## Импульсные нейронные сети в генеративных моделях.

Импульсные нейронные сети являются третьим поколением нейронных сетей, которые имитируют нейронную динамику мозга. В импульсных нейронных сетях информация представлена в виде бинарных временных рядов и имеет событийно-ориентированную архитектуру. Поэтому они могут работать со сверхвысокой скоростью и сверхнизким энергопотребление на нейроморфных чипах. На Loihi [8], TrueNorth [1] и Neurogrid [4] импульсные нейронные сети могут экономить энергию на порядки по сравнению с искусственными нейронными сетями. Например, на TrueNorth вычислительное время приблизительно в 1/100 раз ниже, а потребление энергии в 1/100 000 раз ниже, чем в обычных нейронных сетях [7]. Кроме того, импульсные нейронные сети показали многообещающие результаты не только на относительно простых задачах таких, как задачи классификации [10], [22], [25] и оценка оптического потока [18], но и на более сложных, таких как обнаружение объекта [16] и слежение за объектом [24]. Однако применение импульсных нейронных сетей не ограничивается указанными задачами. Например, генеративные модели на основе импульсных нейронных сетей не были достаточно изучены. По данному направлению существует лишь несколько работ, основные из которых мы рассмотрим в этой лекции и проведем сравнительный анализ.

#### 1 Импульсная генеративно-состязательная нейросеть.

Генеративно-состязательная нейронная сеть (GAN) [12] является одной из наиболее популярных и широко исследуемых генеративных моделей, которая имеет широкое применение (обзор можно посмотреть в [2]). GAN состоит из двух сетей: генератора и дискриминатора, которые конкурируют друг с другом. Генератор пытается обмануть дискриминатор, генерируя синтетические/ложные сэмплы из априорного шума. Дискриминатор, в свою очередь, пытается отличить сэмплы, сгенерированные генератором, от реальных данных, классифицируя их как ложные или настоящие.

В работе [17] предложена импульсная генеративно-состязательная сеть – первая генеративная модель, которая использует только импульсные нейронные сети. В архитектуре используется модель импульсного IF нейрона (integrate-and-fire) с высоким рефракторным периодом, который гарантирует генерацию нейроном максимум одного импульса.



Рис. 1: Архитектура импульсной генеративно-состязательной нейронной сети.

Генератор: Случайный шумовой вектор сэмплируется из дискретного равномерного распределения  $Uni\{0, t_{max}\}$ , после чего преобразуется энкодером в импульсную последовательность с помощью метода временного кодирования time-to-first-spike (подробнее в [17]). Генератор принимает эту импульсную шумовую последовательность z на вход и выводит ложную импульсную картинку G(z). Эта картинка имеет такую же размерность,

как и входящие импульсные картинки. Далее ложная импульсная картинка декодируется в ложную картинку по схеме, противоположной энкодеру.

Дискриминатор: Тренировочные картинки кодируются в настоящие импульсные картинки энкодером и подаются на вход дискриминатору. Дискриминатор – это бинарный классификатор, на вход которому подаются либо настоящая импульсная картинка, либо ложная импульсная картинка из генератора. Он имеет два выходящих нейрона: ложный (0) и настоящий (1). Метка класса определяется с помощью времени импульса. Если настоящий нейрон произвел импульс раньше ложного нейрона, тогда входящая картинка классифицируется, как настоящая, в противном случае – ложная.

Для тренировки применяется стохастический градиентный спуск и метод обратного распространения ошибки на временной области, а функция потерь наименьших квадратов (least squares loss) используется для определения целевой функции. Чтобы применить метод обратного распространения ошибки необходима дифференцируемость в классическом понимании функции активации нейрона. Однако ипульсные IF нейроны имеют недифференцируемую функцию активации, поскольку импульс – это дискретное событие, которое либо 1, когда генерируется импульс, либо 0. Для преодоления этой проблемы рассматривается функция активации ReLU (Rectified Linear Unit), которую аппроксимирует функция активации импульсного нейрона. Предполагая эквивалентность между функцией активации ReLU и временем срабатывания IF-нейрона, можно вычислить производные целевой функции относительно каждого синаптического веса.

Генератор и дискриминатор обучаются вместе поочередно. Во время обучения дискриминатора используются ложные картинки, сгенерированные генератором, но веса генератора не меняются. Аналогично при обучении генератора используется прогноз дискриминатора и потеря распространяется назад из дискриминатора в генератор, но веса дискриминатора остаются неизмененными. И генератор, и дискриминатор тренируются один раз каждую эпоху. После обучения генератор, который теперь приблизительно производит выборку из того же распределения, что и набор тренировочных данных, используется для создания ложных сэмплов.

Импульсная генеративно-состязательная сеть тренируется на базе MNIST для каждой отдельной цифры. Генератор, как и дискриминатор, являются двухслойными импульсными нейросетями, где каждый нейрон второго слоя связан со всеми нейронами первого слоя. Ниже показаны некоторые сэмплы, сгенерированные после тренировки нейросети для 50 эпох и сэмплы, сгенерированные обычной генеративно-состязательной сетью.

0	١	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---

(а) картинки из базы MNIST (тренировка)



(b) картинки, сгенерированные обычной GAN



(с) картинки, сгенерированные импульсной GAN

Рис. 2: Сравнение картинок, сгенерированных обычной генеративно-состязательной сетью и аналогичной импульсной.



Рис. 3: Некоторые сэмплы, сгенерированные импульсной генеративно-состязательной сетью, имеют неоднозначную интерпретацию.

Качество сгенерированных картинок невысокое. Кроме того, некоторые сгенерированные сэмплы имеют неоднозначную интерпретацию. Отметим, что данная импульсная генеративно-состязательная сеть применялась

только к относительно простой базе MNIST.

Основное различие между импульсными генеративно-состязательными нейронными сетями и обычными GAN состоит в том, что первые генерируют изображение на каждом временном шаге. Качество картинок, сгенерированных на разных временных шагах, отличается, что приводит к искажениям. Чтобы решить эту проблему в [11] предложена импульсная генеративно-состязательная нейросеть со специальным методом декодирования (SGAD). Этот метод динамически присваивает различные веса различным временным шагам, тем самым создавая изображения более высокого качества посредством динамического процесса комбинирования. Экспериментальные результаты показали улучшение качества сгенерированных сэмплов для базы MNIST. Кроме того, предложенный подход достигает неплохих значений FID (Fréchet Inception Distance) для Fashion MNIST, CIFAR-10 и CelebA.



Рис. 4: Картинки, сгенерированные SGAD.

# 2 Импульсный вариационный автоэнкодер.

Вариационный автоэнкодер (VAE) [9] – это генеративная модель, состоящая из энкодера, который преобразует входные данные в скрытые переменные, имеющие более сжатое и компактное представление, и декодера, который использует эти скрытые переменные для реконструкции входных данных. Однако, в отличии от обычных автоэнкодеров, которые преобразуют каждый объект входных данных в одну скрытую переменную, вариационный автоэнкодер преобразует каждый объект входных данных в набор параметров, которые определяют распределение вероятностей в скрытом пространстве, а затем делается выборка из этого распределения для реконструции входных данных. Это позволяет не только точно реконструировать входные данные, но и генерировать новые сэмплы, похожие на входные данные. Механизм работы вариационного автоэнкодера связан с гипотезой о том, как мозг изучает визуальную информацию [13]. По этой причине, создание импульсного вариационного автоэнкодера – естественный процесс.

В работе [15] был предложен импульсный вариационный автоэнкодер (FSVAE) для генерации изображений. Самый сложный аспект создания импульсного вариационного автоэнкодера – это создание пространства скрытых переменных. В случае обычного вариационного автоэнкодера скрытое пространство является непрерывным, которое часто представленно нормальным распределением. Однако в рамках импульсных нейронных сетей сэмплирование из нормального распределения невозможно, поскольку импульс – это бинарное событие. Поэтому в работе [15] предложены авторегрессионные модели на основе импульсных нейронных сетей, а скрытые переменные случайным образом выбираются из выходных данных этих моделей, позволяя сэмплировать из процессов Бернулли. В архитектуре используется итеративная модель импульсного LIF нейрона (leaky integrate and fire) [23].



Рис. 5: Тренировка импульсного вариационного автоэнкодера.

Во время обучения, входящая картинка x преобразуется в импульсную последовательность x(t),  $1 \le t \le T$  с помощью прямого кодирования [20] и последовательно подается в импульсный энкодер, выходом из которого является импульсная последовательность  $x_t^E$ . Для генерации скрытых импульсных переменных  $z_t$  моделируется апостериорное распределение вероятностей с помощью авторегрессионной импульсной нейронной сети (Рис. 6). На вход авторегрессионной импульсной нейронной сети подается не только  $x_t^E$ , но и предыдущие сгенерированные скрытые переменные  $z_{t-1}$ . Скрытые переменные  $z_t$  выбираются случайным образом одно за другим из выходных данных авторегрессионной импульсной нейронной сети, позволяя сэмплировать из распределения Бернулли. Далее скрытые импульсные переменные  $z_t$  последовательно подаются в импульсный декодер, который выводит импульсную последовательность  $\hat{x}(t)$ . Для декодирования импульсной последовательности  $\hat{x}(t)$  в картинку  $\hat{x}$  используются неактивные нейроны в выходящем слое.



Рис. 6: Слева – авторегрессионная модель на основе импульсной нейронной сети  $f_q$  для апостериорного распределения вероятностей. Справа – подробное описание импульсной нейронной сети  $f_q$ . Сэмплирование из распределения Бернулли производится с помощью случайного выбора C каналов из kC, где C – размерность  $z_t$ , а  $k \ge 2$  – натуральное число.

Во время сэмплирования априорное распределение вероятностей генерирует скрытую импульсную переменную  $z_t$ , которая затем подается на вход декодеру для генерации картинки  $\hat{x}$  (Рис. 7). Авторегрессионная модель для априорного распределение вероятностей имеет такую же архитектуру, как и для апостериорного (Рис. 6), однако на вход подаются только импульсные скрытые переменные  $z_{t-1}$ . Предполагается, что  $z_0 = 0$ .



Рис. 7: Во время сэмплирование априорное распределение вероятностей генерирует скрытые импульсные переменные  $z_t$  последовательно, которые затем декодируются в картинку  $\hat{x}$ .

Целевая функция определяется как сумма среднеквадратической ошибки (MSE) между исходной картинкой x и реконструированной  $\hat{x}$ , и квадрат максимального среднего расхождения (MMD) между апостериорным распределение  $Ber(\pi_{q,t})$  и априорным  $Ber(\pi_{p,t})$ .

Для тренировки импульсного вариационного автоэнкодера применяется пространственно-временной алгоритм обратного распространения ошибки с пакетной нормализацией, зависящей от порога (STBP-tdBN) [25]. Для его применения необходима дифференцируемость в классическом понимании функции активации LIF нейрона. Из-за ее отсутствия рассматривается аппроксимация её производной.

Эксперименты проводятся с базами MNIST, FashionMNIST, CIFAR-10 и CelebA. Для MNIST и Fashion MNIST использовались 60 000 изображений для обучения и 10 000 для тестирования. Размер входных изображений был изменен до 32×32. Для CIFAR-10 использовались 50 000 изображений для обучения и 10 000 для тестирования. Для CelebA использовались 162 770 изображений для обучения и 19 962 для тестирования. Размер входных изображений был изменен до 64×64. Энкодер и декодер имеют одинаковое количество свёрточных слоев, равное 4 для MNIST, FashionMNIST, CIFAR-10, и равное 5 для CelebA. После каждого слоя производится пакетная нормализация, зависящая от порога. Модель тренируется 150 эпох со скоростью обучения 0.001 и таким же весовым убыванием.



Рис. 8: Картинки, сгенерированные VAE (слева) и импульсным VAE (справа).

В сравнении с обычным автоэнкодером импульсный вариационный автоэнкодер генерирует чуть более четкие изображения, но размытость все-таки сохраняется, что ограничивает возможности его применения в практических приложениях. Размытость изображений связана с формулировкой оптимизации, а именно, вариационные автоэнкодеры оптимизируются путем максимизации нижней границы доказательств (ELBO) относительно сетевых параметров. ELBO состоит из суммы двух слагаемых: первое – это реконструкционная потеря, которая гарантирует, что наблюдаемый сэмпл может быть отображен в скрытое пространство через апостериорную модель и реконструирован обратно из его возможного скрытого представления, второе – это расстояние Кульбака-Лейблера (KL divergence), которое представляет близость априорного и апостериорного распределения вероятностей. Так как скрытое пространство имеет меньшую размерность, то идеальная реконтрукция невозможна. Реконструкционная потеря определяет какие признаки будут иметь наибольший вес. В рассматриваемом импульсном вариационном автоэнкодере ELBO состоит из суммы реконструкционной потери в виде среднеквадратической ошибки (MSE) и максимального среднего расхождения (MMD), которое является более подходящей метрикой расстояния для импульсных последовательностей. Среднеквадратическая ошибка (MSE) не присваивает веса специфическим признакам, все признаки будут реконструированны с похожим весом. Так как спектр мощности естественных изображений находится на более низкой частоте, то низкочастотные признаки, или другими словами размытые, будут доминировать в реконструкционной потере. Следовательно, реконструкционная точность в высокочастотных признаках будет менее важна [5].

## 3 Импульсные диффузионные модели шумоподавления.

Диффузионные вероятностные модели шумоподавления или просто диффузионные вероятностные модели – это класс генеративных моделей, которые на каждом временном шаге добавляют гауссовый шум к данным (процесс прямой диффузии), а затем обучаются обращать процесс диффузии, чтобы получить исходные данные (процесс обратной диффузии). Процесс генерации сэмпла состоит из множества небольших шумоподавляющих шагов.

Впервые диффузионные модели были предложены в [21]. В [14] авторы предложили определенную параметризацию модели, которая не только упрощает обучение, но и способствует генерации высококачественных сэмплов, иногда даже превосходящих по качеству сэмплы, сгенерированные другими моделями. После этой работы интерес к диффузионным вероятностным моделям вырос. Появились работы в таких областях, как аудиомоделирование, видеомоделирование, прогнозирование временных рядов, а также преобразование текста в речь. В [6] предложена импульсная диффузионная вероятностная модель (SDDPM), которая сочетает в себе энергоэффективность импульсных нейронных сетей с превосходной генеративной производительностью, сравнимой с обыкновенными нейронными сетями. В архитектуре использовалась итеративная модель импульсного LIF нейрона.



Рис. 9: Архитектура импульсной диффузионной вероятностной модели.

В кодирующем слое случайный Гауссовский шум преобразуется в последовательность импульсов. Далее импульсная последовательность вместе с "time embeddings" (дополнительный вектор, кодирующий временные отметики) подаются на основную компоненту архитектуры – импульсную U-net. Она состоит из сужающегося пути (слева) и расширяющегося пути (справа). На каждом этапе сужающегося пути разрешение картинки уменьшается, а количество каналов свойств увеличивается. Каждый шаг расширяющегося пути восстанавливает разрешение и уменьшает количество каналов. Эти два пути связаны пропускными соединениями, которые необходимы для восстановления пространственной информации, потерянной при уменьшении разрешения. Сужающийся и расширяющийся пути состоят из нескольких остаточных блоков (Pre-Spike Residual Block), каждый из которых содержит импульсные слои, свёрточные слои и пакетную нормализацию.



Рис. 10: Доимпульсный остаточный блок.

Импульсная U-net передает только импульсы, представленные 0 или 1. Далее, импульсы из U-net передаются в декодирующий слой, где воспроизводится картинка. Процесс сэмплирования состоит из множества шагов шумоподавления, где каждый шаг проходит через импульсную диффузионную вероятностную модель.



Рис. 11: сэмплирование.

Для тренировки используется SpikingJelly – пакет Python для обучения импульсных нейронных сетей, в котором алгоритм обучения основан на методе суррогатного градиента https://spikingjelly.readthedocs.io/ zh-cn/latest/index.html#index-en. Чтобы преодолеть недифференцируемость в классическом понимании функции активации импульсного LIF-нейрона, рассматривается производная суррогатной функции, которая по форме похожа на функцию активации LIF-нейрона, а также является непрерывной и гладкой. В качестве целевой функции рассматривается среднеквадратическая ошибка (MSE).

Эксперименты проводятся с базами MNIST, Fashion MNIST, CIFAR-10, CelebA и LSUN bedroom. Размер входных изображений был изменен до 32 × 32 для MNIST, Fashion MNIST, CIFAR-10 и до 64 × 64 для CelebA и LSUN bedroom. Для тренировки использовались 60 000 изображений для баз MNIST, Fashion MNIST и CIFAR-10, 200 000 для базы CelebA и 59 миллионов для базы LSUN. Для тестирования использовались 10 000 изображений для всех баз.



Рис. 12: Картинки, сгенерированные импульсной диффузионной вероятностной моделью.

Визуальное представление картинок, сгенерированных импульсной дифузионной вероятностной моделью, говорит о высоком качестве сэмплов по сравнению с другими импульсными генеративными моделями. Перед тем, как представить сравнительный анализ в метрике IS (Inception Score) и FID (Fréchet Inception Distance), хотелось бы упомянуть еще одну работу, в которой также рассматривается импульсная диффузионная вероятностная модель.

В работе [19] авторы предложили дискретную диффузионную генеративную модель с векторным квантованием на основе импульсных нейросетей (Spiking-Diffusion). Первый этап включает в себя построение импульсного вариационного автоэнкодера с векторным квантованием (VQ-SVAE). Основное отличие обычного вариационного автоэнкодера (VAE) от аналогичного с векторным квантованием (VQ-VAE) в том, что в первом – скрытое пространство непрерывно, а во втором – представленно в виде дискретных кодов. Преобразование данных в набор дискретных кодов или дискретную матрицу осуществляется с помощью процедуры векторного квантования, в которой используется кодовая книга. В случае VQ-SVAE для создания кодовой книги необходимо преобразовать импульсную последовательность во вложенные признаки, так как хранение импульсной последовательности непосредственно в кодовой книги требует большого объема памяти. Для этого используется постсинаптический потенциал (PSP) и частота импульсов (SFR), а также оператор *k* для оптимизации весов PSP и SFR. Чтобы предотвратить потерю информации при восстановлении импульсной последовательности из вложенных признаков строится адаптивный импульсный генератор. На втором этапе производится поглощающая диффузия, которая заключается в преобразовании дискретной матрицы в маску за фиксированное число временных шагов [3]. После этого строится импульсный диффузионный декодер, который восстанавливает дискретную матрицу, обращая процесс поглощающей диффузии.

Эксперименты, проводимые на базе MNIST, Fashion MNIST, CIFAR-10, показывают превосходство данной модели над импульсным вариационным автоэнкодером (FSVAE). Картинки, сгенерированные FSVAE, размыты, в то время как картинки, сгенерированные Spiking-Diffusion, имеют более четкие границы.



Рис. 13: Картинки, сгенерированные FSVAE и Spiking-Diffusion на базе MNIST и Fashion MNIST.

Однако, импульсная диффузионная вероятностная модель шумоподавления превосходит все упомянутые модели по качеству сгенерированных картинок на всех базах. Следующая таблица представляет сравнительный анализ генеративных импульсных моделей SGAD, FSVAE, SDDPM, Spiking-Diffusion и обычной диффузионной модели DDPM [14] на базах MNIST, Fashion MNIST, CIFAR-10, CelebA, LSUN bedroom.

Например, на базе CelebA импульсная диффузионная модель шумоподавления обеспечивает 4-х кратное и 6-и кратное улучшение FID по сравнению с FSVAE и SGAD. На базе CIFAR-10 улучшение еще более существенное даже по сравнению с моделью Spiking-Diffusion. Более того, с увеличением импульсных шагов качество сгенерированных сэмплов повышается. Если сравнивать с обычной диффузионной моделью DDPM, то для некоторых баз данных качество сгенерированных SDDPM сэмплов сравнимо, а в некоторых случаях даже превосходит, например, для базы CIFAR-10.

Из всех импульсных генеративных моделей импульсная диффузионная вероятностная модель демонстрирует производительность, сопоставимую с производительностью аналогичной неимпульсной диффузионной модели, обеспеченивая сохранение энергии.

Dataset	Model	Method	#Param (M)	Time Steps	IS↑	FID↓
	DDPM [14]	ANN	64.47	/	-	28.70
	FSVAE [15]	SNN	3.87	16	6.209	97.06
MNIST	SGAD [11]	SNN	-	16	-	69.64
	Spiking-Diffusion [19]	SNN	-	16	-	37.50
	SDDPM [6]	SNN	63.61	4	-	29.48
	DDPM [14]	ANN	64.47	/	-	20.24
Fashion	FSVAE [15]	SNN	3.87	16	4.551	90.12
MNIST	SGAD [11]	SNN	-	16	-	165.42
	Spiking-Diffusion [19]	SNN	-	16	-	91.98
	SDDPM [6]	SNN	63.61	4	-	21.38
	DDPM [14]	ANN	64.47	/	-	20.34
CelebA	FSVAE [15]	SNN	6.37	16	3.697	101.60
	SGAD [11]	SNN	-	16	-	151.36
	SDDPM [6]	SNN	63.61	4	-	25.09
LSUN bedroom	DDPM [14]	ANN	64.47	/	-	29.48
	SDDPM [6]	SNN	63.61	4	-	47.64
	DDPM [14]	ANN	64.47	/	8.380	19.04
	FSVAE [15]	SNN	3.87	16	2.945	175.50
	SGAD [11]	SNN	-	16	-	181.50
CIEAP 10	Spiking-Diffusion [19]	SNN	-	16	-	120.50
CIFAR-10	SDDPM [6]	SNN	63.61	4	7.440	19.73
	SDDPM [6]	SNN	63.61	8	7.584	17.27

Рис. 14: Сравнительный анализ качества сэмплирования основных моделей, измеренный в терминах IS (Inception Score) и FID (Fréchet Inception Distance).

## Литература

- F. Akopyan, J. Sawada, A. Cassidy, R. Alvarez-Icaza, J. Arthur, P. Merolla, N. Imam, Y. Nakamura, P. Datta, G.-J. Nam, et al. Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip, IEEE TCAD, 34(10), 2015.
- [2] H. Alqahtani, M. Kavakli-Thorne, G. Kumar, Applications of generative adversarial networks (gans): An updated review. Archives of Computational Methods in Engineering 28(2), 2021.
- [3] J. Austin, D. D. Johnson, J. Ho, D. Tarlow and R. van den Berg, Structured denoising diffusion models in discrete state-spaces, Advances in Neural Information Processing Systems, vol. 34, 2021.
- [4] B. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. Chandrasekaran, J. Bussat, R. Alvarez-Icaza, J. Arthur, P. Merolla and K. Boahen, Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations. Proceedings of the IEEE, 102(5), 2014.
- [5] G. Bredell, K. Flouris, K. Chaitanya, E. Erdil, Ender Konukoglu, Explicitly Minimizing the Blur Error of Variational Autoencoders, arXiv:2304.05939v1, 2023.
- [6] J. Cao, Z. Wang, H. Guo, H. Cheng, Q. Zhang, R. Xu, Spiking Denoising Diffusion Probabilistic Models, arXiv: 2306.17046, 2023.
- [7] A.S. Cassidy, R. Alvarez-Icaza, F. Akopyan, J. Sawada, J.V. Arthur, P. Merolla, P.Datta, M.G. Tallada, B. Taba, A. Andreopoulos, A. Amir, S.K. Esser, J. Kusnitz, R. Appuswamy, C. Haymes, B. Brezzo, R. Moussalli, R. Bellofatto, C.W. Baks, M. Mastro, K. Schleupen, C.E. Cox, K. Inoue, S.E. Millman, N. Imam, E. McQuinn, Y.Y. Nakamura, I. Vo, C. Guok, D. Nguyen, S. Lekuch, S. W. Asaad, D. J. Friedman, B. L. Jackson, M. Flickner, W.P. Risk, R. Manohar and D.S. Modha, RealTime Scalable Cortical Computing at 46 Giga-Synaptic

OPS/Watt with ~  $100 \times$  Speedup in Time-to-Solution and ~  $100,000 \times$  Reduction in Energy-to-Solution. International Conference for High Performance Computing, Networking, Storage and Analysis, 2014.

- [8] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, Sri Harsha Choday, G. Dimou, P. Joshi, N. Imam, Sh. Jain, et al. Loihi: A neuromorphic manycore processor with on-chip learning, IEEE Micro, 38(1), 2018.
- [9] D. P. Kingma, M. Welling, Auto-Encoding Variational Bayes, arXiv: 1312.6114v11, 2022.
- [10] P.U. Diehl and M. Cook, Unsupervised Learning of Digit Recognition using Spike-Timing-Dependent Plasticity, Frontiers in Computational Neuroscience, 9(99), 2015.
- [11] L. Feng, D. Zhao, Yi Zeng, Spiking Generative Adversarial Network with Attention Scoring Decoding, Neural Networks, V. 178, 2024.
- [12] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems, V. 2, MIT Press, Cambridge, MA, USA, 2014.
- [13] K. Han, H. Wen, J. Shi, KH. Lu, Y. Zhang, D. Fu and Z. Liu, Variational autoencoder: An unsupervised model for encoding and decoding fMRI activity in visual cortex, NeuroImage 198, 2019.
- [14] J. Ho, A. Jain and P. Abbeel, Denoising diffusion probabilistic models, 34th Conference on Neural Information Processing Systems, Vancouver, Canada, 2020.
- [15] H. Kamata, Y. Mukuta and T. Harada, Fully spiking variational autoencoder. In AAAI, volume 36, 2022.
- [16] S.J. Kim, S. Park, B. Na and S. Yoon, SpikingYOLO: Spiking Neural Network for Energy-Efficient Object Detection, Proceedings of the AAAI Conference on Artificial Intelligence, 34(07), 2020.
- [17] V. Kotariya and U. Ganguly, Spiking-gan: A spiking generative adversarial network using time-to-first-spike coding, International Joint Conference on Neural Networks, IEEE, 2022.
- [18] C. Lee, A. Kosta, A.Z. Zhu, K. Chaney, K. Daniilidis and K. Roy, Spike-FlowNet: Event-Based Optical Flow Estimation with Energy-Efficient Hybrid Neural Networks, Computer Vision – ECCV 2020: 16th European Conference, 2020.
- [19] M. Liu, J. Gan, R. Wen, T. Li, Y. Chen, H. Chen, Spiking-Diffusion: vector quantized discrete diffusion model with spiking neural networks, arXiv:2308.10187v4, 2023.
- [20] B. Rueckauer, I.-A. Lungu, Y. Hu, M. Pfeiffer, and S.-C. Liu, Conversion of Continuous-Valued Deep Networks to Efficient Event-Driven Networks for Image Classification, Frontiers in Neuroscience, V. 11, 2017.
- [21] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, Deep Unsupervised Learning using Nonequilibrium Thermodynamics, Proceedings of the 32nd International Conference on Machine Learning, V. 37, France, 2015.
- [22] J. Wu, Y. Chua, M. Zhang, H. Li and K.C. Tan, A Spiking Neural Network Framework for Robust Sound Classification. Frontiers in Neuroscience, 12(836), 2018.
- [23] Y. Wu, L. Deng, G. Li, J. Zhu, Y. Xie and L. Shi, Direct Training for Spiking Neural Networks: Faster, Larger, Better, In AAAI, 2019.
- [24] J. Zhang, Bo Dong, H. Zhang, J. Ding, F. Heide, B. Yin, and Xin Yang. Spiking transformers for eventbased single object tracking, In Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition, 2022.
- [25] H. Zheng, Y. Wu, L. Deng, Y. Hu and G. Li, Going Deeper With Directly-Trained Larger Spiking Neural Networks. Proceedings of the AAAI Conference on Artificial Intelligence, 35(12), 2021.