

AEGNN: Asynchronous Event-based Graph Neural Networks

作者：Simon Schaefer, Daniel Gehrig, Davide Scaramuzza

期刊：CVPR

年份：2022

Project page: <https://uzh-rpg.github.io/aegnn/>

1、问题背景/研究动机

本文研究的主要问题：有效利用事件信息稀疏性和异步特性的事件处理范式

- 现有方法存在的问题：

1、事件信息的特性与传统计算机视觉算法**不兼容**。传统方法将事件处理为密集、类似图像的表示，**无法有效地利用事件的稀疏性和异步特性**，导致**计算冗余**。

2、若采用几何学习方法（递归点云处理/异步稀疏卷积）

- Pros：基于图的方法只考虑事件之间的稀疏连接，并将消息传递限制在这些连接上，使计算复杂度降低。

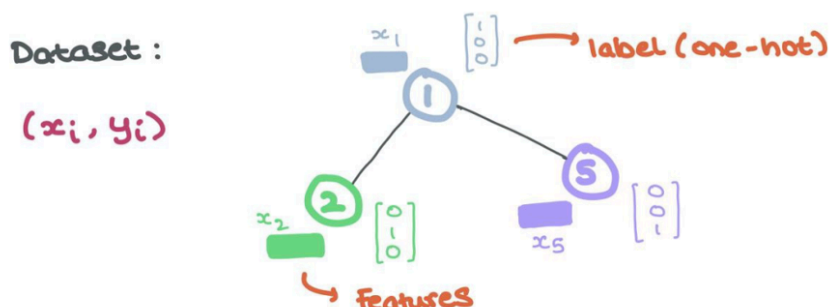
- Cons：造成**缺乏层次化学习能力**，限制了其在复杂任务上的**可扩展性**；**丢失事件的时间信息**，导致性能下降。

- 图神经网络（GNN）

图深度学习背后的思想是**学习具有节点和边的图的结构和空间特征**，这些节点和边表示实体及其交互。



- 图 $G(V,E)$ 是包含一组顶点 $i \in v$ 和一组连接顶点 i 和 j 的边 $e_{ij} \in E$ 的数据结构。
- **节点**表示一个实体或对象，节点具有所表示实体的一系列属性。这些节点属性形成了节点特征。



- **边**也可以有特征。虽然节点本身有各自的特征，但边可以有不同的边特征。通常，具有相似特征或属性的节点相互连接。GNN利用学习特定节点如何以及为什么相互连接，GNN会查看节点的邻域。

2、解决方法

- (1) 提出**基于事件的异步图神经网络** (AEGNN)，将事件建模为**时空图**，保留时间信息，充分利用其稀疏性和高时间分辨率。
- (2) 使用高效的更新规则，**将计算限制在受到新事件影响的节点**，只对少数节点进行计算，从而显著减少计算和延迟。
- (3) 将本文方法应用于目标识别和检测任务，并进行实验评估。

3、具体实现

- 图构建

事件相机具有独立的像素，每当它们感知到亮度变化时，都会触发事件。将一个时间窗口 ΔT 内的一组事件，表示为有序列表：

$$(x_1, y_1, t_1, z_1), (x_2, y_2, t_2, z_2), \dots, (x_n, y_n, t_n, z_n)$$

- 1、通过**均匀采样**将事件流进行子采样，得到一部分事件。

2、将这些子采样的事件组成一个事件图。其中，每个事件作为一个节点，事件之间的连接关系由它们的邻居关系决定。

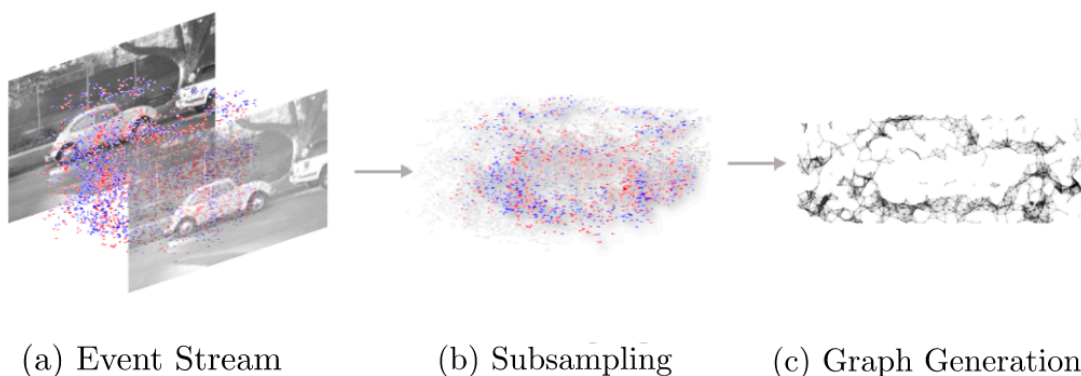
3、将事件的时间位置归一化为一个因子 β ，将其映射到与空间坐标相似的范围内。

4、对于每一对节点 i 和 j ，如果它们在时空距离 R 内，则生成它们之间的一条边 e_{ij} 。

5、将最大邻域节点数限制为 D_{max} 。

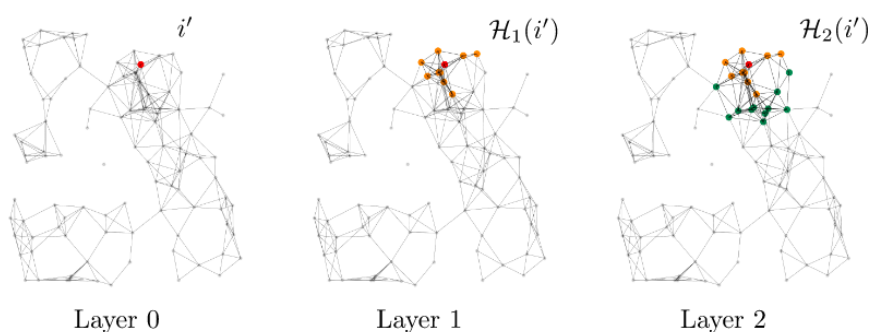
6、赋予初始节点特征 $x_i = p_i$ 和对应于连接顶点之间相对位置的边特征。

由此，得到一个初始的稀疏图，图的顶点表示事件，边表示事件之间的连接关系。与事件数据的其他表示方法相比，该方法保留了事件流的完整时间分辨率。



- 异步处理

当新事件进入时间窗口时，传统方法需要重新计算所有的激活，造成计算浪费。为了解决这一问题，文中提出了一个递归更新规则，根据事件的发生顺序和时间戳进行逐个处理，以最小的计算量包含新的事件。



具体而言，当新事件到达时，只需要更新与该事件相关的节点和边，而不需要重新计算整个图。这是通过将事件视为图的一部分，并根据事件的邻居关系进行局部更新来实现的。这种递归更新规则使得每次图的更新仅限于新事件的k-hop子图，k表示图卷积的层数。

通过这种方式，AEGNN能够实现低延迟的事件处理，且灵活地适应事件的发生顺序和时间间隔的变化，显著减少计算量，提高处理效率。

• 图卷积

包括两个步骤：

1) 在初始化过程中，将卷积应用于全图，存储生成的图，即存储顶点和边以及它们的特征。

2) 在处理步骤中，每插入一个新的顶点到图中，图只发生局部变化。一个全图的更新等价于应用等式从新顶点开始更新其k跳子图。

$$\begin{aligned}\hat{\mathbf{z}}_i &= \sum_{j \in \mathcal{N}(i)} \psi_{\Theta}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{e}_{ij}) \\ \hat{\mathbf{x}}_i &= \gamma_{\Theta}(\mathbf{x}_i, \mathbf{z}_i) \text{ for all } i \in \mathcal{N}(i')\end{aligned}$$

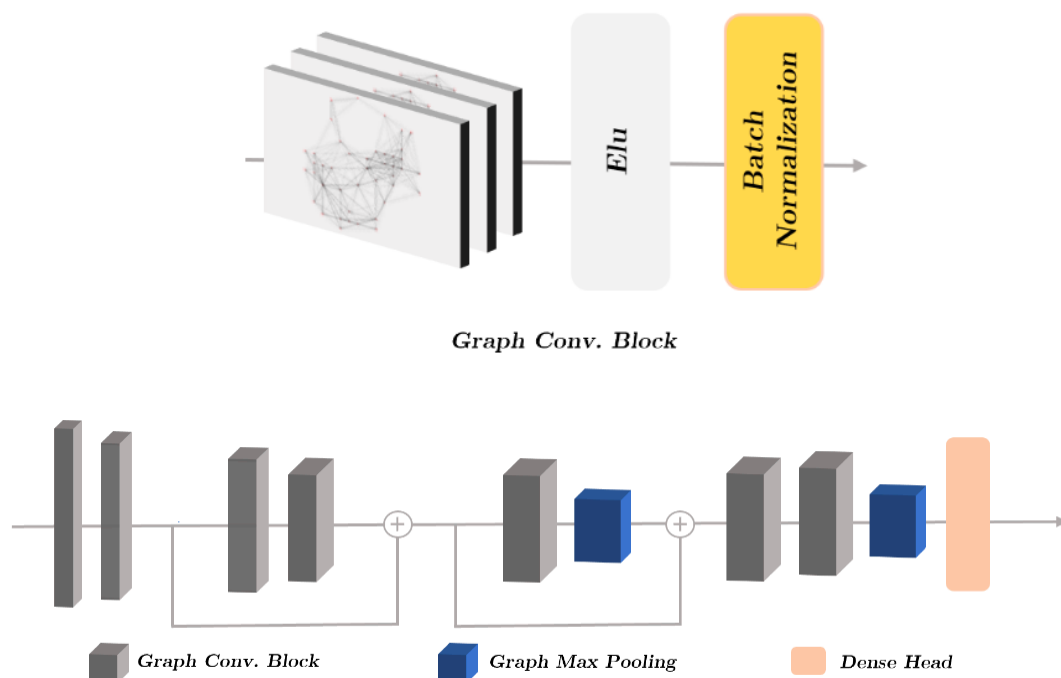
• 图池化

图池化操作将一个图转化为一个更粗糙的图。在这项工作中，重点研究基于簇的池化方法，将图节点聚集成簇，这些聚类中心的新特征通过聚合每个聚类中的特征来计算：

将这些操作作为层进行堆叠，可以实现丰富的、高层次的特征计算

- 网络架构

由7个图卷积块和2个池化层组成



AEGNN采用图神经网络(GNN)架构来处理基于事件的视觉数据。网络架构由多层组成，它将**图卷积与池化层**结合，每个图卷积块由若干个图卷积组成，依次进行激活函数和批归一化。

- **输入**：基于事件流构造的时空图，其中每个节点表示一个事件，边缘表示事件之间的时间关系。
- **结构**：遵循分层学习方法，信息在多层图中传播。其中，每一层包括三个主要步骤：**传递**、**聚合**和**更新**。
 1. 传递：图中的每个节点接收来自相邻节点的信息，并用于更新节点的特征。
 2. 聚合：对接收到的特征信息进行聚合，有助于获得图的全局上下文。
 3. 更新：将聚合的特征与当前节点的特征相结合，更新节点的表示。然后将更新后的表示传递给下一层进行进一步处理。

AEGNN体系结构允许**分层学习**和**捕获事件信息中的时间关系**。最后一层的输出可用于各种任务，eg. 目标识别或检测。

4、实验对比

文中将AEGNN应用于目标识别和目标检测任务中，通过叠加每一层的计算复杂度，推导该模型的前向通道的计算复杂度。

- 目标识别

• 数据集

1. **神经形态N - Caltech101数据集**：包含用真实事件相机记录的事件流，代表8246个事件序列中的101个对象类别，每个事件序列长300 毫秒。
2. **N - Cars数据集**：有汽车或背景真实事件。它有24029个事件序列，每个事件序列长100毫秒。

• 实验

与几种具有不同事件表示的异步和同步方法进行比较，其中事件表示包括脉冲、时间平面、直方图与图。其中，**EST**是一种可学习的、密集的事件表示，与下游任务联合优化。虽然获得了很好的识别精度，但它使用学习到的表示引入了额外的数据处理，且不能异步建模。**AsyNet**提出了一种基于事件直方图的异步稀疏网络，但它没有明确表示事件的时间关系。**NVS - S**和**EvS - B**使用了基于图的事件表示。与EvS - B中使用的标准图卷积不同，而AEGNN使用样条卷积，编码的是空间信息。

Methods	Representation	Async.	N-Caltech101		N-Cars	
			Accuracy ↑	MFLOP/ev ↓	Accuracy ↑	MFLOP/ev ↓
H-First [39]	Spike	✓	0.054	-	0.561	-
HOTS [28]	Time-Surface	✓	0.210	54.0	0.624	14.0
HATS [51]	Time-Surface	✓	0.642	4.3	0.902	0.03
DART [43]	Time-Surface	✓	0.664	-	-	-
YOLE [7]	Event-Histogram	✓	0.702	3659	0.927	328.16
EST [17]	Event-Histogram	✗	0.817	4150	0.925	1050
SSC [20]	Event-Histogram	✗	0.761	1621	0.945	321
AsyNet [35]	Event-Histogram	✓	0.745	202	0.944	21.5
NVS-S [31]	Graph	✓	0.670	7.8	0.915	5.2
EvS-S [31]	Graph	✓	0.761	11.5	0.931	6.1
Ours	Graph	✓	0.668	0.369	0.945	0.03

从表中，可以看出AEGNN在达到最先进的性能的同时，总体上具有最低的计算复杂度。在N - Cars上取得了最好的精度，与次优的异步方法AsyNet相比，计算复杂度降低了20倍。

- 目标检测

基于事件的目标检测旨在从事件流中分类和检测目标边界框

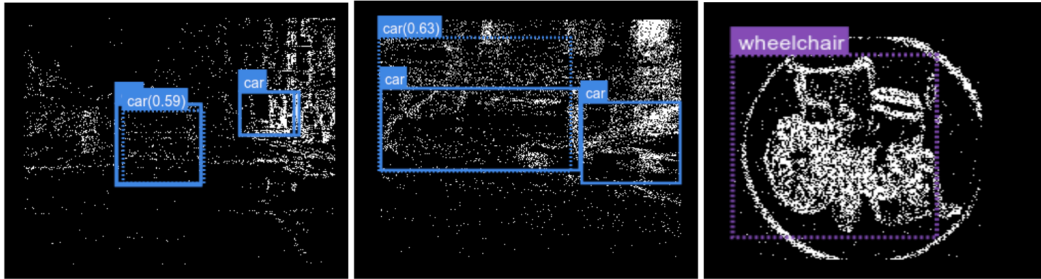
• 数据集

1. **神经形态N - Caltech101数据集**: 包含用真实事件相机记录的事件流, 代表8246个事件序列中的101个对象类别, 每个事件序列长300 毫秒。
2. **Gen1数据集**: 以城市环境中的汽车场景为目标, 并标注了行人和汽车, 包含228123个汽车边界框和27658个行人边界框。

• 实验

使用mAP得分以及每个事件的计算复杂度, 与同步和异步方法进行了比较, 并进行定量与定性分析。在定性结果中, AEGNN的预测显示为虚线, 标签为实线。

Methods	Representation	Async.	N-Caltech101		Gen1	
			mAP \uparrow	MFLOP/ev \downarrow	mAP \uparrow	MFLOP/ev \downarrow
YOLE [7]	Event-Histogram	✓	0.398	3682	-	-
Asynet [35]	Event-Histogram	✓	0.643	200	0.129	205
RED [42]	Event-Volume	✗	-	-	0.40	4712
NVS-S [31]	Graph	✓	0.346*	7.8	0.086*	7.8
Ours	Graph	✓	0.595	0.37	0.163	0.39



AEGNN比NVS - S提高了7.7 %的mAP, 同时减少了21倍的计算量。因为NVS-S使用了标准的图卷积, 有一个仅限于其直接邻域的感受野, 影响了检测性能。

与RED相比, AEGNN取得了较低的准确率, 但计算量明显优于RED。因为RED使用了一个密集的、同步的递归网络, 不能进行逐事件的处理。

与AsyNet相比, 在性能相当的情况下, AEGNN实现每个事件减少520 - 540倍的计算量。

5、总结

- 提出了基于事件的异步图神经网络AEGNN，用于处理基于事件的视觉数据。
- AEGNN将事件表示为**时空图**，并使用**高效的更新规则**来限制网络的更新计算，实现高效的事件处理。
- 在目标识别和检测任务中进行实验，AEGNN取得**显著的性能提升**，不仅在**准确率**方面优于现有方法，而且在**计算效率**方面也具有明显的优势。
- 未来的研究方向包括可以将AEGNN应用于**其他视觉处理任务**，以及实现将其在**专用硬件**上进一步提高性能。

6、启发

- **事件数据的表示**：将事件数据表示为**时空图**，可以获得事件之间的时空关系。这种表示保留了事件数据的稀疏性和高时间分辨率，允许高效处理并减少计算量。
- **异步处理**：为基于事件的视觉任务引入异步处理范式。通过**更新仅受新事件影响的节点的网络**，可以实现低延迟事件处理，并减少计算冗余。这种异步处理策略可以应用于各种GNN架构。
- **分层学习**：通过利用GNN，可以在同步输入上进行训练，并在测试期间以异步模式进行，实现分层学习，使其可扩展到更复杂的任务。