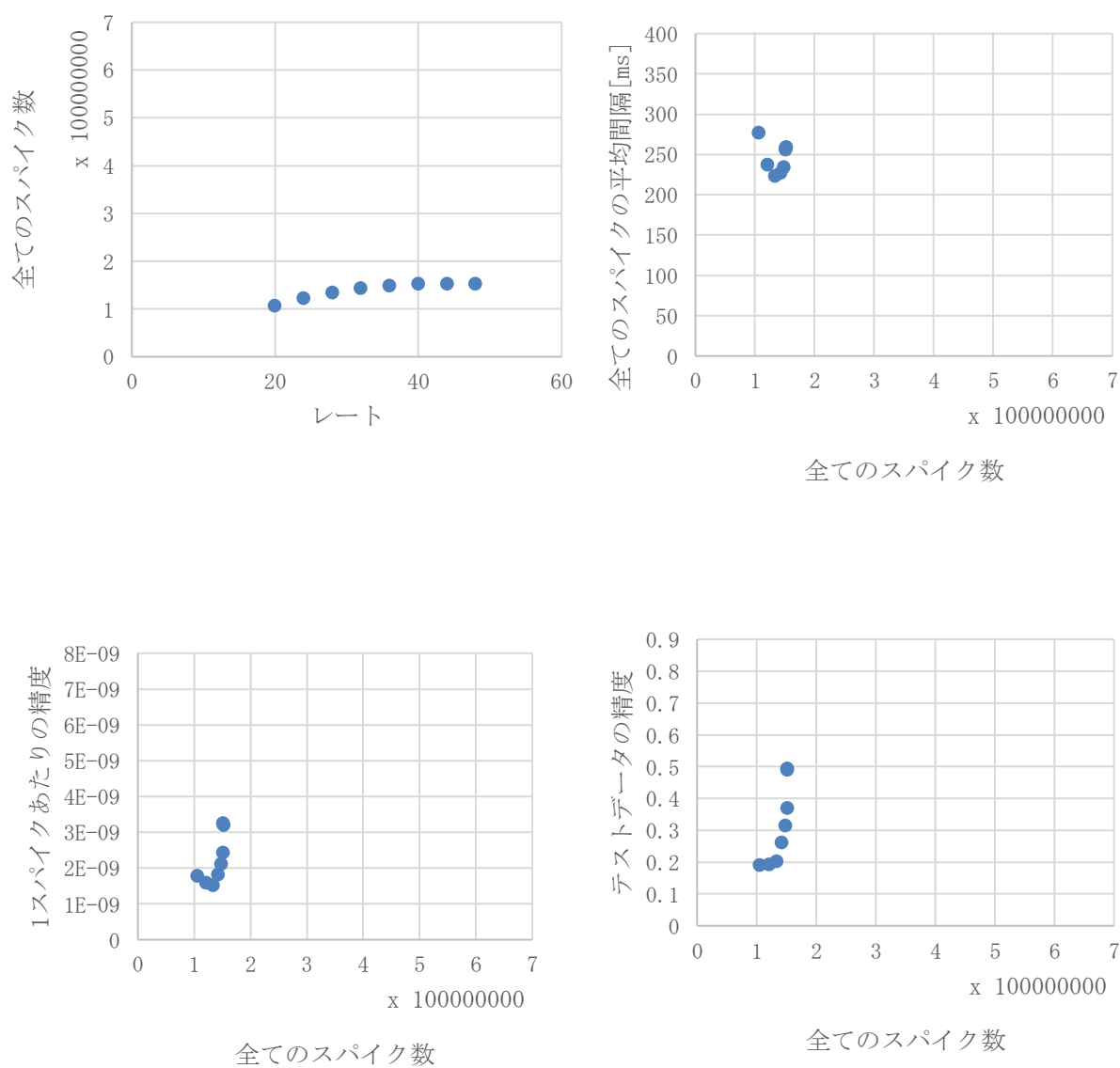
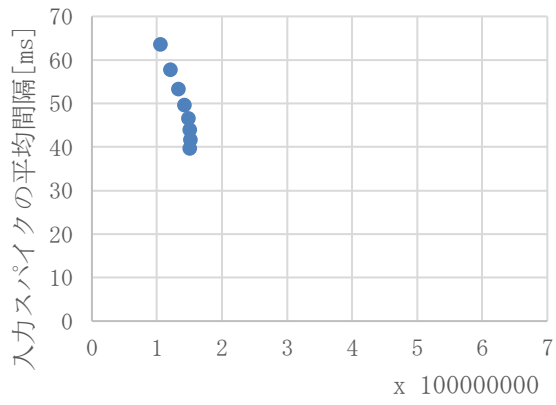
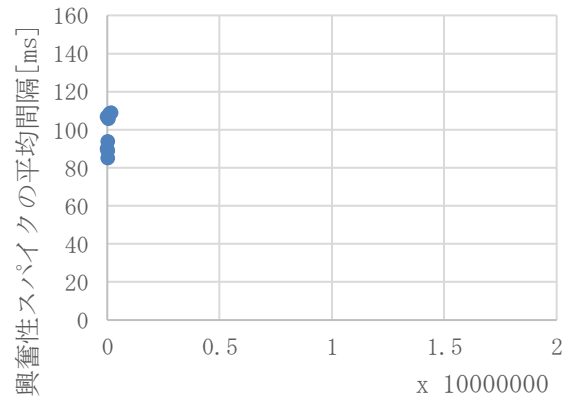


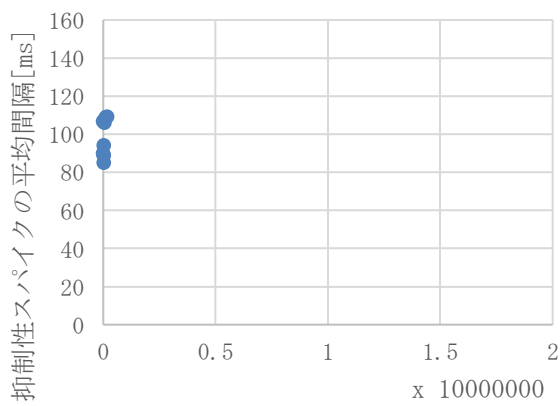
図 30 抽出範囲変更 1×ガンマスパイクの結果 1



入力スパイク数



興奮性スパイク数



抑制性スパイク数

図 31 抽出範囲変更 $1 \times$ ガンマスパイクの結果 2

抽出範囲変更 2×ガンマスパイクの結果

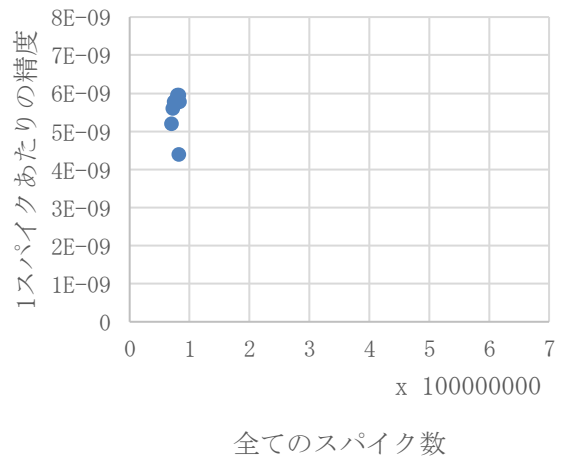
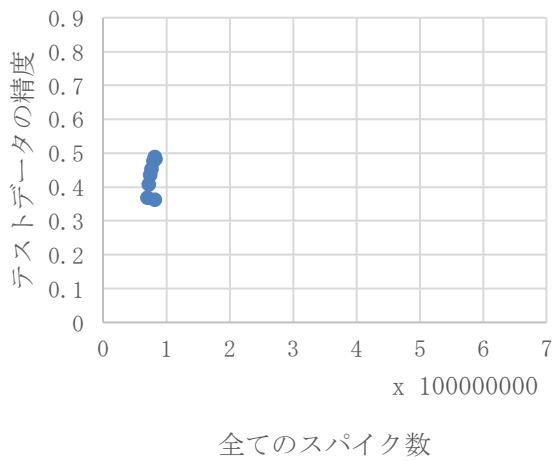
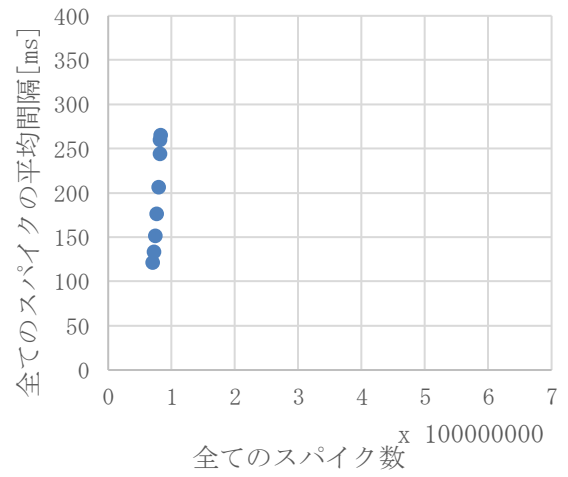
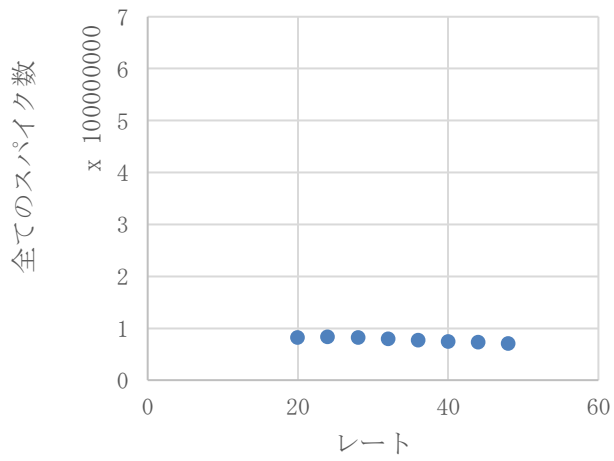


図 32 抽出範囲変更 2×ガンマスパイクの結果 1

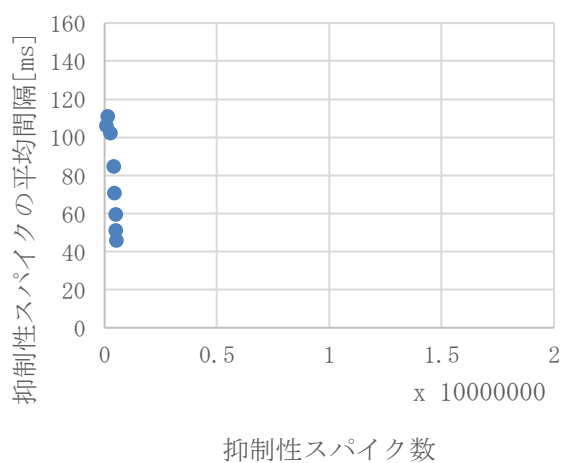
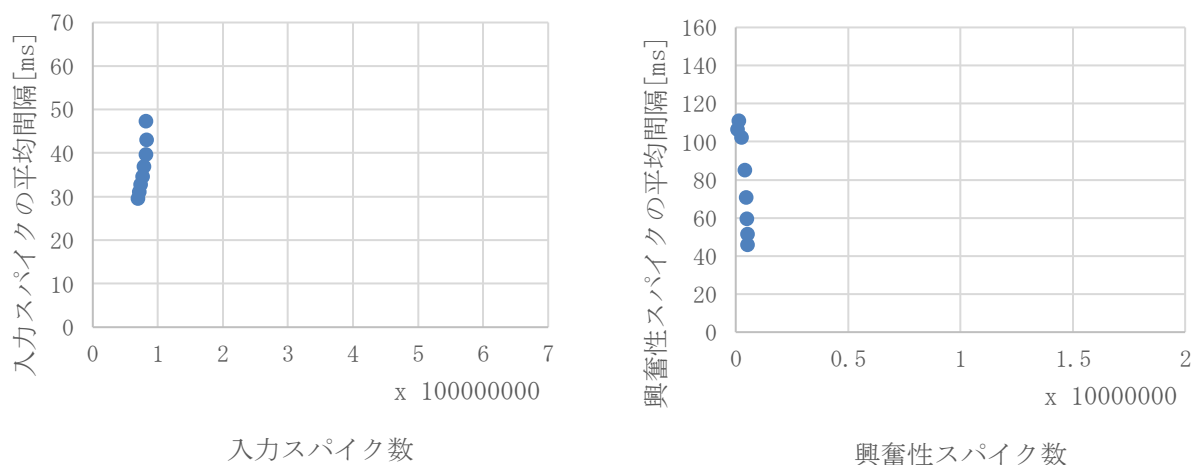


図 33 抽出範囲変更 2×ガンマスパイクの結果 2

【考察】

入力前処理を行わない符号化手法同士を比較する。

画素値に掛け合わせるレートを一定間隔で変更し、異なるスパイク数で出力した。ネットワークを学習する際、使用および出力されたスパイクの総数は、全体としてガンマ分布より出力されたガンマスパイクが少ないことが確認できる。ガンマスパイクはレートが上がるにつれスパイク数も増加しているが、緩やかな変化であった。再入力機構が存在しない、ポアソン分布より出力されたポアソンスパイクはレートが上がるにつれスパイク数が上昇した。再入力機構が有効であるポアソンスパイクの数が最も多かったが、レート 30 付近まで増加した後、減少している。レート 30 付近でピークを迎えていると考えられるが、これはレート 30 付近が最も再入力されていることを意味する。30 付近までは興奮性ニューロンの出力スパイクが設定した数より少なかったために学習中に一つの画像に対して複数回再入力起きたことが原因だと考えられる。その後減少しているのは、レートの増加とともに入力スパイク数が増え、それに伴い興奮性ニューロンの出力スパイクも一定数を超えたため再入力回数が少なくなったためだと推測できる。

- ポアソンスパイク (再入力機構あり) ▲ ポアソンスパイク (再入力機構なし)
- ガンマスパイク

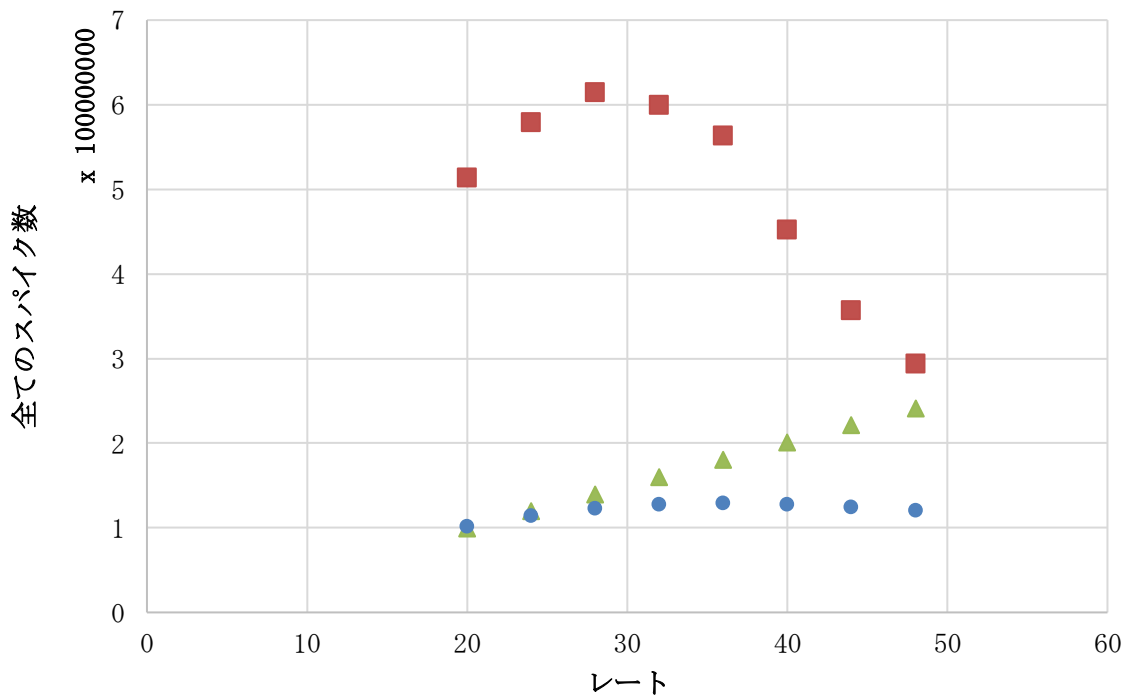


図 34 分布同士の比較 1

テストデータの分類精度を確認する。本研究で提案したガンマ分布による符号化手法は低い精度を示す場合もあるが、一部では他の手法と同様の精度を示している。よってガンマ分布を利用して符号化は可能であることが示された。

手法同士を比較すると再入力機構ありポアソン分布手法が最も高い精度を達成している。これは再入力する過程で各画素を表すのに適応したスパイク列が全ての画素でまた十分な数で出力されているためだと考えられる。再入力機構なしポアソン分布手法やガンマ分布手法は出力スパイクが一定数を下回っていても次の入力データの学習に移行するため、すべての画素で学習に十分な数のスパイクを出力できていなかった場合が考えられる。

そのためガンマ分布手法にも再入力機構と同様の仕組みを導入することでより精度が上昇する可能性が考えられる。

一方、全体として精度の上昇とともに使用および出力されたスパイクの総数も増加する。すなわち精度と消費電力が比例してしまう。3つのうち最も効率の良い手法を見つけるため1スパイクあたりの精度を確認する。

- ポアソンスパイク (再入力機構あり) ▲ ポアソンスパイク (再入力機構なし)
- ガンマスパイク

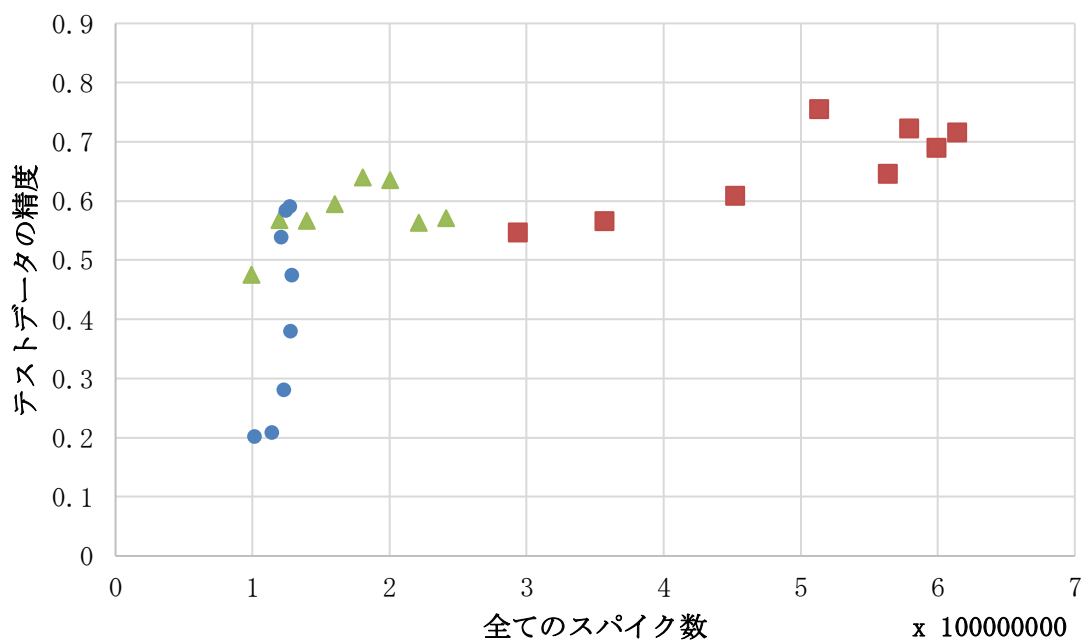


図 35 分布同士の比較 2

1 スpikeあたりの精度に関しては、最も低い値を示したのは再入力有りポアソン分布手法であった。再入力なしポアソン分布手法、ガンマ分布を使用した手法は同程度の値を示した。使用および出力されたSpike全体のSpike間隔を確認すると、ガンマ分布手法は密なSpike列を少ないニューロンで出力し、ポアソン分布手法は疎なSpike列を多くのニューロンで出力していることがわかる。

- ポアソンSpike (再入力機構あり) ▲ ポアソンSpike (再入力機構なし)
- ガンマSpike

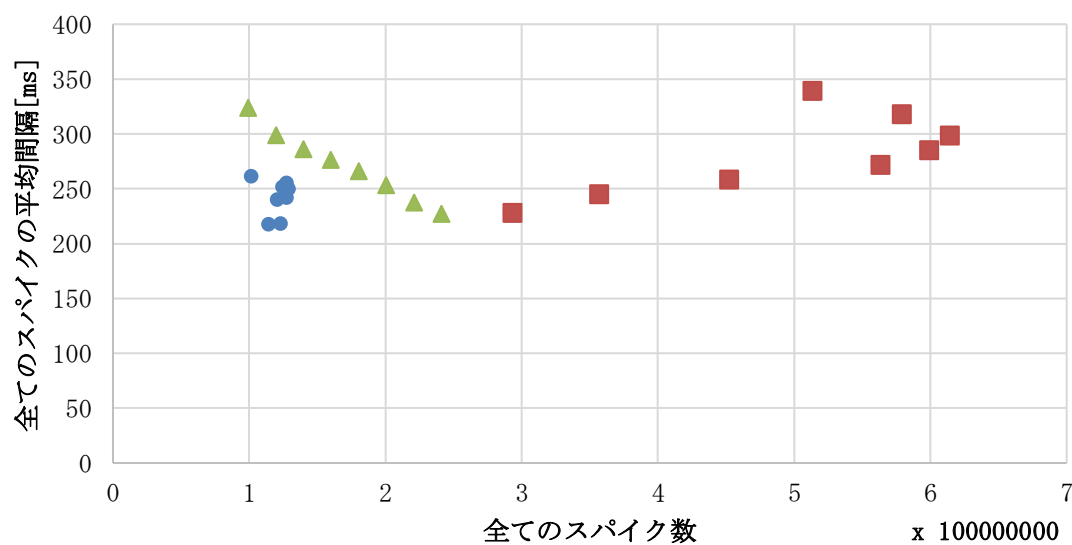
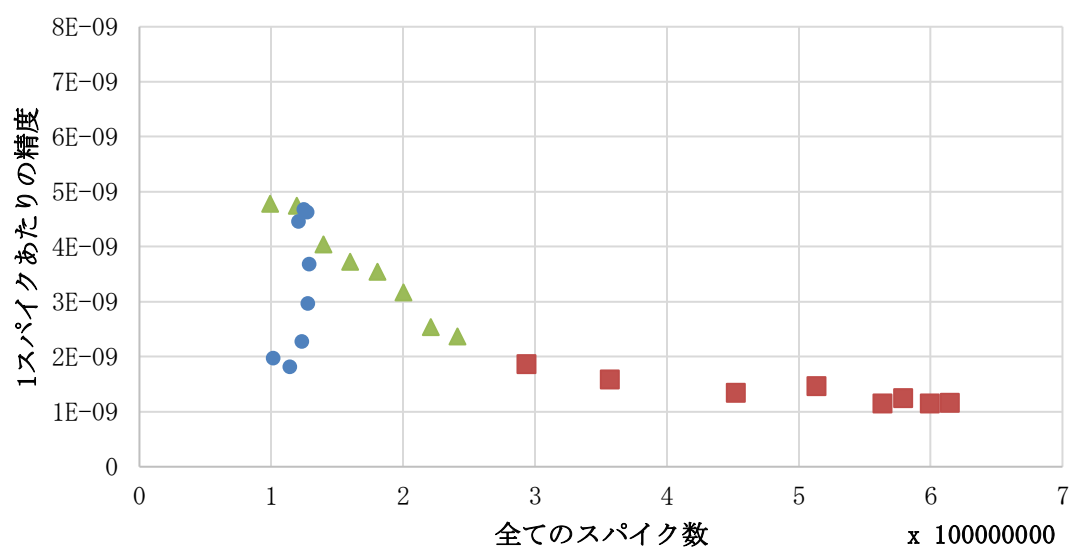


図 36 分布同士の比較 3

入力画像から作成される入力スパイクのスパイク間隔を確認すると、3つの手法が同程度の範囲に収まった。また興奮性スパイクの平均間隔と数が抑制性スパイクとおおよそ同じ数値を示すのは、1対1で対応しているためであると考えられる。またスパイクの平均間隔が100ms付近に多いのは、Diehl & Cookモデルの興奮性ニューロンの膜電位の時定数が100msに設定されていることに起因すると考察できる。

- ポアソンスパイク (再入力機構あり) ▲ ポアソンスパイク (再入力機構なし)
- ガンマスパイク

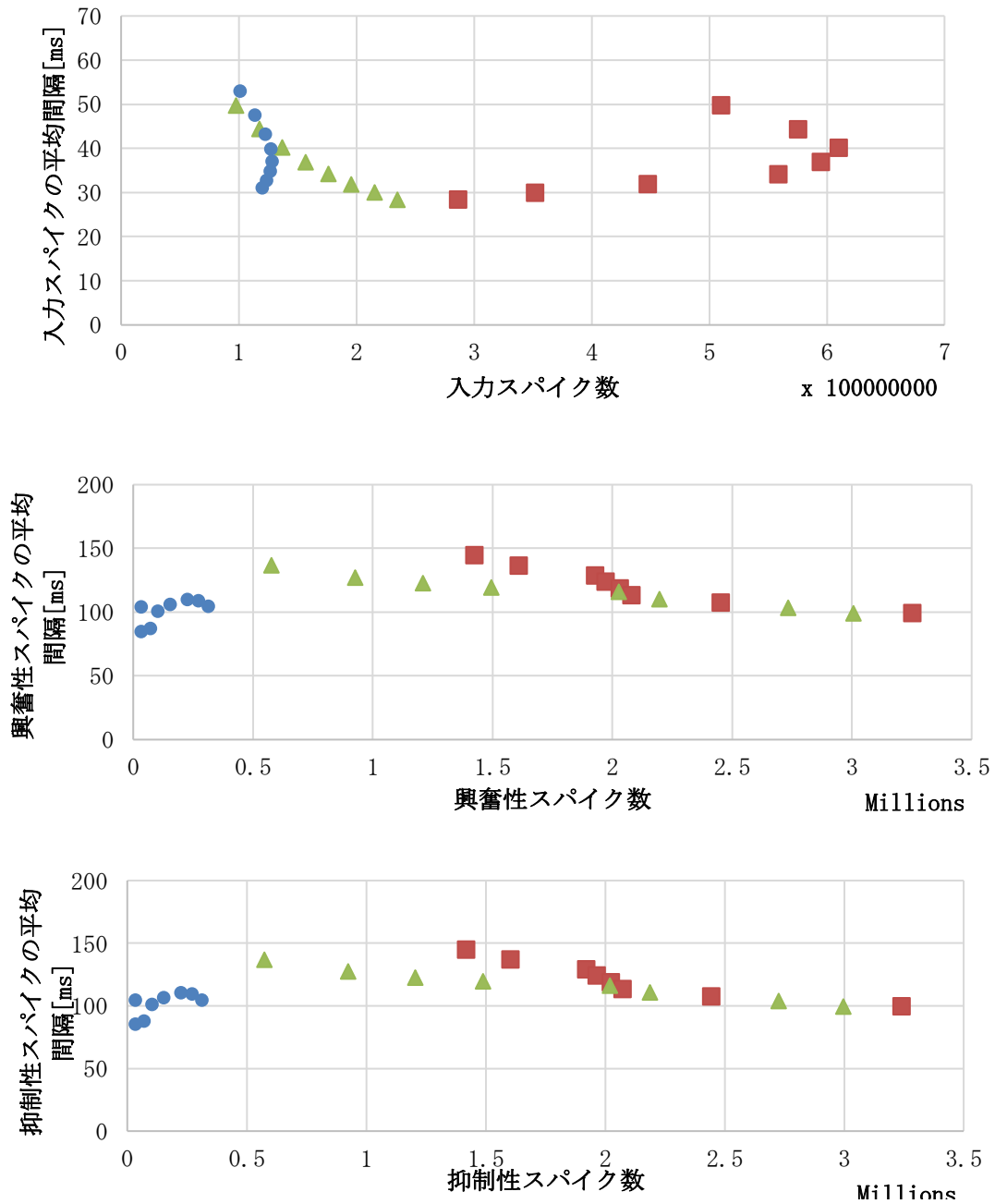


図 37 分布同士の比較 4

ここまで入力前処理を行わない符号化手法同士を比較してきたが、画素からスパイクへの変換手法は同じで入力前処理が異なる手法を比較する。まずはポアソンスパイク(再入力機構あり)を比較する。

入力前処理がない手法に比べ、エッジフィルターを適用した手法の再入力回数が増加した。これは画像にエッジフィルターを適用したことで、画素値の値が減少したり、0になる部分が出現し、入力スパイクが少なく出力され、興奮性ニューロンの反応が下がったため再入力が増加したと考えられる。「レートと全てのスパイク数の関係」「全てのスパイク数と全てのスパイクの平均間隔」「入力スパイク数と入力スパイクの平均間隔」のグラフを確認すると抽出範囲変更2のプロットが同じ形を描いているが、これは画像の再入力に起因する。例えばレート20において少ないスパイクを出力したため、レートが16追加される。レート36においてスパイクを出力し、レート20のスパイク数と足し合わせ全てのスパイク数が加算される。このように合計スパイク数が算出されるのでこのような結果になると推測できる。

- 入力前処理なし&ポアソンスパイク(再入力機構あり)
- ▲ エッジフィルター&ポアソンスパイク(再入力機構あり)
- × 抽出範囲変更1&ポアソンスパイク(再入力機構あり)
- 抽出範囲変更2&ポアソンスパイク(再入力機構あり)

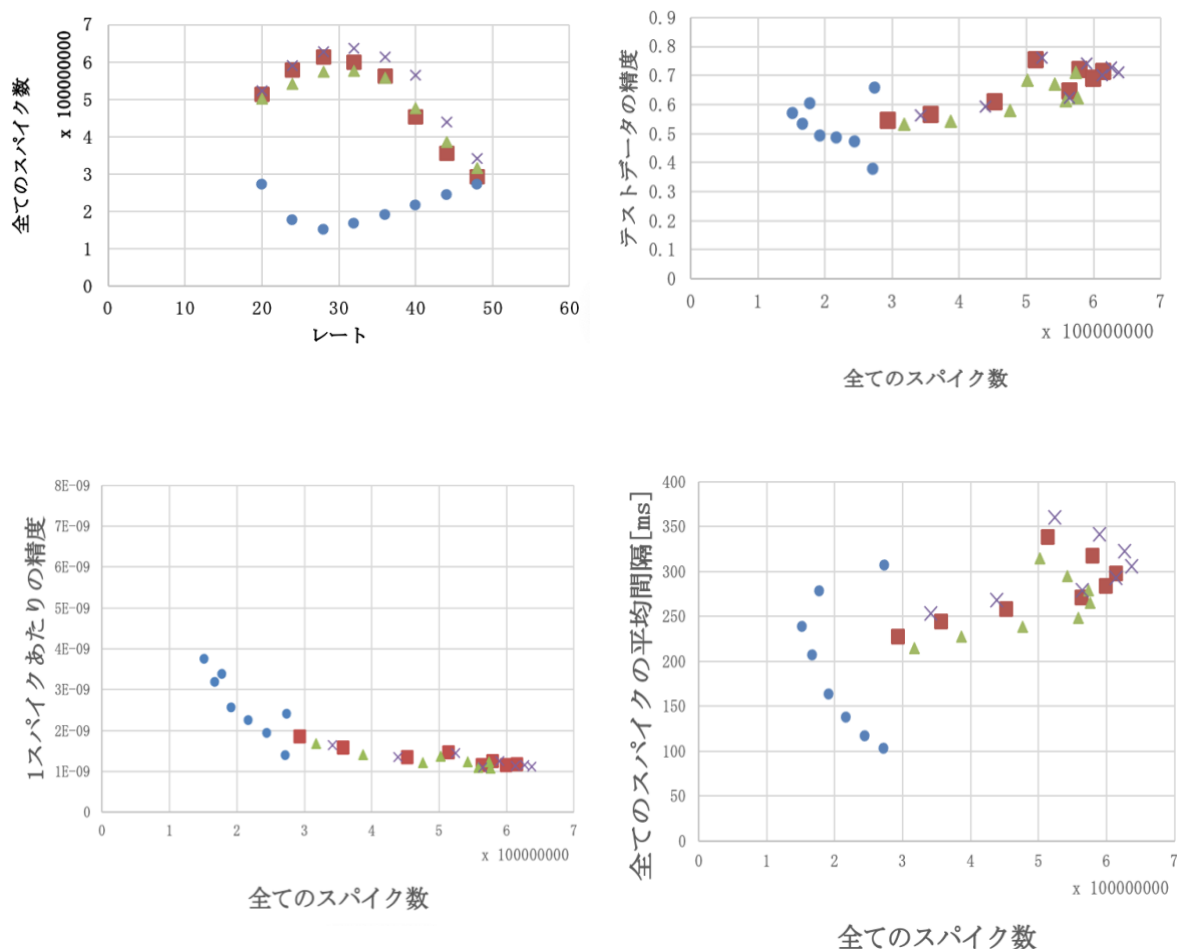


図 38 ポアソンスパイク(再入力機構あり)の比較 1

- 入力前処理なし&ポアソンスパイク (再入力機構あり)
- ▲ エッジフィルター&ポアソンスパイク (再入力機構あり)
- × 抽出範囲変更1&ポアソンスパイク (再入力機構あり)
- 抽出範囲変更2&ポアソンスパイク (再入力機構あり)

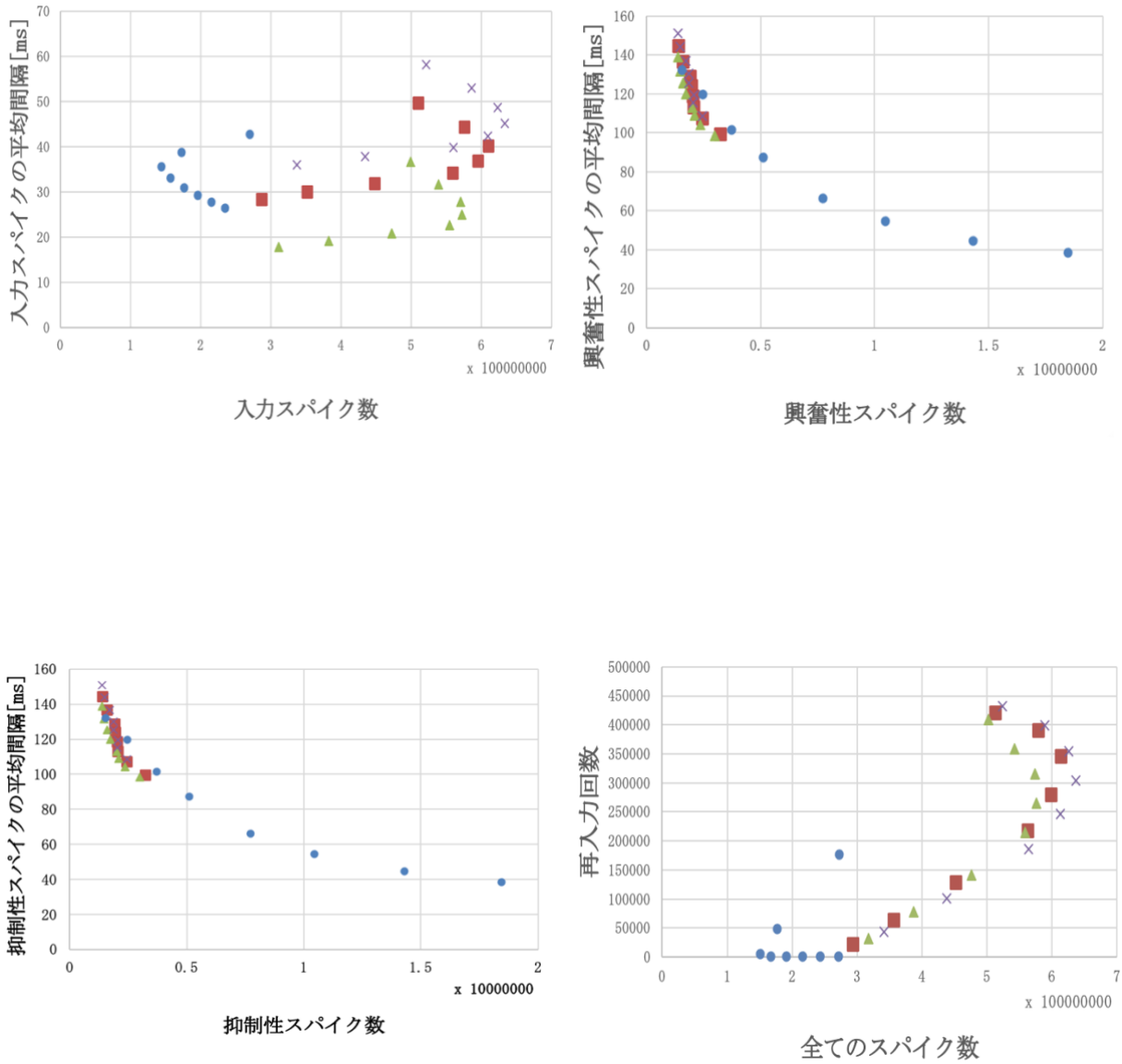


図 39 ポアソンスパイク (再入力機構あり) の比較 1

ポアソンスパイク (再入力機構なし) を比較する。ポアソン分布による手法では、レートの上昇とともに合計スパイク数が増加していることが確認できる。レートが増加し入力スパイクが出力されるほど、スパイク間隔が小さくなることが確認できる。「テストデータの精度と全てのスパイク数の関係」のグラフを確認すると、精度のピークが確認できる。これらのことからポアソンスパイクにおいて、符号化に有用な入力スパイク数と入力スパイクの平均間隔が存在することが考えられる。スパイク数が少なすぎても多すぎても、またスパイクの平均間隔が近すぎても遠すぎても精度が出ないと考えることができ、符号化においては一定の乱雑さが必要であると考えられる。

- 入力前処理なし&ポアソンスパイク (再入力機構なし)
- ▲ エッジフィルター&ポアソンスパイク (再入力機構なし)
- × 抽出範囲変更1&ポアソンスパイク (再入力機構なし)
- 抽出範囲変更2&ポアソンスパイク (再入力機構なし)

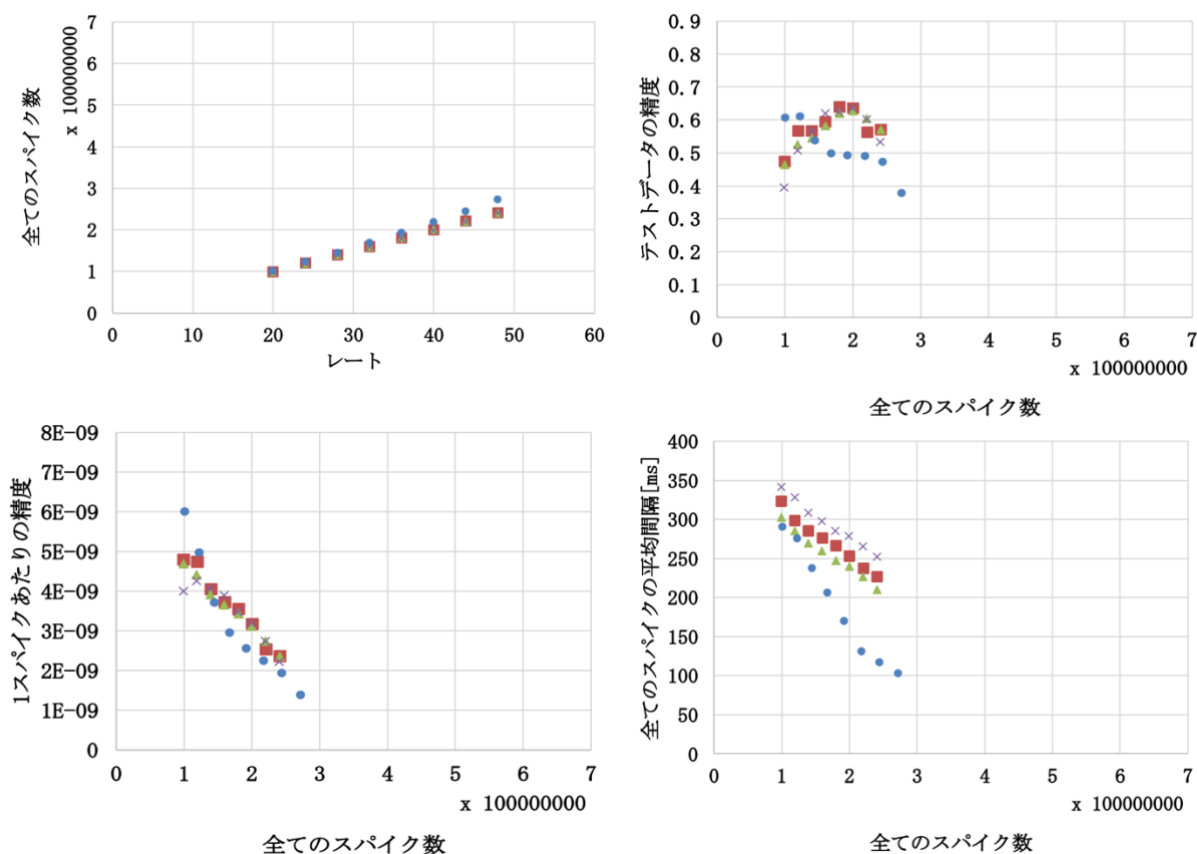


図 40 ポアソンスパイク (再入力機構なし) の比較 1

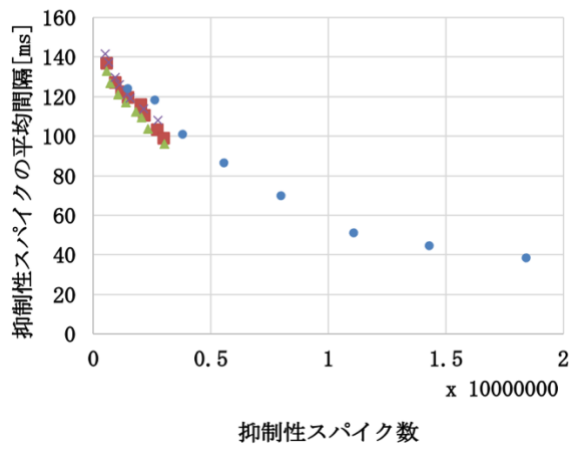
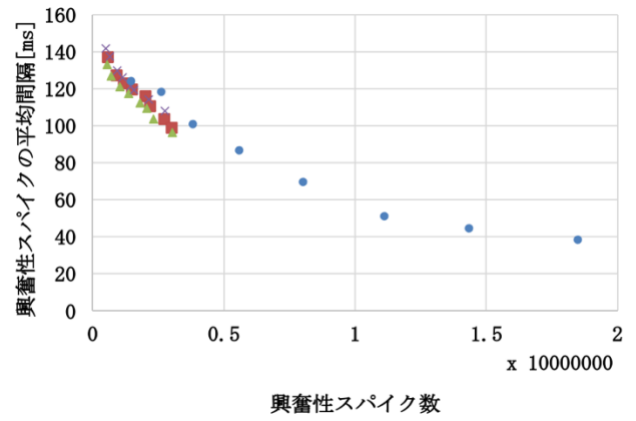
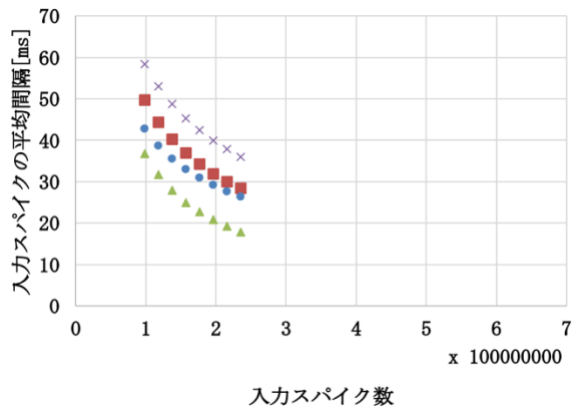


図 41 ポアソンスパイク (再入力機構なし) の比較 2

ガンマスパイクを比較する。ガンマスパイクにおいてはレートを変化によるスパイク数の変化は少なく、また入力前手法の違いにおける変化も少ないものであった。またテストデータの精度に関してはエッジフィルターとガンマスパイクの組み合わせが最も高い値を示した。また1スパイクあたりの精度に関しても最も高い値を示した。入力スパイクの平均間隔最も小さかったのもエッジフィルターとガンマスパイクの組み合わせである。入力スパイク数の大小関係と精度の大小関係が一致していないため、エッジフィルターとガンマスパイクの組み合わせが高効率を示したのは、入力スパイクの平均間隔が関係していると考えられる。

- 入力前処理なし&ガンマスパイク ▲ エッジフィルター&ガンマスパイク
- × 抽出範囲変更1&ガンマスパイク ● 抽出範囲変更2&ガンマスパイク

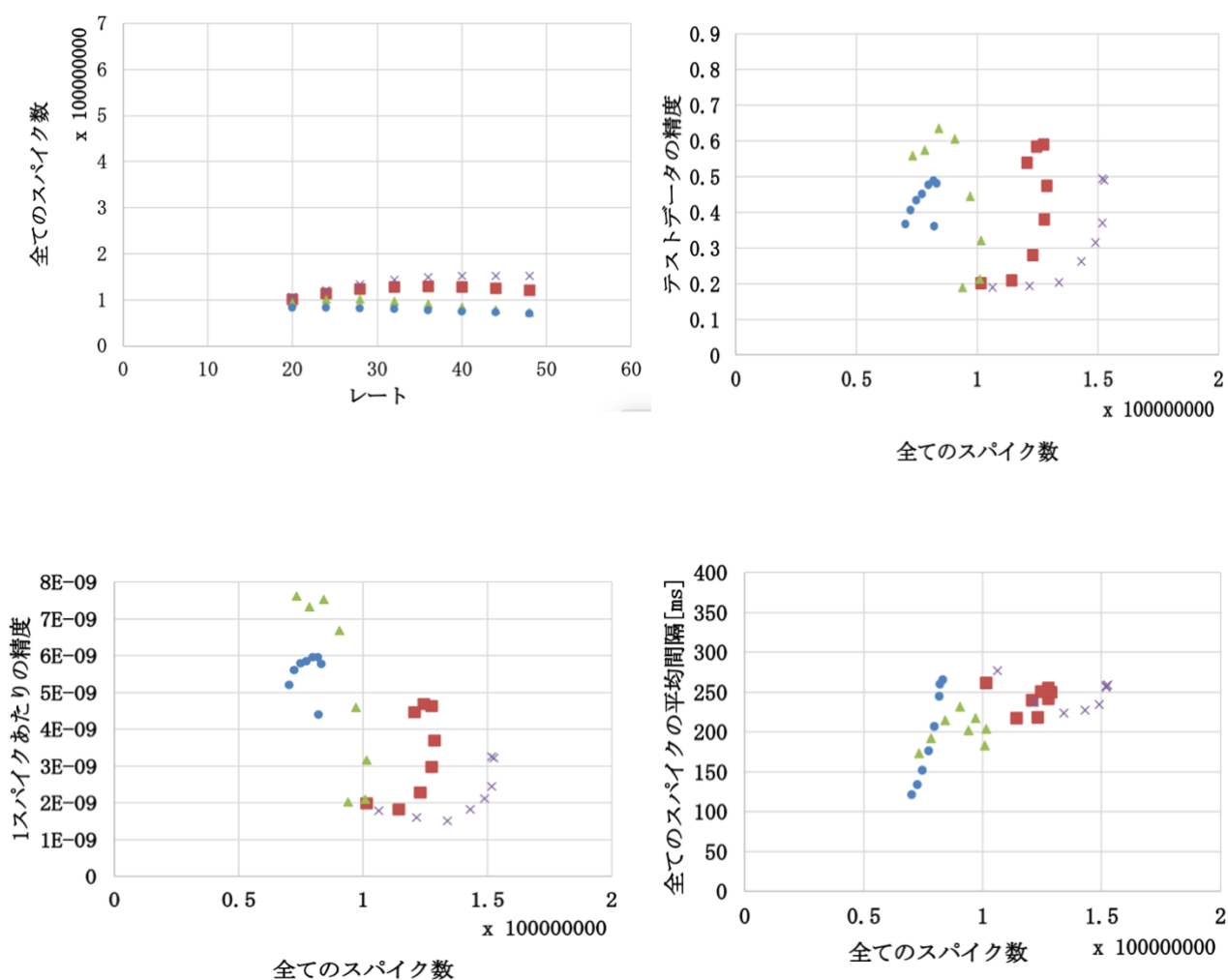


図 42 ガンマスパイクの比較 1

- 入力前処理なし&ガンマスパイク
- ▲ エッジフィルター&ガンマスパイク
- × 抽出範囲変更1&ガンマスパイク
- 抽出範囲変更2&ガンマスパイク

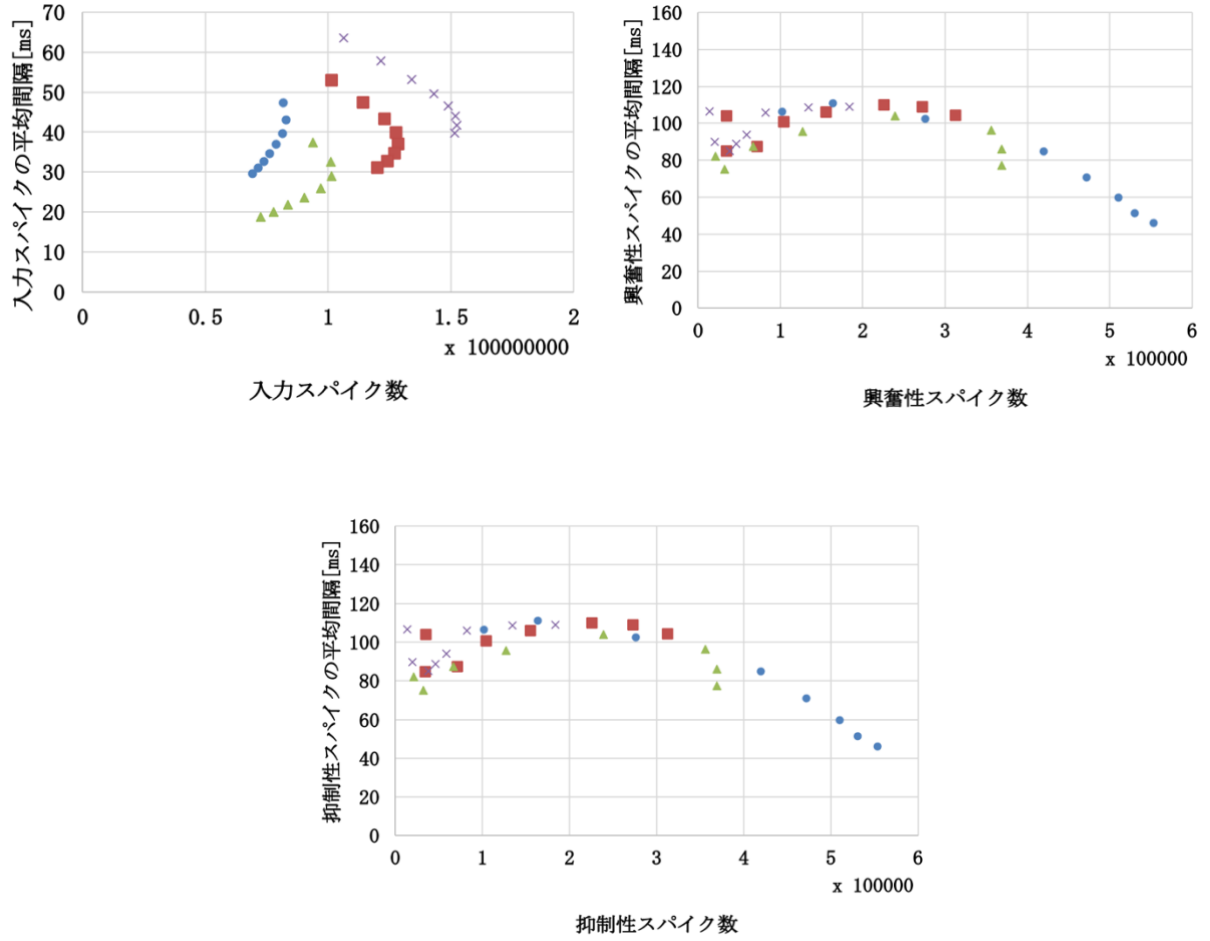


図 43 ガンマスパイクの比較 2

次にテストデータの精度が 0.5 以上を記録した手法を抽出し、考察をまとめる。

- ▲ 入力前処理なし×ポアソンスパイク(再入力機構あり) × エッジフィルター×ポアソンスパイク(再入力機構あり)
- ✖ 抽出範囲変更1×ポアソンスパイク(再入力機構あり) ● 抽出範囲変更2×ポアソンスパイク(再入力機構あり)
- + 入力前処理なし×ポアソンスパイク(再入力機構なし) ■ エッジフィルター×ポアソンスパイク(再入力機構なし)
- 抽出範囲変更1×ポアソンスパイク(再入力機構なし) ◆ 抽出範囲変更2×ポアソンスパイク(再入力機構なし)
- 入力前処理なし×ガンマスパイク ● エッジフィルター×ガンマスパイク

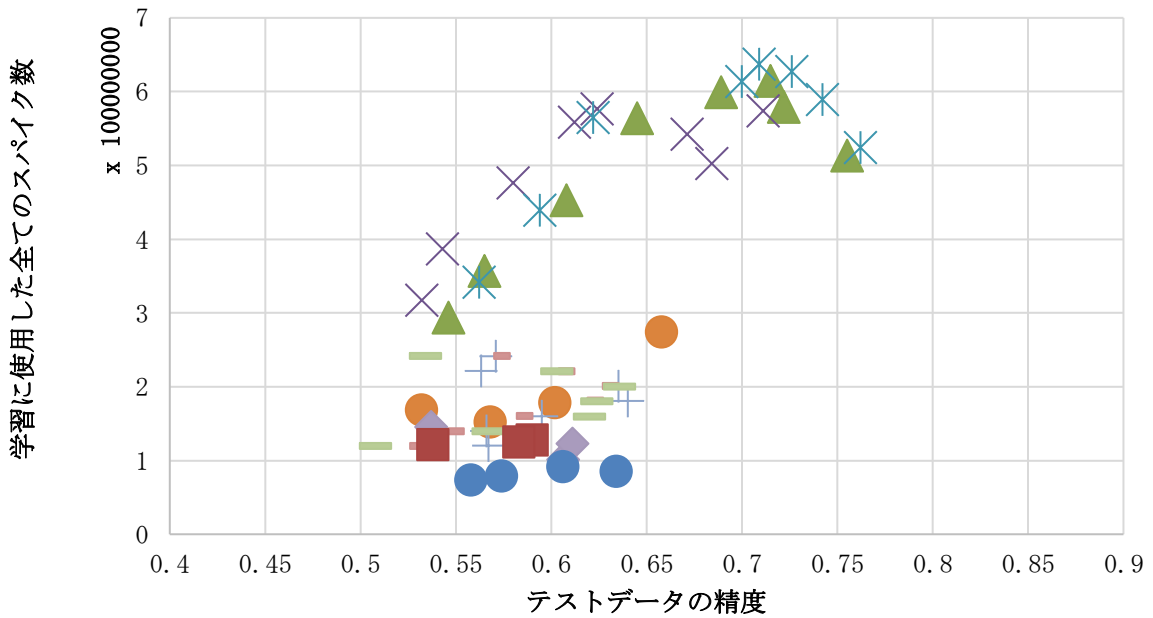


図 44 学習に使用した全てのスパイク数とテストデータの精度の関係

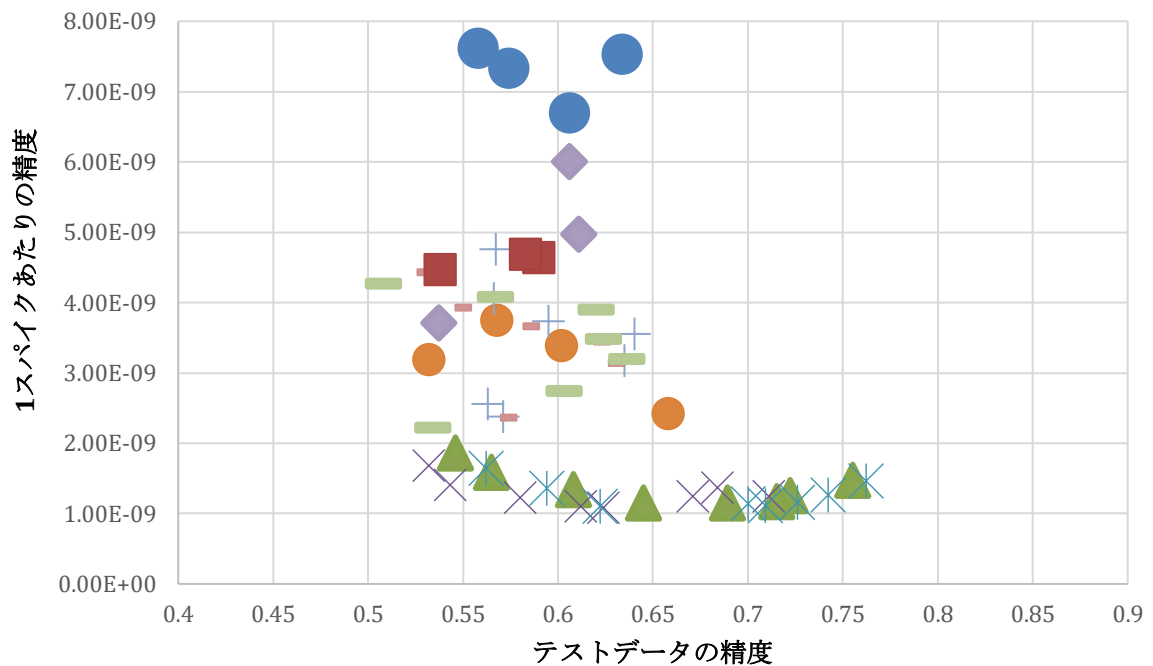


図 45 1 スパイクあたりの精度とテストデータの精度の関係

入力前処理なしのガンマスパイクの精度は0.5以上を示しており、画像からスパイクへの符号化に活用できることが示されている。抽出範囲の変更1とガンマスパイク、抽出範囲の変更2とガンマスパイクの組み合わせは精度が0.5を上回ることがなかった。入力前処理が精度に影響を与えることがわかる。テストデータの精度が最も高い手法は、抽出範囲の変更1と再入力があるポアソンスパイクの組み合わせであった。この手法は、学習に使用した全てのスパイク数も最も多い結果となった。一方、抽出範囲の変更1と再入力がないポアソンスパイクの組み合わせは最も精度が低い結果となった。抽出範囲の変更1とポアソンスパイクの組み合わせは十分なスパイクの出力ができず、再入力機構の存在が精度に影響を与えたことが考えられる。また抽出範囲の変更2の手法も全ての場合において、全てのスパイク数は低下しているものの精度も同様に低下しており、精度や効率性の向上に寄与しないことが考えられる。

学習に使用した全てのスパイク数が最も少なかったのはエッジフィルタとガンマスパイクの組み合わせであった。電力効率を評価するために1スパイクあたりの精度を確認するとどの手法よりも高い値を示した。エッジフィルタにより入力画像の情報が削減され、精度や効率性が向上したと考えられるが、ポアソンスパイクにおいては向上が確認されなかったため、スパイクの種類によって適切な入力前処理が異なることが考えられる。

5 章 結論

5-1 結論

本研究ではSNNの符号化手法に着目し、手法を考案、従来手法との学習効率を評価・比較を行った。またネットワークの学習に使用されるスパイク数や1スパイクあたりの精度を求め評価した。符号化手法を変更することでネットワークの精度が変化すること、提案したガンマ分布による符号化手法は従来使用されてきたポアソン分布による手法と同様に符号化に使用することができることを示した。

また従来手法であるポアソン分布を活用した符号化手法より、エッジフィルタとガンマ分布を使用した符号化手法が効率良く学習を行うことが確認できた。

本研究を通して、SNNにおいては、符号化手法は精度や効率について大きな影響を及ぼすことが確認された。さらなる符号化手法の研究を行うことでより大きな改善が可能である符号化手法の開発ができると予想できる。

5-2 今後の展望

今回、電力効率が良い手法を考案できたものの、精度の面では劣っていたため、高い精度と電力効率を両立する手法を考案する必要がある。

学習に使用したスパイクの数より、電力効率を測定したが、実際には画像加工処理の過程やシナプス結合の強度調整にも電力は必要である。全ての処理に使用される電力を算出することで正確な電力効率を比較する必要がある。また使用する分布の違いによりスパイクの生成にどのような影響があるのか、より深める研究が必要だ。

今回の研究ではMNISTデータセットを使用したがる、カラー画像のような複雑なデータセットや他のネットワークでも同様に有効であるか影響を検証する必要がある。

より優れた符号化手法を開発することで、SNNの用途はさらに広がると予想される。

6 章 参考文献

- [1] Auge, Daniel & Hille, Julian & Mueller, Etienne & Knoll, Alois.(2021).A Survey of Encoding Techniques for Signal Processing in Spiking Neural Networks. *Neural Processing Letters*.53.10.1007/s11063-021-10562-2.
- [2] Pfeiffer, M., & Pfeil, T. (2018). Deep learning with spiking neurons: Opportunities and challenges. *Frontiers in neuroscience*, 12, 774.
- [3] P. U. Diehl, M. Cook, (2015). Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in computational neuroscience*, 9, 99.
- [4] Nunes, J. D., Carvalho, M., Carneiro, D., & Cardoso, J. S. (2022). Spiking neural networks: A survey. *IEEE Access*, 10, 60738-60764.
- [6] 山崎匡, 五十嵐潤. (2021). はじめての神経回路シミュレーション:1 ニューロンからヒト全脳モデルまで. 森北出版
- [7] Fatahi, M., Ahmadi, M., Shahsavari, M., Ahmadi, A., & Devienne, P. (2016). evt_MNIST: A spike based version of traditional MNIST. *arXiv preprint arXiv:1604.06751*.
- [8] Lee, J. H., Delbruck, T., & Pfeiffer, M. (2016). Training deep spiking neural networks using backpropagation. *Frontiers in neuroscience*, 10, 508.
- [9] Yang, X., Zhang, Z., Zhu, W., Yu, S., Liu, L., & Wu, N. (2020). Deterministic conversion rule for CNNs to efficient spiking convolutional neural networks. *Science China Information Sciences*, 63, 1-19.
- [10] Zhu, L., Dong, S., Li, J., Huang, T., & Tian, Y. (2020). Retina-like visual image reconstruction via spiking neural model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1438-1446).
- [11] 岩沢宏和. (2016). 分布からはじめる確率・統計入門. 東京図書株式会社.
- [12] 馬場真哉. (2018). Python で学ぶあたらしい統計学の教科書. 翔泳社.
- [13] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [14] 近藤峰生. (2020). 眼科診療ビジュアルラーニング 5, 網膜、硝子体. 中山書店
- [15] 小島比呂志. (2014). 脳とニューロンの生理学-情報伝達・発生・意識. 丸善出版
- [16] Maimon, G., and Assad, J. (2009). Beyond poisson: Increased spike-time regularity across primate parietal cortex *Neuron*.62,426-440

[17] Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *The Journal of comparative neurology*, 292(4), 497–523.

<https://doi.org/10.1002/cne.902920402>

[18] <https://briansimulator.org/>

[19] <https://pytorch.org/>

7 章 謝辞

本論文は多くの方々のご協力により完成させることができました。

本研究を実施するにあたりミケレット・ルジェロ教授には研究や発表についての熱心なご指導を賜りました。深く御礼申し上げます。

そして日々の議論を通じて多くの知識やアドバイスを頂戴いたしましたミケレット研究室の皆様には深く感謝いたします。

最後に家族や友人、本論文執筆に携わってくれた全ての方々に感謝の意を表し、謝辞とさせていただきます。