

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр

(назва освітнього ступеня)

на тему: Розроблення автоматизованої системи прогнозування траєкторії
руху транспортного засобу для різних сценаріїв автономного водіння

Виконав(ла): студент(ка) IV курсу, групи КТ-41

спеціальності 151 – Автоматизація та комп'ютерно-

інтегровані технології

(шифр і назва спеціальності)

Гарасимів В.М.

Керівник

(підпис)

Золотий Р.З.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Чихіра І.В.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Голотенко О.С.

(прізвище та ініціали)

Рецензент

(підпис)

Сисак І.М.

(прізвище та ініціали)

Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності,			
основи охорони праці			

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	20.05.2026	
2.	Підбір джерел по темі роботи	22.05.2026-24.05.2026	
3.	Опрацювання публікацій та збір даних по темі роботи	24.05.2026-27.05.2026	
4.	Виконання роботи згідно мети	27.05.2026-29.05.2026	
5.	Оформлення першого та другого розділів	29.05.2026-31.05.2026	
6.	Оформлення третього розділу	31.05.2026-11.06.2026	
7.	Оформлення четвертого розділу	11.06.2026-13.06.2026	
8.	Підготовка презентації	14.06.2026	
9.	Підготовка доповіді	15.06.2026	
10.	Оформлення кваліфікаційної роботи	15.06.2026	
11.	Нормоконтроль	16.06.2026	
12.	Перевірка на плагіат	16.06.2026	
13.	Попередній захист кваліфікаційної роботи	16.06.2026	
14.	Захист кваліфікаційної роботи	24.06.2026	

Студент

(підпис)

Гарасимів В. М.

(прізвище та ініціали)

Керівник роботи

(підпис)

Золотий Р.З.

(прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота бакалавра складається з пояснювальної записки та графічної частини (ілюстративний матеріал – слайди).

Об'єм графічної частини кваліфікаційної роботи становить _____.

Об'єм пояснювальної записки складає ____ друкованих сторінок формату А4.

В роботі використано ____ літературних джерел.

У першій частині роботи, після рекапітуляції існуючих методів, що не пов'язані з навчанням, у рамках прогнозування траєкторій вивчено різні існуючі представлення та підходи до прогнозування руху на основі навчання. Потім запропоновано вирішити проблему прогнозування траєкторії за допомогою теплових карт ймовірностей для полегшення мультимодальності. розроблено три різні способи створення цих теплових карт та оцінено їх одна відносно одної та існуючого сучасного рівня. Також запропоновано повний метод вибірки для вилучення фактичних траєкторій з цих теплових карт та вивчено переваги та недоліки цих методів теплових карт порівняно з іншими поширеними фреймворками. У наступному розділі розглянуто багатоагентне прогнозування, а точніше, прогнозування на узгоджених виходах на рівні сцени, для цих типів моделей теплових карт за допомогою методу самопланування та вивченні пост-обробки.

Ключові слова: АВТОНОМНЕ ВОДІННЯ, АВТОМАТИЗОВАНА СИСТЕМА, ПРОГНОЗУВАННЯ, МАШИННЕ НАВЧАННЯ.

ЗМІСТ

ВСТУП	5
1. АНАЛІТИЧНА ЧАСТИНА	6
1.1 Огляд проблеми та постановка завдання.....	6
1.2 Огляд систем прогнозування траєкторії руху без навчання.....	9
2. ПРОЄКТНА ЧАСТИНА	20
2.1. Машинне навчання для прогнозування траєкторії.....	20
2.2 Розробка моделі прогнозування траєкторії за допомогою теплових карт	38
3 СПЕЦІАЛЬНА ЧАСТИНА.....	54
3.1 Графова нейронна мережа для вхідних даних HD-Map.....	54
3.2 Генерація теплових карт за допомогою растрових зображень на рівні смуги руху	55
4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ	63
4.1 Ергономічні проблеми безпеки життєдіяльності при роботі за	63
комп'ютером.....	63
4.2 Організація безпечної роботи електроустаткування задіяного при роботі системи електронного навчання	66
ВИСНОВКИ.....	69
ПЕРЕЛІК ПОСИЛАНЬ	70

ВСТУП

Нещодавні досягнення в методах машинного навчання дозволили досягти величезного прогресу в автономному водінні, зокрема завдяки етапу сприйняття завдяки глибокому навчанню та глибоким нейронним мережам у поєднанні з прогресом в сенсорах, картографуванні та методах пропріоцепції. Тому зараз увага зміщується до наступних кроків у автономному конвеєрі, де прогнозування відіграє важливу роль. Після того, як навколишні дорожні агенти виявлені та відстежені, система керування повинна передбачити їхню майбутню траєкторію та відповідно спланувати курс без зіткнень.

Це прогнозування траєкторії має відповідати кільком вимогам. По-перше, воно, очевидно, має бути точним і достовірним, щоб його результати можна було надійно використовувати в наступних процесах. Майбутнє може пропонувати безліч можливостей, з яких не завжди можливо зробити висновок, спираючись виключно на минулі історичні дані. Тому прогноз має бути мультимодальним, шляхом прогнозування кількох одночасних ймовірних майбутніх подій. Оскільки прогноз має бути зроблений для всіх навколишніх агентів, а поведінка цих агентів значною мірою залежить від їхньої взаємодії один з одним, модель повинна враховувати ці взаємодії, а її мультимодальні прогнози повинні бути узгоджені один з одним. Нарешті, для безпеки та надійності прогноз траєкторії має бути легким для інтерпретації, ретельно оціненим, здатним забезпечувати достовірні оцінки та розробленим з урахуванням його кінцевого використання в конвеєрі.

1. АНАЛІТИЧНА ЧАСТИНА

1.1 Огляд проблеми та постановка завдання

Автономний транспортний засіб має спільну архітектуру з більшістю автономних роботизованих систем, яку можна підсумувати у 3 основних кроках, проілюстрованих на рис.1.1. Цей конвеєр працює в замкнутому циклі, і кожен його крок повторюється один за одним з певною частотою, щоб система могла стверджувати свою продуктивність та адаптуватися до змінюваного середовища.

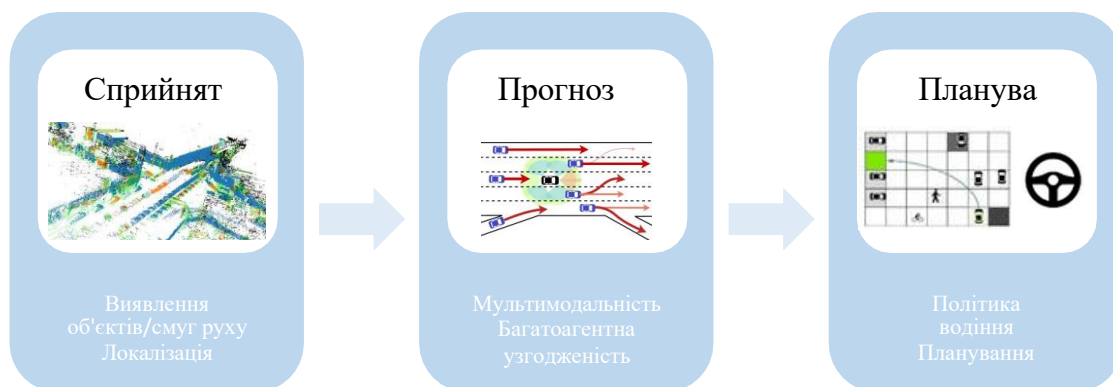


Рисунок 1.1 - Конвеєр автономного водіння

Система повинна бути обізнаною зі своїм оточенням. Для цього вона оснащена набором датчиків (камери, радары, лідары тощо). Дані з цих датчиків обробляються за допомогою алгоритмів, щоб витягти з нього значущу семантичну інформацію, таку як оцінка глибини, проїзна зона, розмітка смуг руху, об'єкти, перешкоди, інші рухомі агенти, а також локалізація. Історично ці методи сприйняття спочатку виконувалися з використанням ручних функцій, але нещодавно методи глибокого навчання показали значний внесок в удосконалення в обробці зображень і поступово замінили кожен крок у стеку сприйняття, будь то виявлення об'єктів, оцінка глибини, семантична сегментація.

Цей перший крок семантичної інформації потім обробляється з використанням часової інформації для виконання завдань міркування вищого рівня, таких як локалізація або відстеження об'єктів.

Після того, як автономний транспортний засіб (АВ) створив зображення обстановки, необхідно міркувати про майбутню еволюцію цієї сцени. Якщо помістити її в рамки, керуючий агент повинен мати змогу оцінити свій майбутній стан, враховуючи свій поточний стан та обрану дію. Однак, хоча більшу частину переходу станів можна легко вивести з кінематики, деякі частини можуть розвиватися незалежно від дій агента, дотримуючись власної логіки та мети. Більш конкретно, інші виявлені дорожні агенти рухатимуться в майбутньому, і автономний транспортний засіб повинен оцінити їхнє майбутнє місцезнаходження, щоб уникнути зіткнення з ними та мати змогу планувати власну траєкторію руху в довгостроковій перспективі.

Враховуючи поточне представлення сцени та її прогнозовану майбутню еволюцію, аудіовізуальний засіб повинен прийняти план, щоб досягти свого головного завдання: дістатися до місця призначення. Однак цієї мети необхідно досягти, дотримуючись деяких підзавдань, таких як дотримання правил дорожнього руху або уникнення зіткнень. Таким чином, стек планування враховує спостережуваний стан та його прогнозоване майбутнє для побудови безпечної та ефективної послідовності дій, яку він оцінив як найкращу відповідно до визначених критеріїв. Перша дія цієї послідовності потім буде застосована через стек керування до входних даних керма та дросельної заслінки.

Прогнозування траєкторії є ключовим етапом автономного засобу, і тому є критично важливим для автономної поведінки. Воно відповідає, мабуть, за його найважливіший критерій: безпеку, оскільки є наріжним каменем для запобігання зіткненням та прогнозування майбутнього, що є важливим для плавної, енергоефективної, комфортної та надійної поведінки.

Хоча деякі підходи, що стосуються водіння в єдиній наскрізній диференційованій моделі для навчання поведінки прямого водіння від входних даних сенсорів. Ці методи жертвують інтерпретованістю та теоретичними

гарантіями заради вигоди, яка здебільшого полягає у спрощеній архітектурі та уніфікованому навчанні міткам, які важко та дорого збирати. Крім того, навчання цим наскрізним моделям є громіздким, оскільки вони вимагають від моделі розуміння значної кількості причинно-наслідкових зв'язків між дуже високорівневими елементами керування та дуже низькорівневими вхідними даними датчиків, а також потребують складних методів, щоб уникнути зсуву розподілу.

Галузь досліджень прогнозування все ще підходить до проблеми прогнозування комплексно, але розділяє завдання всередині архітектури моделі. Починаючи безпосередньо з даних давачів та обробляючи сприйняття й відстеження, перш ніж робити висновки про майбутнє прогнозування. Ми вирішили ізолювати проблему прогнозування траєкторії від ідеального відстеження сприйняття, як це передбачено в більшості наборів даних про траєкторії та бенчмарків.

Ми поставили нашу проблему в проміжному представленні, де сусідні об'єкти були виявлені та відстежені для вилучення їхніх часових послідовностей координат, а навколишня карта була локалізована та виведена з даних давачів. Як показано на рис.1.2, Модель прогнозування приймає локальну карту та траєкторії навколишніх/цільових агентів як вхідні дані, щоб передбачити майбутню послідовність координат з регулярною частотою та до заздалегідь визначеного часового горизонту (наприклад, 5 секунд при 10 Гц) для одного або всіх агентів, що знаходяться поблизу. Як ми покажемо далі в цій роботі, ці вхідні дані можна використовувати або в їх необробленій векторизованій формі, або шляхом растрезації для отримання зображень з висоти пташиного польоту.

Модель прогнозування траєкторії повинна враховувати широкий спектр факторів. Її конструкція може змінюватися залежно від типу агента, якого потрібно передбачити. Після кодування часової інформації se—Щоб зрозуміти послідовності кожного агента, він повинен бути здатним розуміти можливі взаємодії між кожним з них, а також правила дорожнього руху, які можуть зумовлювати ці взаємодії або додавати додаткові обмеження. Карта

навколишнього середовища з її дорожніми смугами, зоною проїзду та перехрестями також має дуже важливий вплив на траєкторію руху транспортних засобів.

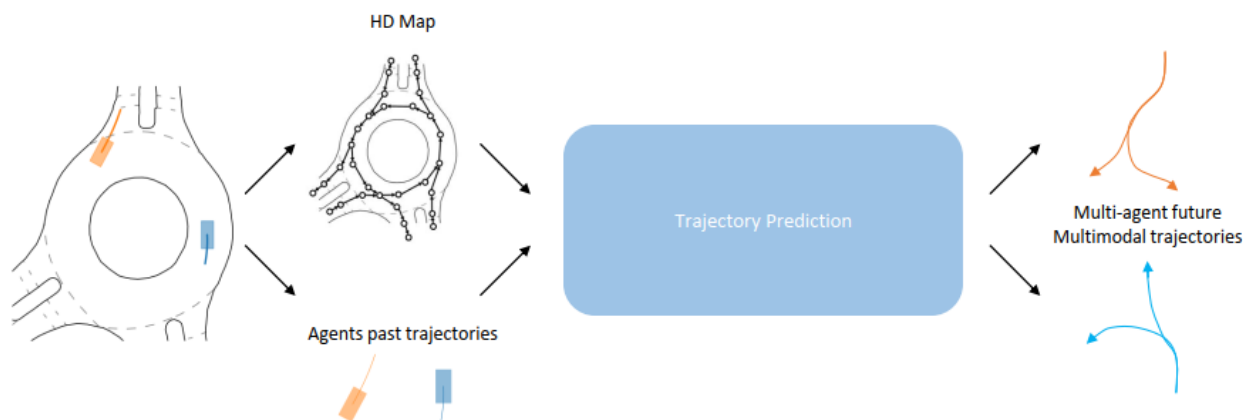


Рисунок 1.2 - Вхідні та вихідні дані для прогнозування траєкторії

Окрім первинної продуктивності, конвеєр прогнозування руху повинен відповідати різним викликам. Щоб належним чином охопити різні ймовірні майбутні сценарії, прогнозування має бути мультимодальним, таким чином, щоб кожне можливе майбутнє було представлено в певній модальності. Більше того, необхідно прогнозувати дії кількох агентів одночасно, достатньо швидко, щоб їх можна було реалізувати в системі реального часу, і таким чином, щоб прогноз на рівні сцени був узгодженим між кожним агентом. Ці численні одночасні обмеження також викликають питання про те, як нам слід оцінювати ці моделі прогнозування, і чи повинні вони також бути зобов'язані надавати оцінки невизначеності для своїх оцінок.

1.2 Огляд систем прогнозування траєкторії руху без навчання

Кінематичні моделі для оцінки руху.

Фізичні моделі.

Найпростіше, прогнозування майбутнього положення транспортного засобу можна просто зробити, використовуючи його швидкість v та напрямок

руху θ , та апроксимуючи їх як константи найближчим часом у моделі постійної швидкості (CV):

$$\mathbf{p}_{t+\Delta t} = \mathbf{p}_t + v\Delta t \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad (1.1)$$

Як наближення першого порядку, ця формула може бути точною у дуже коротких часових горизонтах, але в довгостроковій перспективі вона не відображає справжньої фізичної поведінки об'єкта, який може прискорюватися або повертатися. Тому її можна уточнити за допомогою моделі постійної швидкості повороту ω та прискорення a (CTRA):

$$\begin{aligned} \mathbf{p}_{t+\Delta t} &= \mathbf{p}_t + v_t \Delta t \begin{pmatrix} \cos \theta_t \\ \sin \theta_t \end{pmatrix} \\ v_{t+\Delta t} &= v_t + a\Delta t \\ \theta_{t+\Delta t} &= \theta_t + \omega\Delta t \end{aligned} \quad (1.2)$$

Оскільки не всі ці величини можуть бути отримані з модуля Сприйняття, використовується фільтр Калмана можна використовувати для отримання цих значень з минулої послідовності позицій. Ці фільтри зазвичай потребують ручно налаштованих гіперпараметрів для початкових значень невизначеності, проте останні підходи намагаються вивчати ці значення шляхом зворотного поширення помилок прогнозування фільтра. Ці моделі потім суттєво забезпечують фазу прогнозування на основі їхнього оціненого стану.

Фізична модель, яка найчастіше використовується у фільтрах Калмана для автономного водіння, — це модель велосипеда. Що дещо відрізняється від СТРА, якщо припустити постійне значення керма δ , як показано на рис. 1.3, та обчислення швидкості повороту на основі рульового керування, колісної бази L та швидкості:

$$\omega = v * \tan(\delta) / L \quad (1.3)$$

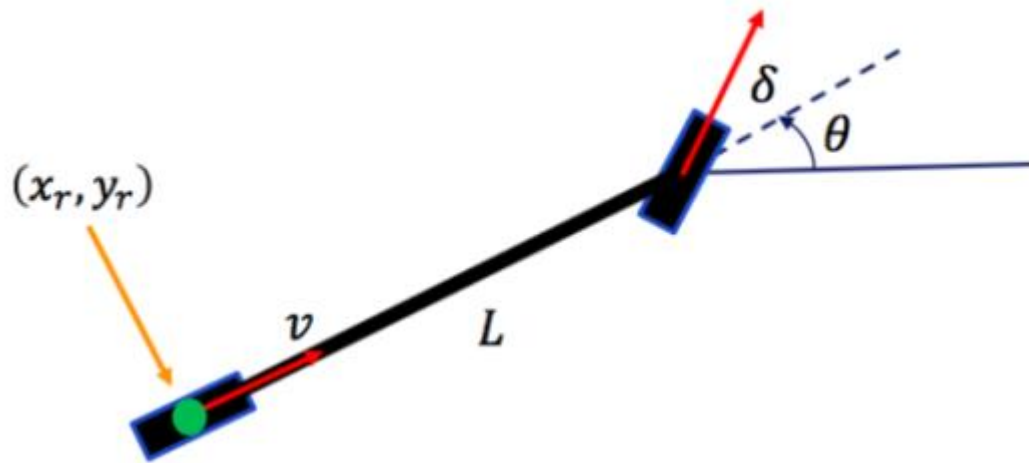


Рисунок 1.3 - Модель велосипеда з постійним кутом повороту керма δ та колісною базою L .

Гладкість та екстраполяція часових рядів.

Завдання прогнозування траєкторії можна розглядати як задачу екстраполяції, де майбутні точки необхідно передбачити на основі історичних даних часового ряду, які підганяються поліномом Чебишева до траєкторії ху-позицій або підганянням 5-градусним поліномом для поперечної складової та 3-градусним поліномом для поздовжньої складової у координатах Френе, як показано на рис. 1.4, також підганяє клотоїди до траєкторій, заснованих на постійно змінній кривизні.

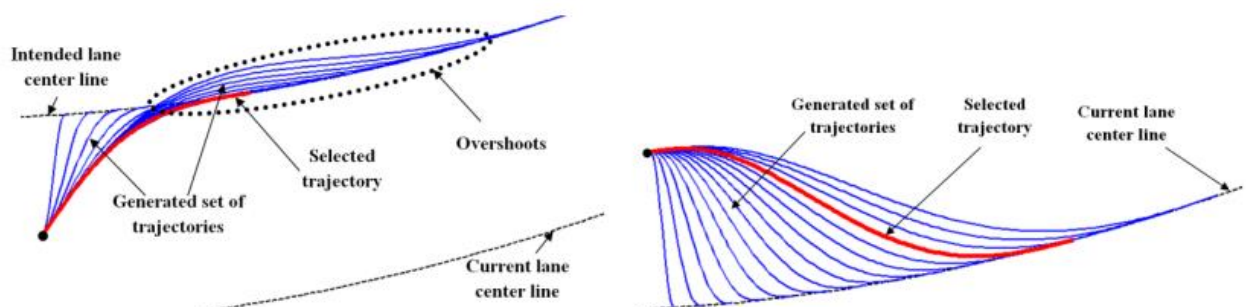


Рисунок 1.4 - Кандидати в поліноміальну регресію з обмежень на початкову точку (координати та похідна) та кінцеву точку (похідні).

Стохастичні моделі.

Приховані марковські моделі.

Приховані марковські моделі (ПММ) представляють траєкторію автомобіля, як послідовність часткових спостережень x , отриманих з прихованої змінної z . У формулюванні ланцюга Маркова, як показано на рис. 1.6 розподіл ймовірностей z_t на кожному кроці часу t залежить лише від попередньої змінної z_{t-1} , а потім спостережуване положення x_t обумовлене лише z_t . Таке моделювання робить ПММ дуже пристосованими для послідовних даних.

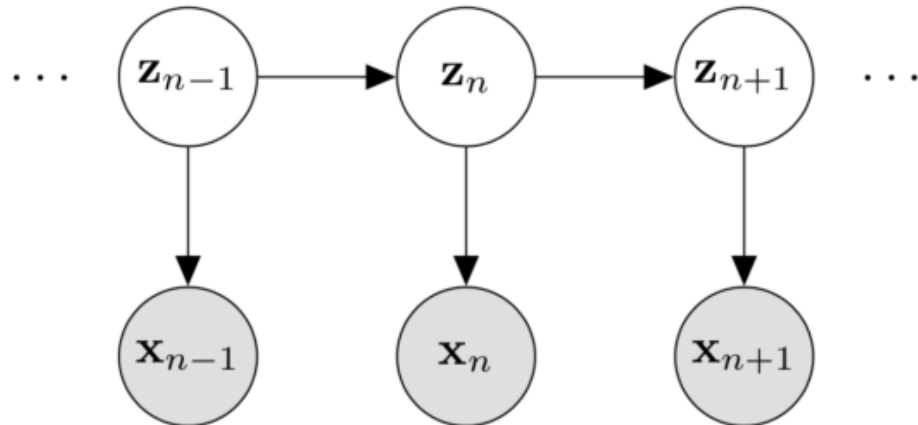


Рисунок 1.5 - Ланцюг Маркова з прихованою змінною z та спостереженням x .

ПММ моделює спільну ймовірність z та x :

$$p(x, z) = \prod_t^T p(x_t | z_t) p(z_t | z_{t-1}) \quad (1.4)$$

де $p(x_t | z_t)$ – ймовірність викиду;

$p(z_t | z_{t-1})$ – ймовірність переходу.

Таким чином, це генеративна модель, яка представляє базовий розподіл ймовірностей.

Моделювання Монте-Карло.

Деякі автори використовувати вибірку Монте-Карло для моделювання можливих майбутніх шляхів з послідовності рівномірних елементів керування. Такі дії, як рульове керування та прискорення, ефективно вибірково вибираються

з урахуванням обмежень крутного моменту та тертя без будь-якої вимоги до лінійної фізичної моделі, одночасно перевіряючи зіткнення, щоб визначити найімовірніший шлях, що уникає зіткнення.

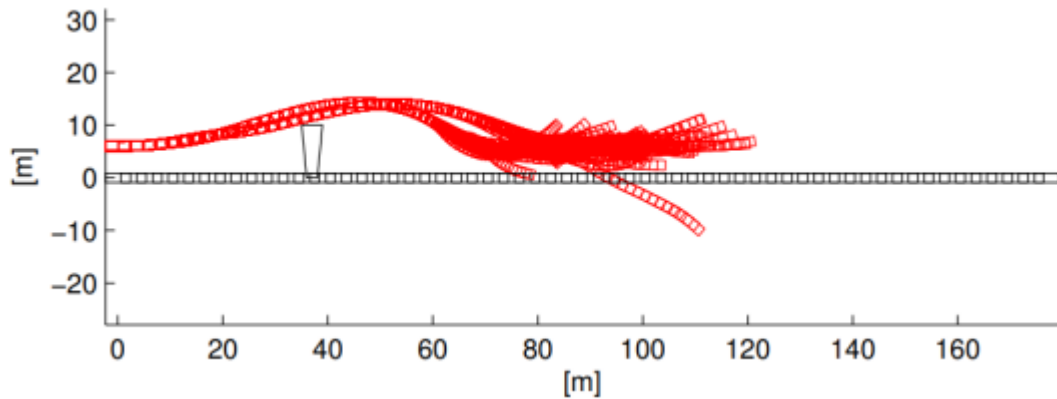


Рисунок 1.6 - Вибірка методом Монте-Карло для уникнення зіткнень.

Умовні випадкові поля

Умовні випадкові поля відрізняються від моделей умовних випадкових полів тим, що вони є дискримінативними моделями лише для умовної ймовірності $p(z|x)$. Їх також можна застосовувати для прогнозування прихованих подій, таких як гальмування або обгін.

Баєсівські мережі

Баєсівську мережу використовують для оцінки ймовірності кожної можливої смуги руху, використовуючи топологію перехрестя за допомогою прямого графіка, що представляє залежності між різними змінними. Порівняно з ПММ, баєсівські мережі менше зосереджені на часовій залежності рішення та багато іншого щодо взаємозв'язків між різними змінними, такими як обраний маневр, сигнал повороту тощо.

Моделі взаємодії

Соціальні сили

Зв'язок з прогнозуванням траєкторії руху пішоходів.

Незважаючи на те, що за кінцевою метою це досить схожі завдання, прогнозування руху пішоходів та автомобілів значно відрізняється за методологією та характером. Пішоходи здебільшого розвиваються в

необмеженому середовищі, на їхній рух головним чином впливають інші пішоходи в натовпі. Автомобілі ж дотримуються дуже суворих правил дорожнього руху, де інші агенти залишаються важливими для запобігання зіткненням та зайнятості простору, але не є основними факторами руху цільового автомобіля.

Можна прогнозувати рух пішоходів шляхом застосування соціальних сил з боку інших пішоходів та перешкод, як показано на рис. 1.7. Подібні методи були застосовані через багаточастинкові системи та гідродинаміку до дорожнього руху, але вони залишаються на макроскопічному рівні та недостатні для ефективного та точного прогнозування рухів окремих автомобілів.

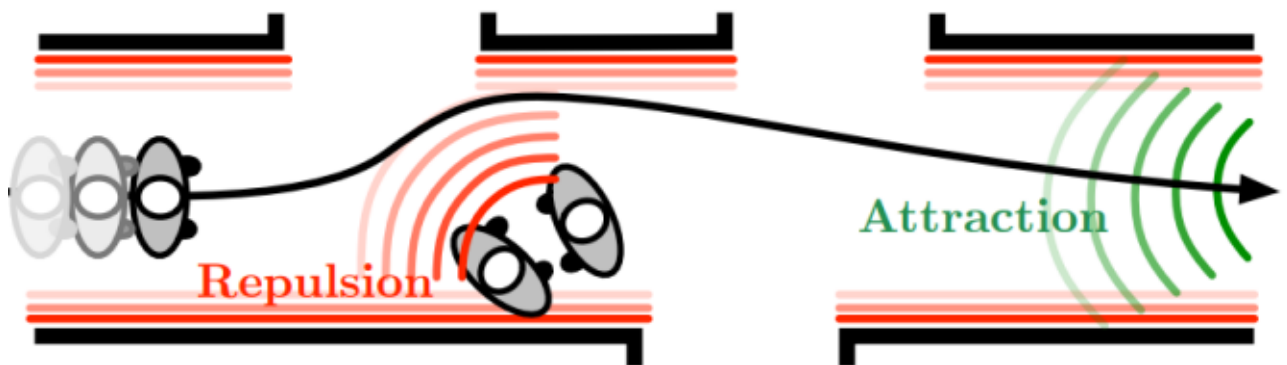


Рисунок 1.7 - Соціальні сили для прогнозування руху пішоходів.

2.3.2 Моделі драйверів

Зв'язок із симуляцією Оскільки моделювання реалістичної поведінки водія зводиться до створення набору можливих траєкторій, якими агент водіння міг би слідувати в реальному житті, симуляції можна розглядати як вибір однієї з можливих траєкторій для агента та її використання для внутрішнього планування. Знову ж таки, під час моделювання пункт призначення агента має бути відомий, але більшість простих моделей моделювання все одно не враховують пункти призначення, зосереджуючись на взаємодії автомобілів.

Найпоширенішою моделлю поведінки є Модель інтелектуального водія (МІВ), яка безпосередньо керує прискоренням автомобіля згідно з наступним рівнянням:

$$\dot{v} = a[1 - (v/v_0)^\delta - (s^*(v, \Delta v)/s)^2] \quad (1.5)$$

Якщо $\Delta v = v - vl$ — швидкість наближення до транспортного засобу, що веде, Δ — показник степеня, який зазвичай дорівнює 4, і де v порівнюється з бажаною швидкістю v_0 , а s — з бажаною відстанню s^* :

$$s^*(v, \delta v) = s_0 + \max(0, vT + \frac{v\Delta v}{2\sqrt{ab}}) \quad (1.6)$$

де T — часовий проміжок до транспортного засобу, що веде;

s_0 — мінімальна відстань, a — максимальне прискорення;

b - комфортне гальмівне уповільнення;

$s_0 + vT$ — тут рівноважний член;

$\frac{v\Delta v}{2\sqrt{ab}}$ — динамічний член, що реалізує «інтелектуальну» стратегію гальмування, де кінематичне прискорення, необхідне для безпеки, саморегулюється в бік комфортного уповільнення.

Можна адаптувати цю модель водія від симуляції до прогнозування поведінки за визначенням. Необхідне визначення попередніх кластерів поведінки водіїв на основі даних та оцінка їх апостеріорних ймовірностей на основі цих гіпотез, враховуючи минулу траєкторію.

Ця модель, як і більшість інших, здебільшого транскрибує просту модель руху автомобіля для симуляції дорожнього руху та уникнення зіткнень, і охоплює лише певні сценарії, такі як автомагістралі. Вона більш адаптивна до навколишнього середовища, ніж фізичні моделі, але жертвує цією адаптивністю заради реалізму.

Теоретико-ігрові підходи

Також можна розглянути прогнозування/планування майбутньої послідовності дій для кількох рушійних агентів як розв'язання гри, де кожен агент хоче максимізувати власну винагороду, іноді також враховуючи винагороду інших, як показано на рис. 1.8. Розв'язок потім отримується шляхом розв'язання інтерактивної гри за допомогою дерева гри або гра, де провідний транспортний засіб спочатку обирає дію, а потім наступний транспортний засіб і так далі, щоб досягти рівноваги Неша, де жоден гравець не може виграти, змінюючи лише свої дії.

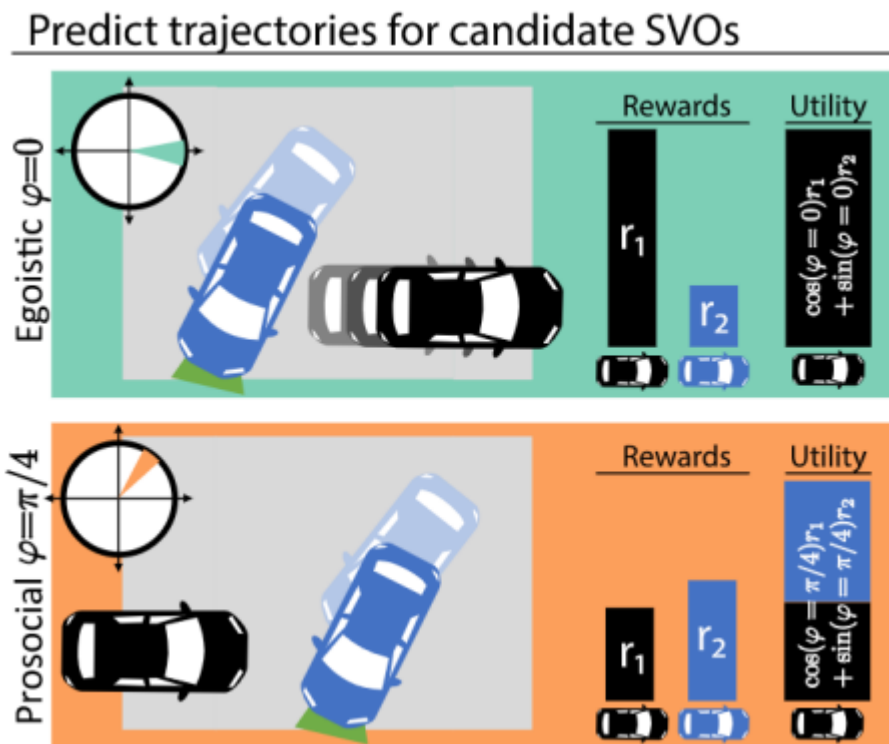


Рисунок 1.8 - Різні моделі поведінки, що впливають з рішень теорії ігор, залежно від того, чи враховує агент лише власну винагороду (егоїстичний підхід), чи також винагороду іншого (просоціальний підхід).

Ієрархічне прогнозування.

Для представлення різних деталей реалістичної траєкторії рух можна розбити на проміжні кроки, які можуть бути або кроками планування до пункту призначення, або заздалегідь визначеними операціями, отриманими на основі реальної людської поведінки.

Пошук по графу.

Зв'язок з плануванням завдання прогнозування траєкторії має багато спільного з плануванням. Дійсно, прогнозування майбутнього одного навколишнього автомобіля майже те саме, що й планування цього автомобіля як керуючого агента. Основні відмінності полягають у тому, що, хоча керуючий агент знає свій пункт призначення для планування, ця мета прибуття повинна бути оцінена під час прогнозування або не дотримуватися жодного припущення щодо конкретної цілі та охоплювати всі можливі пункти призначення.

Графік може представляти неперервний простір, як у Hybrid A*, що розширює евристичний метод пошуку графів A* до пошуку безперервних станів шляхом зберігання безперервних станів у досягнутих дискретних комірках або швидкого дослідження випадкових дерев (RRT), яке ітеративно конструює власні вузли в просторі дій, гарантуючи кінематичну доцільність. Оскільки навколишня дорожня карта зазвичай доступна у вигляді графіка, майбутній шлях автомобіля також можна оцінити, досліджуючи досяжні вузли на карті. Профіль швидкості можна врахувати пізніше або під час пошуку графа за допомогою просторово-часового графіка, як показано на рис. 1.9.

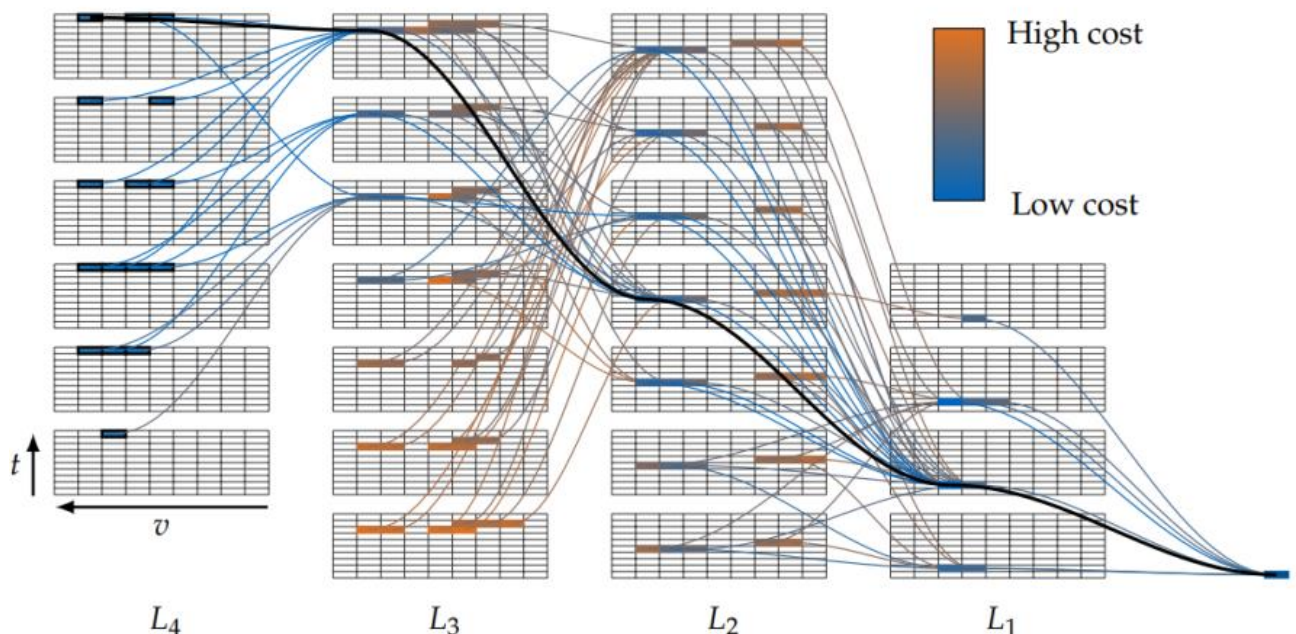


Рисунок 1.9 - Просторово-часовий пошук графів для планування траєкторії.

Прогнозування на основі маневру.

Щоб зменшити складність простору прогнозування, у деяких роботах пропонується обмежити можливості дискретним набором маневрів.

Якорі.

Минуле траєкторії можна використовувати для знаходження її найближчого сусіда в кластері опорних траєкторій, а потім використовувати репрезентативну траєкторію кластера для оцінки майбутнього. Якщо кожен кластер надає вагу або ймовірність відносно історії траєкторії, то виведений шлях також може бути середньозваженим значенням усіх найближчих кластерів. Також можна використовувати метод «Найближчі сусіди» як базову точку для свого еталону прогнозування траєкторії руху та досліджує використання декартових або криволінійних координат як запиту. Однак ці методи залишаються досить обмеженими для обробки всіх варіацій, присутніх у русі руху або нових схемах перехресть.

Намір маневру.

Для моделювання більш абстрактних та загальних понять, у деяких роботах було вирішено моделювати окремі наміри, такі як зміна смуги руху, поступитися дорогою, поворот праворуч тощо. Також можна відстежувати кілька гіпотез маневру за допомогою фільтра Калмана та вибирає найбільш ймовірну гіпотезу апостеріорі, використовувати приховані марковські моделі для оцінки ймовірності руху транспортного засобу прямо або повороту. Однак більшість цих методів оцінки маневру спираються виключно на траєкторію руху цільового агента і тому не враховують взаємодії. Можна також додати штраф за траєкторії зіткнення під час зіставлення взаємодіючих транспортних засобів, але вимагати обмеження перевірки на зіткнення простими парами транспортних засобів або врахування експоненціальної кількості комбінацій серед усіх агентів, як показано на рис. 1.10.

Незалежно від їхньої методології, більшість цих методів, що не базуються на навчанні, не враховують реальні вимоги адаптивності та гнучкості для адаптації до різних схем доріг та взаємодій поза межами заздалегідь визначеної сфери застосування.

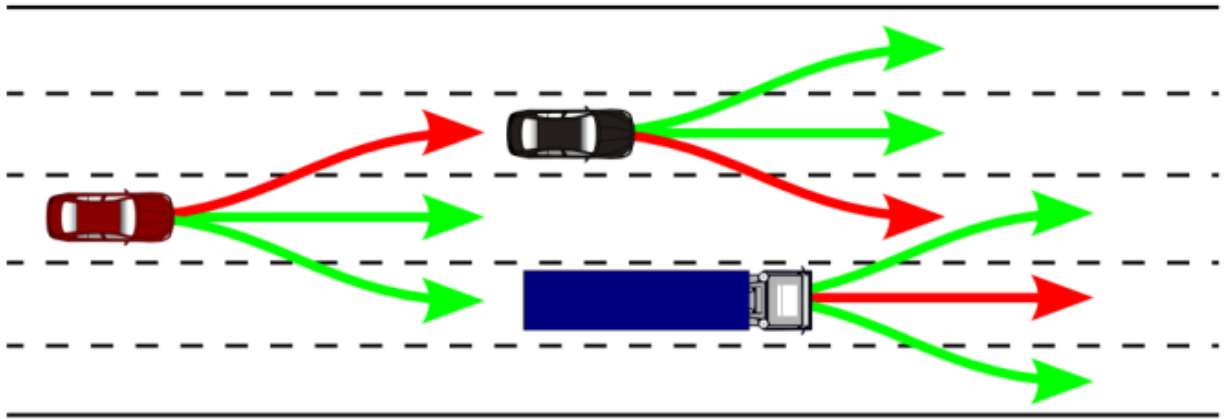


Рисунок 1.10 - Можлива комбінація маневрів між кількома агентами на автомагістралі.

2. ПРОЄКТНА ЧАСТИНА

2.1. Машинне навчання для прогнозування траєкторії

Машинне навчання, і особливо глибоке навчання, нещодавно надало моделям неперевершену здатність адаптуватися до будь-яких спостережуваних нових сценаріїв, не вдаючись до постійно зростаючого масиву попередньо визначених фіксованих моделей поведінки. Тут ми розглядаємо існуючі методи глибокого навчання, що застосовуються для прогнозування траєкторії, класифікуючи чотири основні необхідні кроки для надійного прогнозування траєкторії:

- Тимчасове кодування, оскільки основні підказки містяться в історії спостережень за цільовим транспортним засобом за останні кілька секунд
- Взаємодія агентів, оскільки нашим головним мотивом є уникнення зіткнень з іншими агентами, а багато обмежень швидкості та положення безпосередньо пов'язані з розподілом навколишніх агентів
- Контекст карти: навіть до взаємодії наше водіння значно обмежене зоною, якою можна проїхати, та місцевими правилами дорожнього руху, яких ми повинні дотримуватися
- Вихідне представлення: після того, як вся необхідна інформація була оброблена на попередньому кроці, навіть з теоретично ідеальною інформацією, ми повинні вибрати форму, в якій наша модель виведе цей прогноз, оскільки кожен вибір матиме свій власний вплив на охоплення та можливі способи використання прогнозу.

Методи машинного навчання, відмінні від нейронних мереж, такі як методи опорних векторів або випадкові лісими також застосовувалися для прогнозування майбутньої поведінки, але оскільки більшість цих застосувань також можуть бути оброблені нейронними мережами, і враховуючи їх нещодавно продемонстровану чудову продуктивність за наявності достатньої

кількості даних, ми зараз зосередимося на методах глибокого навчання, які останнім часом значно розширюються в дослідницькому співтоваристві. Серед методів, представлених у попередньому розділі, деякі також можна розглядати як методи навчання, такі як ПММ, де ймовірності переходу з одного стану в інший вивчаються з даних.

Тимчасове кодування

Багатошаровий перцептрон

Можна застосовувати багатошаровий перцептрон (MLP) до послідовності минулих станів для прогнозування параметрів цільової смуги руху та зміни смуги. Хоча цей метод здається простим і не враховує можливі адаптації для врахування часового виміру, насправді він працює відносно добре для коротких послідовностей, оскільки дозволяє моделі мати повний огляд усієї послідовності з інформацією про положення для кожного кроку. Однак, зі збільшенням розміру послідовності збільшується і розмір параметрів моделі, і мережа стає занадто повільною та нездатною узагальнювати дані. Можна також використовувати MLP для прогнозування параметрів гаусової моделі суміші, але застосовує її лише до поточного кроку часу, щоб дотримуватися припущення Маркова.

Згортканейронна мережа (CNN)

У глибокому навчанні згортковий шар – це ковзний фільтр, який витягує ознаки на основі вивченого ядра. Їхня конструкція робить їх інваріантними до трансляції та ефективними за параметрами, таблично адаптовано для зображень та часових рядів.

2D згортки.

У широкому сімействі моделей, які приймають на вході зображення, що представляє семантичну інформацію з висоти пташиного польоту, також необхідно враховувати еволюцію часу для представлення рухомих сусідніх агентів та історії цільового транспортного засобу. Цю зміну часу зазвичай можна зобразити двома способами. По-перше, пам'ятаючи, що вхідне зображення не обмежене 3 каналами RGB, для кожного часового кроку можна використовувати окремий канал, що показує положення кожного агента в цей момент часу, як

зображено на рис. 2.1. Однак це може стати досить громіздким з великою кількістю часових кроків, і швидшого часу виведення з меншою кількістю параметрів моделі можна досягти, представивши цю еволюцію у простому RGB-зображенні, де рухомі транспортні засоби показані лінією, що згасає, як на рис. 2.1, причому найдальші часові кроки є найбільш прозорими.

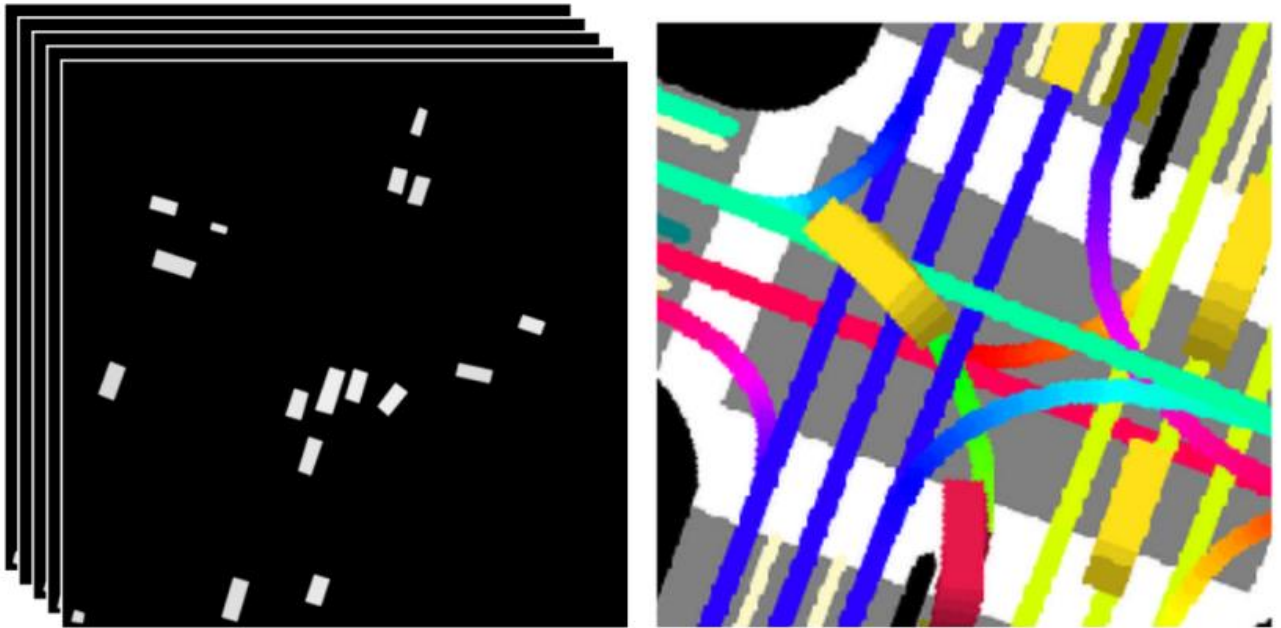


Рисунок 2.1 - Вхідні дані зображень для моделей на основі CNN з різним представленням часової еволюції.

1D згортки.

Згортки також можна використовувати у векторизованому середовищі, де вхідними даними є послідовність координат для послідовних часових кроків. У цьому випадку вхідні дані можуть бути сформовані як тензор (T, D) , де T — кількість історичних часових кроків, а D — розмір ознаки ($D = 2$ для x у щонайменше). T вимір потім можна використовувати для застосування 1D згортки як ковзного вікна вздовж часового виміру, щоб витягти характеристики вищого рівня, такі як швидкість або прискорення.

Рекурентна нейронна мережа (RNN)

RNN рекурентно застосовує ту саму комірку до кожної часової ознаки один крок за іншим. Комірка має пам'ять, яка оновлюється під час кожного

застосування, одночасно вибираючи, яку інформацію з попередньої пам'яті зберігати.

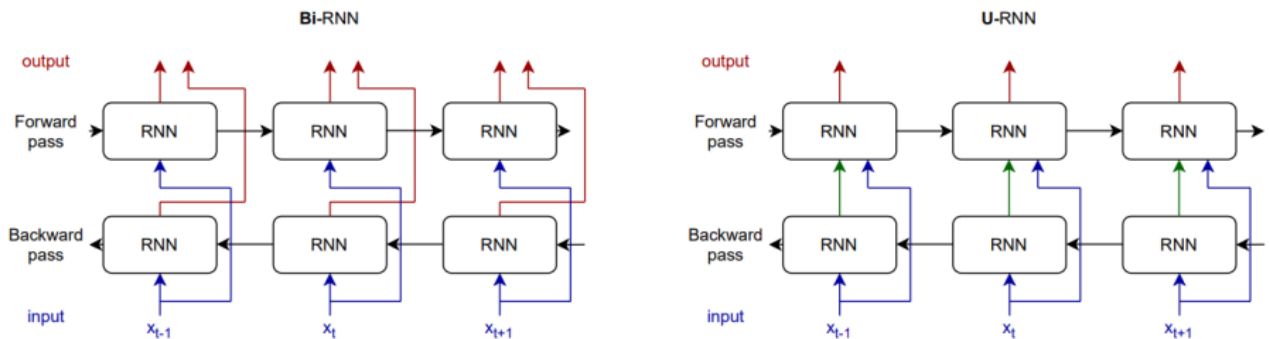


Рисунок 2.2 - Двонаправлені RNN, застосовані до послідовного входу з використанням симетричного (Bi-RNN) або асиметричного (U-RNN) розташування.

Поширеними рекурентними архітектурами комірок є GRU та та довгостроковій системі управління персоналом, які використовують залежні активації для адаптації частини пам'яті, збереженої з попередніх часових кроків, та частини, яка буде оновлена з новим, загалом допомагаючи довготривалій пам'яті.

PointNet.

Рисунок 3.3: Ілюстрація роботи та створення екземпляра хмари точок

В інших роботах послідовність позицій розглядається як група точок, які потім можна закодувати за допомогою поширених методів хмари точок, таких як PointNet. Коротше кажучи, ці моделі почергове агрегування інформації хмари точок між точками, зазвичай шляхом усереднення їхніх ознак, розподілюють агреговану інформацію шляхом об'єднання її з кожною точкою та кодування нової інформації за допомогою багаточасового перцептрона, як показано на рис. 2.3.

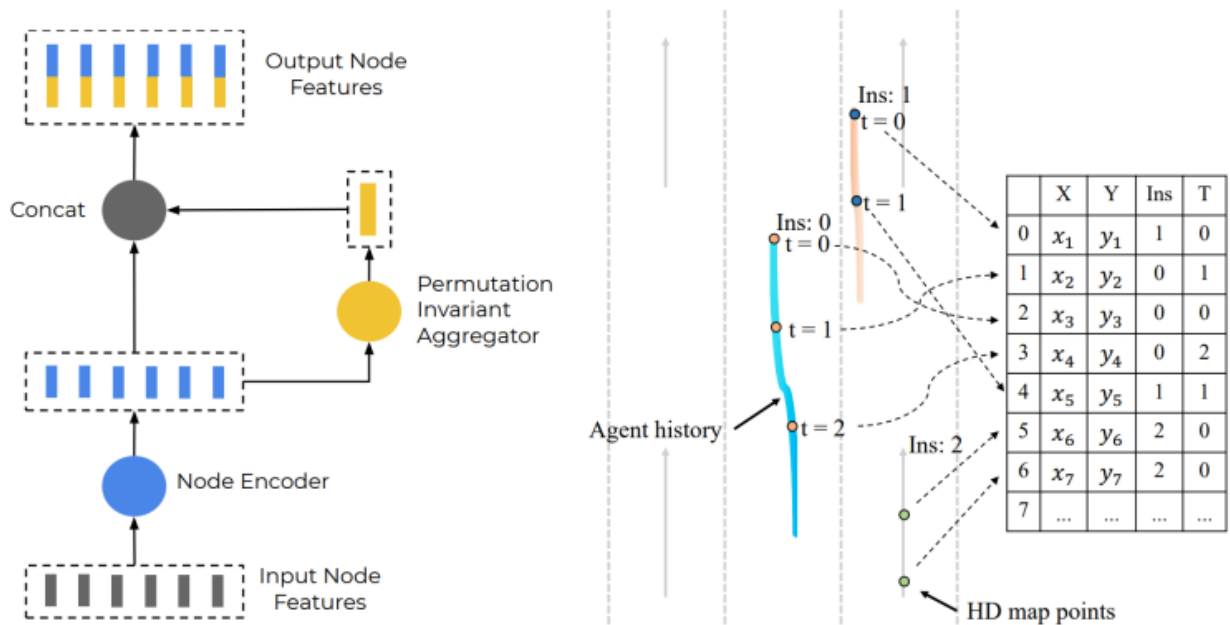


Рисунок 2.3 - Ілюстрація роботи та створення екземпляра хмари точок

Увага та Трансформери

Увага працює як словниковий запит між токенами: ми хочемо, щоб кожен токен вибирав інші важливі токени та використовував лише їхню інформацію. Для цього з токенів генеруються три тензори за допомогою лінійної проєкції: запити Q , ключі K та значення V . У випадку часового кодування всі ці тензори будуть похідними від кроків часу T і тому матимуть форму (T, d) , де d - розмір ознаки. Цей випадок, коли всі запити, ключі та значення походять з одного тензора, називається самоуважністю. Коли запити надходять з іншого джерела, ніж контекстні ключі та значення, операція називається перехресною увагою, яку ми побачимо пізніше в цьому розділі. Потім оцінка уваги α обчислюється через масштабований скалярний добуток між запитами та ключами, до якого застосовується softmax для нормалізації:

$$\alpha = \text{softmax}\left(\frac{Q \cdot K^T}{d}\right) \quad (2.1)$$

Оновлення уваги до токенів тоді є зваженою сумою значень, зважених за показником уваги:

$$X = \alpha V \quad (2.2)$$

Оскільки вся операція є інваріантною відносно перестановок, до кожного токена зазвичай додається позиційне вбудовування, щоб розрізнити їх (у нашому випадку це буде вбудовування позначки часу).

Перевагою цієї методики, спільної з CNN, є те, що, враховуючи всі можливі часові кроки залишається інваріантним до їх загальної кількості, оскільки кінцевий контекст є середньозваженим, підсумованим до одного вектора ознак, і єдина вимога полягає в тому, щоб ключі та значення мали однаковий перший вимір.

Т, що вони мають за власним призначенням. Таким чином, часовий кодер на основі уваги теоретично міг би обробляти історію будь-якого розміру.

Після демонстрації високої ефективності в галузі обробки природної мови, увага також все частіше застосовується в системі прогнозування траєкторії, замінюючи рекурентні нейронні мережі (RNN). Завдяки увазі один часовий крок може отримати доступ до контекстної інформації з будь-якого іншого часового кроку, незалежно без розбору, не страждаючи від проблем короточасної пам'яті, як у випадку з RNN, або обмежених рецептивних полів, як у випадку з CNN. Об'єднані разом у чергуванні з шарами прямої дії, ці моделі уваги називаються трансформаторами. Такі трансформатори були застосовані як до пішоходів та транспортного засобу для прогнозування траєкторії.

Взаємодія агентів.

Під час керування автомобілем ми постійно звертаємо увагу на інших учасників дорожнього руху, щоб уникнути зіткнення з ними, дотримуватися пріоритету руху або дотримуватися їхньої швидкості. Цю інтерактивну поведінку необхідно відобразити в нашій моделі прогнозування.

Виправлено найближчі та сусідні транспортні засоби.

Складність використання інформації про сусідні транспортні засоби полягає в їхній різній кількості та відносному розташуванні. Дорога може бути як порожньою, так і заповненою рухом, і більшість традиційних повністю підключених шарів глибокого навчання вимагають фіксованого розміру вхідних даних, звідси необхідність у певних правилах щодо вибору сусіднього

автомобіля. Можна використовувати k найближчих транспортних засобів відповідно до евклідової відстані або попередньо визначені зони інтересів на основі смуг руху, де транспортні засоби будуть розглядатися, якщо вони знаходяться в межах зони, як показано на рис. 2.4.

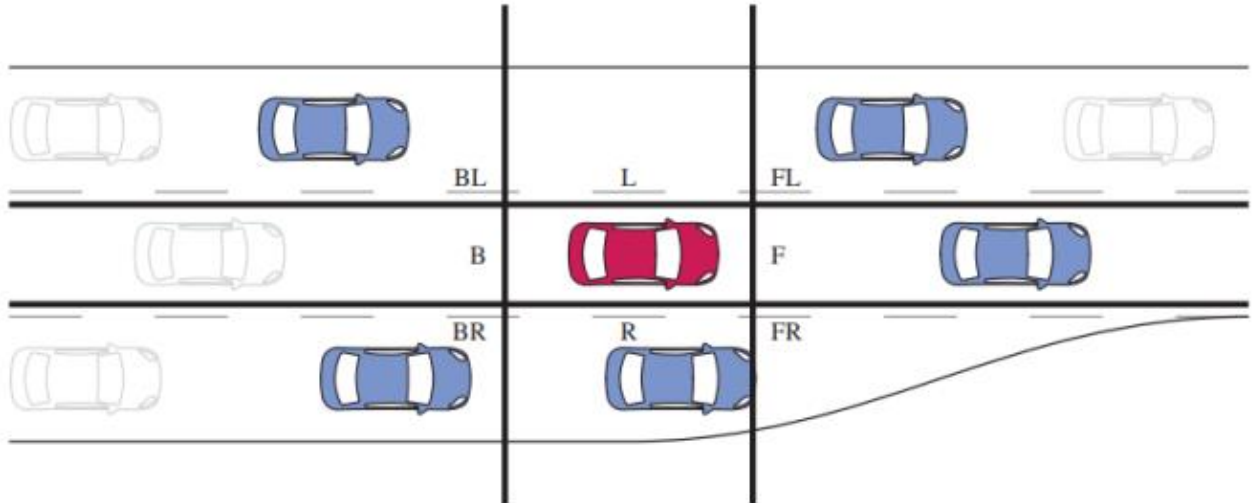


Рисунок 2.4 - Виправлено врахування сусідів на основі смуг руху.

Спільний доступагреговане представництво.

Щоб впоратися з обмеженою кількістю агентів, можна об'єднувати прихований стан від сусідів всередині просторової сітки шляхом їх підсумовування та передає результуючу об'єднану сітку цільовому агенту через конкатенацію та повністю зв'язаний шар до всіх позицій сітки.

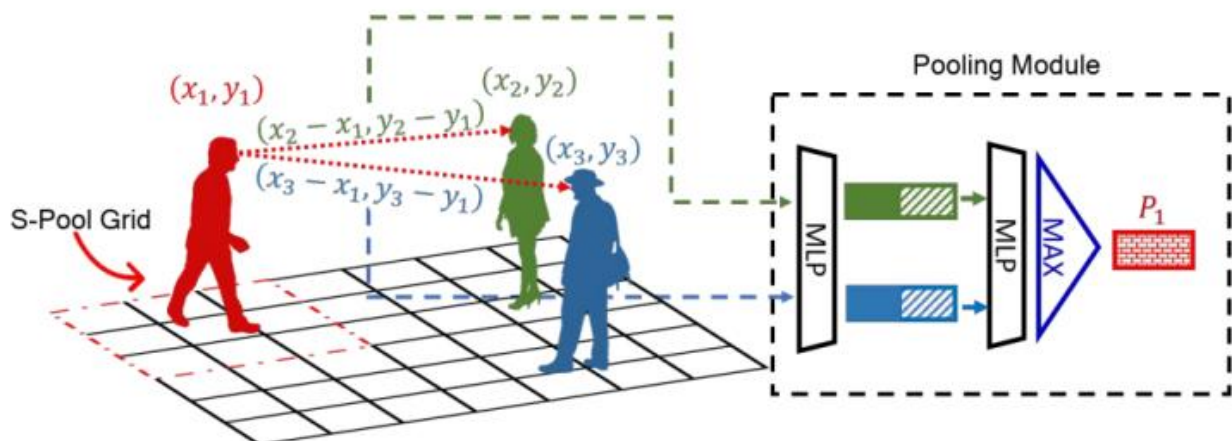


Рисунок 2.5 - Соціальне об'єднання для вивчених розріджених взаємодій

Можна використовувати подібну сітку, але застосовувати згорткові шари для збереження просторової структури, спростіть цю операцію, замінивши її на MLP, а потім на MaxPooling, як зображено на рис. 2.5 для кращої швидкості та можливостей взаємодії на відстані.

Графові нейронні мережі (GNN).

Різних агентів у сцені можна розглядати як вузли N графа (N, E) , зображеного на рис.3.6, де ребра E представляють взаємодії між учасниками дорожнього руху. Ребра можна з'єднати відповідно до відстані між агентами або повністю підключити, враховуючи відносно невелику кількість агентів.

Графова нейронна мережа (GNN) може бути використана для агрегації інформації між вузлами через ребра та оновлення характеристик вузлів на основі їхніх з'єднаних ребер. Наприклад агрегувати інформацію про зв'язані вузли за допомогою операції підсумовування або використовувати передачу повідомлень та обчислює вектор ознак для кожного ребра з об'єднаних станів сусідів через MLP, а потім застосовує MaxPooling для кожного ребра вузла.

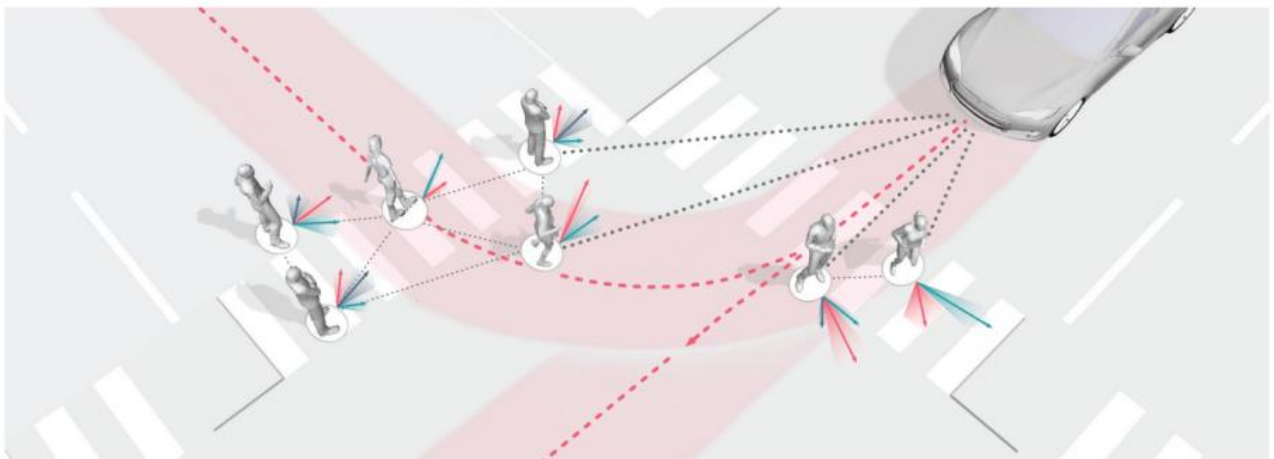


Рисунок 2.6 - Графік взаємодії між дорожніми агентами.

GNN дозволяють параметрам моделі не залежати від кількості агентів, зберігаючи при цьому специфічну для агента інформацію без глобальної агрегації, що жертвує представницькою владою.

Увага.

У рамках взаємодії увагу можна розглядати як конкретний випадок повністю зв'язаної GNN. Оператор агрегації, який підсумовує всі ознаки сусіднього агента, є операцією вивченої уваги, де запит є цільовим агентом, а ключі та значення – його сусідами. Розподілений кожному агенту в сцені, це знову зводиться до самоуваги, де кожен агент дивиться на кожного іншого агента. Як показано на рис. 2.7, увага дозволяє цільовому агенту вибрати, який сусід є важливішим для прогнозування його майбутньої траєкторії, наприклад, автомобіль прямо попереду. Для цього нанесіть шар уваги для кількох голів на основі функцій агентів, закодованих у LSTM, для моделювання їхньої взаємодії, що дозволяє взаємодіяти без локальних обмежень для змінної кількості агентів. Подальші роботи узагальнюють це шляхом об'єднання кількох уваг шарів в архітектурі Transformer

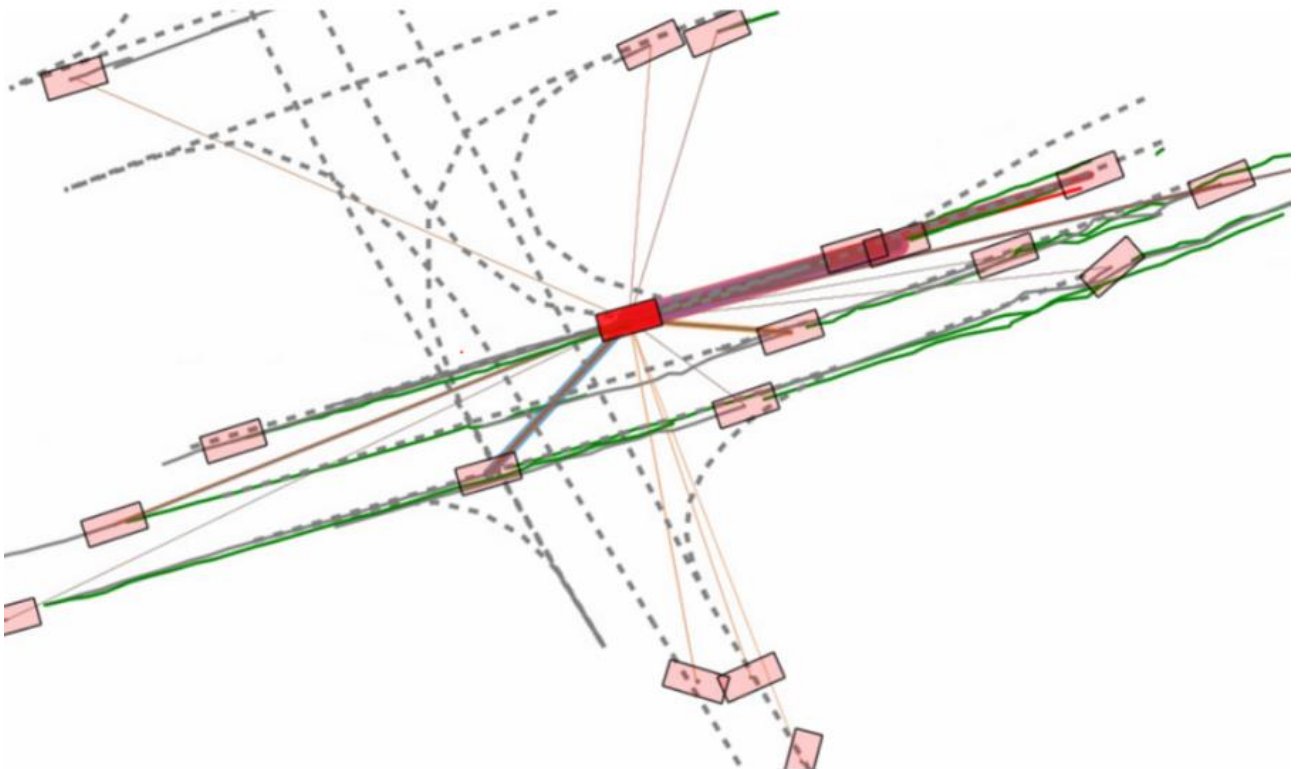


Рисунок 2.7 - Увага між агентом цільового центру та його сусідами. Увага представлена прямими лініями, з різним кольором відповідно до точки уваги та шириною, масштабованою залежно від оцінки уваги.

Розмивання межі між часовим та агентним вимірами.

У нещодавніх роботах було зроблено спробу узагальнити часове кодування та кодування взаємодії, враховуючи, що кожен агент на заданому кроці часу може бути окремим, токен швидкості взаємодіє з іншими агентами на інших кроках часу. Встановлення цього зв'язку безпосередньо збільшує можливості моделі, але якщо факторизувати час і соціальні уваги, використовуючи одне й те саме представлення, та чергувати обидві операції, то можна отримати можливість покращити складність моделі.

Контекст карти

Хоча кодування взаємодії є спільною частиною прогнозування як пішоходів, так і транспортних засобів, статичне контекстне кодування, а також фізичні обмеження, є тим, що їх відрізняє. Транспортні засоби більше обмежені навколишнім розташуванням, тоді як люди розвиваються в набагато менш обмеженому середовищі.

Хоча початкові набори даних траєкторії зазвичай були зосереджені на прямих ділянках автомагістралей з фіксованою кількістю смуг, де не було потреби в додатковій інформації на карті, нещодавні набори міських даних містили дані з перехресть, кільцевих розв'язок та вставок-ситуації, де потрібен контекст для проїзної зони.

Зображення CNN.

Простий спосіб полягає у представленні локальної карти навколо агента у вигляді зображення, так само, як GPS-пристрій відображав би її для людини. Отримане зображення являє собою вид з висоти пташиного польоту (BEV) дорожнього покриття в певному діапазоні навколо транспортного засобу, зазвичай орієнтоване відповідно до напрямку руху агента. Це представлення може містити різні форми та шари, такі як зона проїзду, центральні лінії, за бажанням забарвлені відповідно до їхнього напрямку, стоп-лінії світлофори, пішохідні переходи або супутникові знімки.

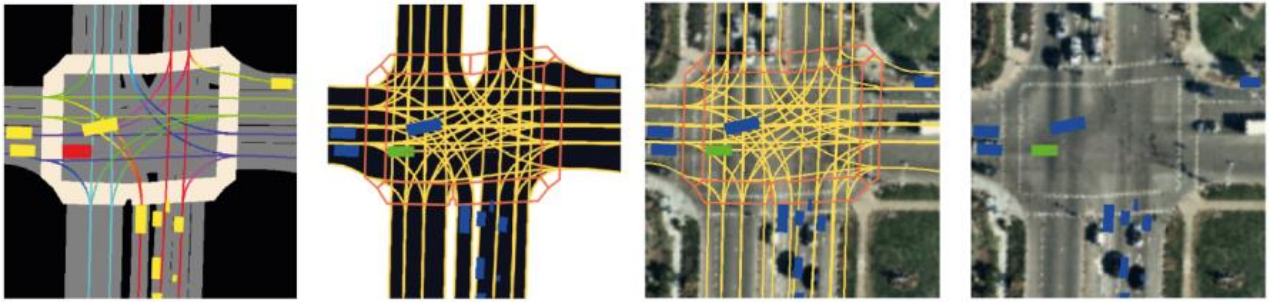


Рисунок 2.8 - Доступні представлення BEV у наборі даних Lyft

Вони дозволяють зберегти двовимірну просторову структуру карти, проте під час об'єднання з послідовними вхідними координатами вони часто сплющуються та втрачають більшу частину своєї просторової інформації.

Найновіші набори даних кодують надані карти у форматі, подібному до Lanelet2, де дорога сегментована на смуги руху. Кожна смуга руху являє собою безперервний фрагмент смуги руху певної довжини та визначається як послідовність точок. Потім смуги руху кодуються разом у граф, де кожна смуга руху, можливо, має попередника, наступника, лівого та правого сусідів. Цей граф зв'язності містить багато інформації про шляхи, до яких може дістатися транспортний засіб, але втрачається, якщо перетворити його на зображення. LaneGCN, тому безпосередньо використовує граф карти як вхідні дані, кодуючи кожен окремий сегмент смуги руху та застосовуючи згортки графа для включення інформації про зв'язність, як показано на рис. 2.9. Потім кожен агент підключається до своїх локальних сегментів смуги руху відповідно до порогового значення відстані.

Можна далі кодуювати кожен характеристичу траєкторії агента в окремий граф дороги, а потім проводити обчислення взаємодії агентів шляхом об'єднання самих графів смуг руху, щоб локалізувати взаємодії. PGP технологія вивчає політику обходу графів, яка потім використовується для вилучення можливих розгортань маршрутів, з яких виводяться та оцінюються траєкторії.

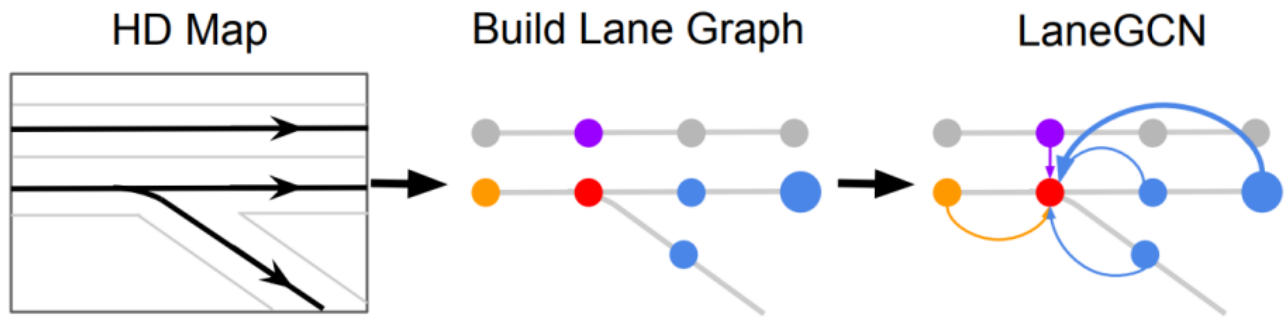


Рисунок 2.9 - Побудова графа карти та кодування зв'язності

Увага.

Агенти водіння можуть запитувати корисну інформацію серед навколишніх сегментів дороги, використовуючи перехресну увагу. Зазвичай розмежування проводиться з увагою до інших інтерактивних агентів шляхом використання окремих альтернативних шарів, як у mmTransformer або SceneTransformer, але деякі роботи намагаються об'єднати фреймворки, як у VectorNet, де глобальна увага без розбору застосовується як до статичних, так і до динамічних елементів, що надалі повторно використовується в TnT.

Вихідне представлення.

Після того, як контекстна інформація була правильно закодована, постає питання про те, як представити прогнозований результат і як його згенерувати. Очевидно, що одним з головних факторів цього рішення є цільове використання після прогнозування, проте ще потрібно зробити багато архітектурних рішень і компромісів залежно від якостей, які ми сподіваємося знайти в прогнозі, будь то точність, різноманітність, охоплення чи доцільність.

Мультимодальне прогнозування.

Хоча історія траєкторії може містити підказки щодо майбутнього руху агента, деякі майбутні рішення залежать від прихованих змінних, які знають лише водій, таких як реальний пункт призначення, уподобання щодо позиції та маршруту, як показано на рис. 2.10. Зміна смуги руху та надання або отримання переваги також залежать від стилю водіння, який залежить від кожної окремої людини та може змінюватися з часом. Зрештою, точні механізми керування з

плином часу ніколи не можуть бути передбачувані з ідеальною точністю, що призводить до зростаючої невизначеності з часом та розкиду можливих координат на кожному кроці часу. Тому можливе майбутнє дорожнього агента є мультимодальним, тобто прогноз повинен містити кілька можливостей, за бажанням з відповідним показником ймовірності.

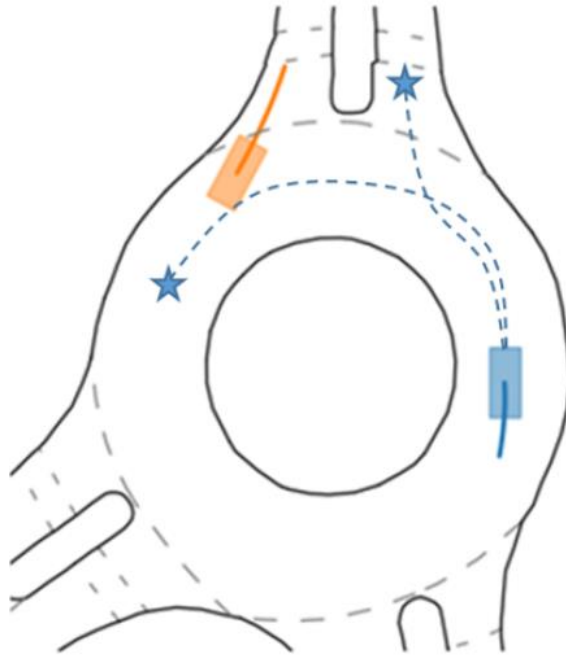


Рисунок 2.10 - Неоднозначні множинні можливі майбутні сценарії з одного історичного контексту

Координати.

Тут ми здебільшого розглядатимемо методи, які намагаються безпосередньо вивести свій прогноз у вигляді послідовності координат, протилежних прогнозуванню розподілу, з якого користувач потім може вибірково отримати точку. VAE є дещо гібридним випадком у цьому відношенні, навчання та висновки систематично вимагають від моделі виведення координат, але проміжний крок спочатку передбачає розподіл, який ефективно вибірково аналізується.

Скалярні координати.

Безпосередньо прогнозують майбутні абсолютні координати вздовж автомагістралі, проте великий діапазон можливих координат вимагає від них

виведення проміжного значення швидкості, а потім інтегрування його для поздовжніх координат (поперечні координати залишаються в обмеженому діапазоні в опорній точці автомагістралі). Щоб уникнути проблеми великого діапазону, необхідно прогнозувати відносні координати початкового положення транспортного засобу як єдине прогнозування максимального А-постеріорного вектора (MAP).

Для мультимодальності можна безпосередньо передбачити набір координат, по одному для кожної модальності. Для їх навчання зазвичай використовується програш за принципом «переможець отримує все», де навчається лише прогноз, найближчий до істини. Для представлення алеаторичної невизначеності деякі методи використовують гаусівську модель, проілюстровану на рис. 2.11, де середнє значення — це просто передбачувані координати точки, а дисперсія являє собою оцінену похибку моделі, яка зазвичай зростає з часом та швидкістю.

Елементи керування Траєкторія++ та Багатошляховий++ прогнозують прискорення та керування, а потім інтегрує їх для отримання фізично можливих траєкторій. Багатошляховий++ спостерігає незначне зниження продуктивності при цьому, але обирає цей компроміс для кращої наближеності до реальності. Використання таких похідних також забезпечує кращу інтерполяцію, як і прогнозування коефіцієнта полінома.

Авторегресивні моделі.

Замість прямого прогнозування набору координат для кожного часового кроку, деякі підходи використовують послідовну природу траєкторій і роблять кожен наступний часовий крок залежним від попереднього через рекурентну архітектуру, як показано на рис. 2.11. Під час навчання вони використовують базові координати замість передбачуваних для обумовлення на попередніх часових кроках.

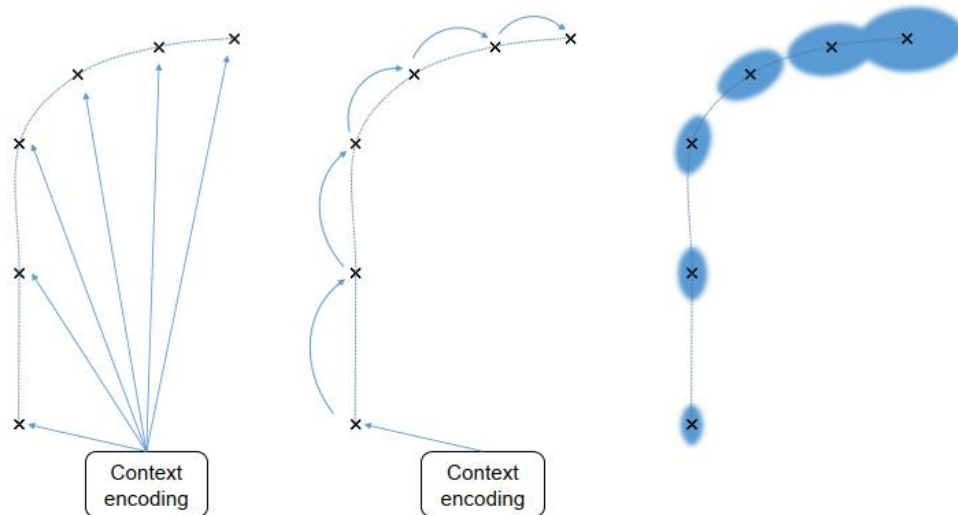


Рисунок 2.11 - Різні параметри для виведення траєкторії

Якорі

Моделі, що прогнозують набір мультимодальних координат, можуть страждати від колапсу мод, якщо всі модальності навчаються одночасно, або від низької ефективності вибірки, якщо одночасно навчається лише одна модальність. Крім того, вивченим модальностям може бракувати різноманітності, оскільки вони будуть зосереджені навколо середніх значень найімовірніших позицій на навчальному наборі. Щоб допомогти з різноманітністю, багатошляховий метод кластеризує навчальні траєкторії у фіксований набір попередньо визначених якорів. Потім модель має класифікувати вибірку за відповідними найближчими якорями та оцінити коефіцієнти регресії для відповідності ставці на якорі. Відносно реальної траєкторії. CovertNet пропонує подібне прогнозування на основі якорів, але додає динамічні якорі траєкторії, адаптовані до динаміки поточного агента. PRANK використовує підхід, більше пов'язаний з методами ранжування та метричним навчанням, шляхом оцінювання прогнозу у великому словнику, що зберігає до 2 мільйонів траєкторій.

Пізніші підходи використовують апріорні дані, надані навколишньою картою, для вибору ймовірних точок прив'язки. TnT генерує пропозиції щодо кінцевих цілей на основі схеми дороги та передає їх у модель для оцінки

ймовірностей та коефіцієнтів регресії. LaneRCNN також використовує граф карти, в який включені ознаки агента, для генерації оцінок ймовірності в кожному сегменті, як показано на рис. 2.12. PRIME переглядає досяжні шляхи з графа зв'язаної карти та видає можливі траєкторії для кожного з них, перш ніж використовувати оцінювач на основі навчання для їх оцінювання. Ці підходи також згадуються як методи, засновані на цілях, оскільки вони дискретизують можливі траєкторії, починаючи з їхньої виведеної мети, а потім оцінюючи повний проміжний рух.

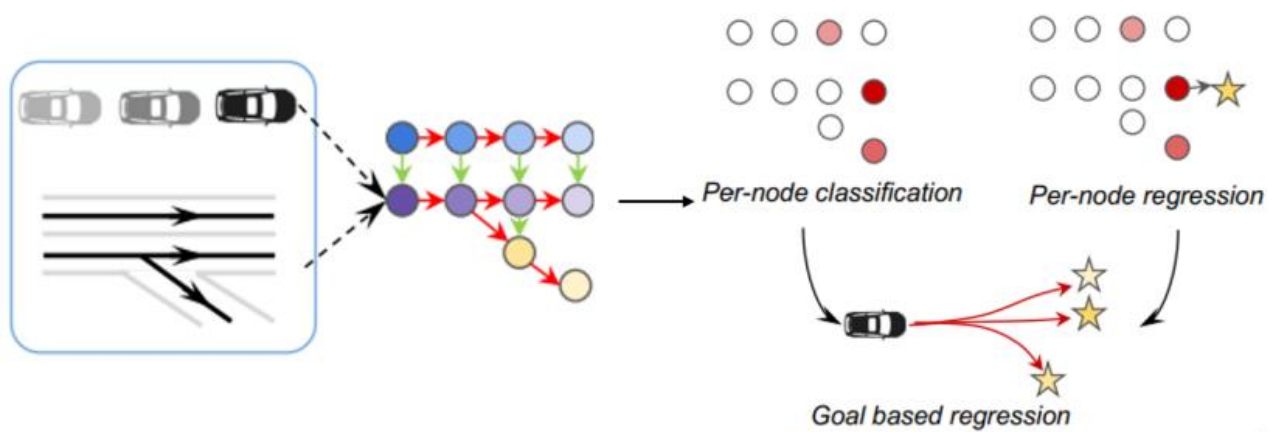


Рисунок 2.12 - Прогнозування траєкторії на основі цілей з використанням попередньої карти.

У новіших підходах перевагу надається використанню вивчених опорних точок, щоб отримати вигоду від їхньої різноманітності без необхідності...приймати емпірично упереджені рішення щодо того, звідки мають братись якорі. Multipath++ [Варадара-[Ян та ін., 2022]та автоботи[Гіргіс та ін., 2021]тому зберігаються вивчені латентні вбудовування, кожне з яких відповідає різній модальності, що робить їх дещо гібридним підходом між прогнозуванням на основі прив'язки та звичайним прямим прогнозуванням координат, оскільки вихід моделі залишається набором передбачуваних координат, а латентні вбудовування не можна візуалізувати для інтерпретації.

Прогнозування траєкторії на основі вибірки

Методи на основі вибірки мають підхід до мультимодальності, відмінний від описаних вище робіт. Замість прогнозування фіксованої кількості координат або ймовірностей, вони використовують внутрішню випадкову вибірку, яка дозволяє їм давати стільки ж прогнозів, скільки разів вони вибірково перевіряються.—Таким чином, кожен висновок створює лише одну траєкторію, але ці висновки не є детермінованими, і тому генеруватимуть іншу ймовірну траєкторію при повторному просуванні.

Варіаційні автокодери (VAE).

Варіаційні автокодери висувають гіпотезу, що майбутнє y обумовлений латентною змінною z . Параметри апіорного розподілу $p(z|x)$ вибірково обробляються для отримання значень z , які потім декодуються генератором $q(y|z, x)$. p та q апроксимуються нейронними мережами, навченими для зіставлення отриманого прогнозованого y з істинним значенням. Цей процес проілюстровано на рис. 2.13 на відміну від звичайного детермінованого процесу. Під час навчання також вивчається апостеріорний розподіл $p(z, y, x)$ та використовується для вибірки z зі знанням майбутнього y , щоб усунути неоднозначність кількох можливих модальностей, а апостеріорний $p(z, y, x)$ та апіорний $p(z|x)$ потім обмежуються тим, щоб мати подібні розподіли. PRECOG вибірки певного латентного значення для кожного часового кроку, але вибірка змінна повинна бути послідовною на кожному кроці часу, щоб представляти керований намір, який має бути однаковим на всіх наступних кроках часу.

Основним недоліком цих генеративних методів є те, що вони не забезпечують оцінку ймовірності, а також не можуть гарантувати, що вибірккові модальності добре охоплюють усі майбутні можливості. Lookout система намагається виправити це, навчаючи семплер для покращення охоплення вибіркових модальностей та оцінювача для оцінки їх ймовірностей.

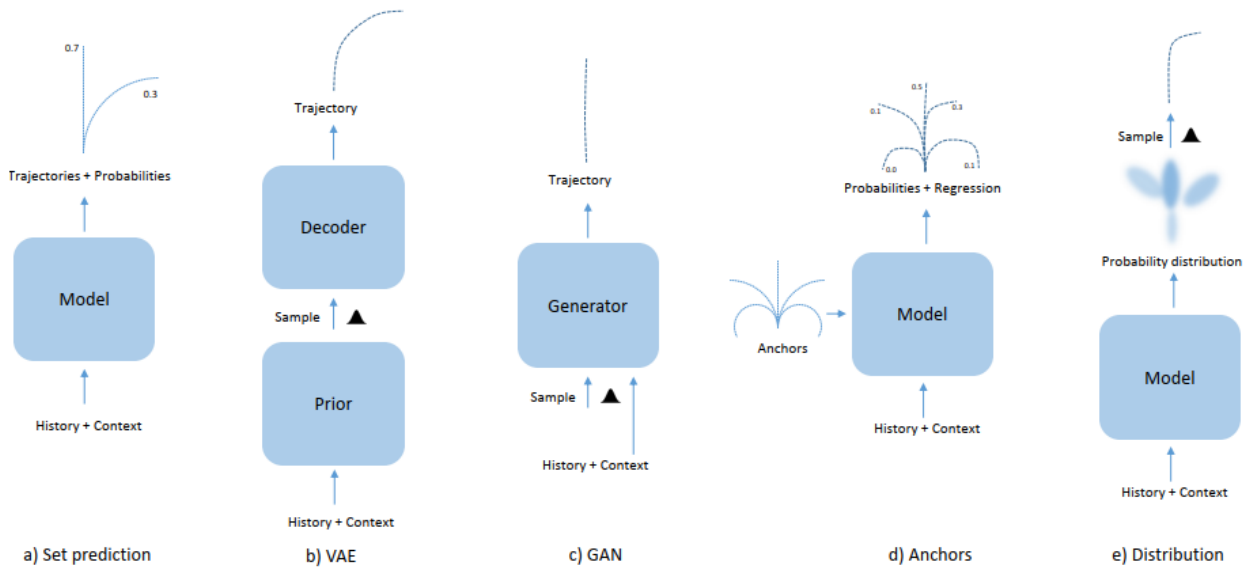


Рисунок 2.13 - Різна архітектура моделі для мультимодального та варіаційного прогнозування

Генеративно-змагальні мережі

Це ще одна генеративна техніка, яка навчає генератор для отримання реалістичних траєкторій з вибіркового вхідного шуму. Генератор навчається за допомогою дискримінатора зіставлення, згенерованого розподілу з істинним. Навчання називається змагальним, оскільки дискримінатор вчиться розділяти бали, що приписуються «фальшивим» або «реальним» траєкторіям, тоді як генератор намагається обдурити дискримінатор, наближаючи два розподіли один до одного, що призводить до протилежних втрат. Порівняно з VAE, латентне значення спочатку вибірково вибирається безпосередньо без будь-якого контексту, як показано на рис. 2.13, замість того, щоб бути похідним від вивченого умовного апіорного розв'язання.

Більшість існуючих методів навчання для прогнозування траєкторії страждають від множинності. Перш за все, ці методи повинні бути мультимодальними, але надані базові дані є унімодальними за своєю природою, а зазвичай використовується багатоголова структура з кількох майбутніх прогнозів не спирається на жодний конкретний диференціатор між кожною головою, що призводить до сильної залежності від випадкової ініціалізації.

Існуючі рішення щодо використання якорів часто спираються на широкі та чутливі до зміщення значення, як припущення та можуть виключати важливі випадкові ситуації, такі як відхилення агента від зони руху. Тому наша мета полягає в розробці моделі прогнозування, яку можна легко навчити та яка дасть природний мультимодальний та необмежений прогноз.

2.2 Розробка моделі прогнозування траєкторії за допомогою теплових карт

Представлення теплової карти

Після багатьох нещодавніх робіт[Чжао та ін., 2021, Гу та ін., 2021, Цзен та ін., 2021], ми зосереджуємося на попередньому—диктуючи останню точку траєкторії y_T , і стверджуємо, що повну траєкторію потім можна інтерполювати не-однозначно з цієї кінцевої точки. Це дозволяє нам спростити задачу, щоб змодельовати лише розподіл ймовірностей цієї останньої точки. Ми представляємо розподіл можливих майбутніх подій за допомогою двовимірної карти ймовірностей, яка дає необмежене наближення ймовірності позиції агента. Така карта тепла з відповідним контекстом та траєкторіями проілюстрована на рис. 2.14.

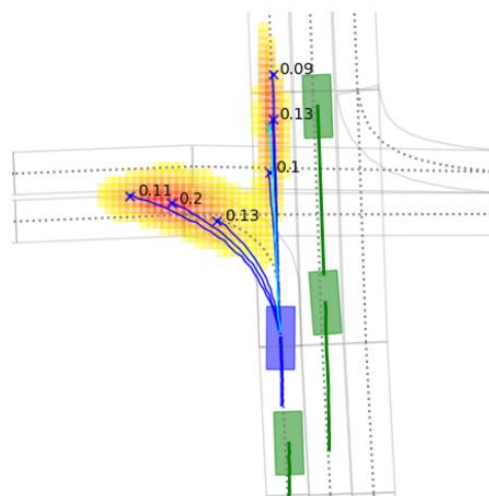


Рисунок 2.14 - Вихід теплової карти, що відображає прогнозовану кінцеву точку траєкторії

Ця тепла карта представлена у вигляді квадратного зображення та природно враховує мультимодальну предикцію, де кожен піксель представляє можливу майбутню позицію цільового агента. Це також дозволяє повністю описувати майбутню невизначеність у розподілі ймовірностей, без необхідності вибирати його моди чи середні значення. Іншими словами, ми переформулюємо звичайну задачу регресії прогнозування траєкторії в задачу класифікації, де кожен піксель сітки є можливим класом.

Декодування теплової карти в траєкторію.

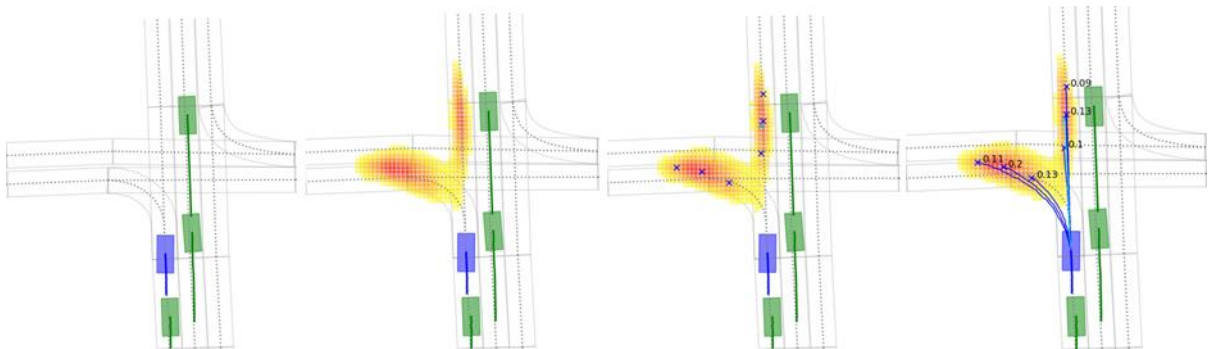


Рисунок 2.15 - Конвеєр для прогнозування траєкторії за допомогою карти ймовірностей. Зліва направо: карта контексту, траєкторії цільового агента (синій) та сусіда (зелений) подаються як вхідні дані для мережі; вихідні дані карти тепла мережі; вибіркові кінцеві точки; траєкторії побудовані для кожної кінцевої точки.

Ми ілюструємо процес використання теплової карти для прогнозування траєкторії на рис. 2.15. Спочатку модель прогнозує просторову теплову карту, враховуючи траєкторію історичного минулого, інших агентів та локальний контекст карти. На другому кроці ми вибираємо з теплової карти скінченну кількість можливих майбутніх місць розташування з можливістю вибрати, яку метрику ми хочемо оптимізувати, без перенавчання моделі. Нарешті, ми будуємо повні траєкторії на основі минулої історії та обумовлені вибіркою кінцевих точок.

Для повноти ми спочатку почнемо з пост-обробки цієї сітки ймовірностей та того, як вона використовується, оскільки це є спільним для всіх різних методів

декодування, які ми пізніше вивчимо, і краще розглянемо контекст використання теплової карти.

Кінцева точка вибірки.

Наша мета полягає в тому, щоб спробувати теплову карту ймовірностей, щоб оптимізувати вибрану нами метрику продуктивності. У більшості наборів даних, для кінцевої прогнозованої точки використовуються дві основні метрики: MissRate (MR) та Final Displacement Error (FDE). MissRate відповідає відсотку прогнозу, який знаходиться далі за певний поріг від еталонного значення, а FDE – це просто середнє значення відстані l^2 між прогнозом та еталонним значенням. Коли вихідний сигнал є мультимодальним, з прогнозами k , використовуються мінімальні Final Displacement Error $\min FDE_k$ та Miss Rate відносно k прогнозів MR_k .

Оптимізація коефіцієнта промахів.

Ми розробляємо метод вибірки для оптимізації коефіцієнта помилок між прогнозованими модальностями та істинним значенням. Випадок визначається як пропущений, якщо істинне значення знаходиться далі ніж на 2 м від прогнозу. Для заданої області A ймовірність того, що істинне значення Y знаходиться в цій області, дорівнює інтегралу розподілу ймовірностей p в цій області

$$P(Y \in A) = \int_{x \in A} p(x) dx \quad (2.3)$$

Отже, для прогнозів k , враховуючи 2D розподіл ймовірностей, вибірка мінімізує очікуваний MR – це той, який максимізує інтеграл розподілу ймовірностей майбутнього на площі, визначеній як кола радіусом 2 м навколо передбачень k :

$$E(\mathbb{1}_{\min_k \|c_k - Y\| > 2}) = 1 - \sum_k \int_{\|c_k - x\| < 2} p(x) dx \quad (2.4)$$

Тому ми обробляємо дані жадібним способом та ітеративно вибираємо місце з найвищим значенням інтегрованої ймовірності в його колі діаметром 2 м. Після отримання такої точки ми встановлюємо на нуль значення теплової карти під визначеним колом і переходимо до вибору наступної точки тим самим методом.

Результат проілюстровано на рис. 2.16. Ми бачимо, що кожну вибірку точку можна оточувати колом радіусом 2 м, яке ледь перекривається з іншими колами. Кожна точка вибірково розташована майже на однаковій відстані від інших, оскільки встановлення ймовірності для попередніх точок рівною нулю встановлює дуже суворе обмеження на мінімальну відстань між точками.

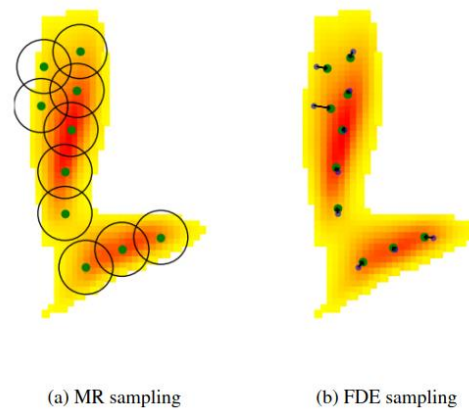


Рисунок 2.16 - Ілюстрація методів вибірки теплової карти

Для реалізації ми обробляємо покриту область для кожної точки, використовуючи шар згортки з фіксованими вагами ядра таким чином, щоб наближено утворити коло діаметром 2 м. На практиці ми використовуємо не радіус 2 метри, а 1,8 метра, як ми з'ясували, для кращої продуктивності. Ми також масштабуємо теплову карту до $0,25 \times 0,25$ м² на піксель з білінійною інтерполяцією для більш точного прогнозування місця розташування.

Оптимізація кінцевої помилки переміщення.

Вихідне зображення можна представити як дискретний розподіл ймовірностей (x_i, p_i) , де x_i представляє центр пікселя p_i – пов'язане з ними

значення ймовірності. Оптимізація кінцевої помилки зміщення за прогнозами k означає знаходження k центроїдів c_k , які мінімізують таку величину:

$$\text{minimize}_c \sum_i p_i \|c - x_i\| \quad (2.5)$$

Для цього ми розробляємо наш алгоритм вибірки для оптимізації FDE.

Лістинг 2.1

Algorithm 1: MR Sampling Algorithm

input: Probability map $p(x)$

K number of predictions

R threshold for Miss Rate

for $k = 1..K$ **do**

 Find c_k maximizing $\int_{\|c_k - x\| < R} p(x) dx$

 Set $p(x) = 0$ for all x such that $\|c_k - x\| < R$

end

Ми замінюємо класичне середньозважене значення $p_i x_i$ для кожного центроїда c_k на $\frac{1}{d_k} \sum_i p_i x_i$, де d_k – це дисперсія відстані між точкою x_i та центроїдом c_k , щоб бути стійкішою до викидів та враховувати оптимізацію норми l_2 замість її квадрата.

По суті, ми оновлюємо кожен прогноз як середньозважене значення його локальної околиці в радіусі 3 м. Коефіцієнт m_i , де m_i – це відстань між точкою x_i та її найближчим центроїдом, дозволяє гнучко визначити межі розділів порівняно з KMeans (де ми б використовували $1/d_k \leq m_i$ замість цього): коли x_i знаходиться в розділенні прогнозу k , його значення дорівнює 1, тоді як коли воно знаходиться поза ним, воно зменшується, дорівнюючи 0 у точній позиції іншого прогнозу k' , де його ніколи не можна покращити зміщенням k .

Ми ініціалізуємо центроїди результатами алгоритму оптимізації Miss Rate та використовуємо число ітерацій L як параметр для налаштування компромісу між коефіцієнтом промахів та FDE: коли L дорівнює нулю, коефіцієнт промахів

оптимізується, тоді як коли L збільшується, MR жертвується для покращення FDE.

Лістинг 2.2

Algorithm 2: FDE Sampling Algorithm

input: Set of points x_i with probability weight p_i
 L number of iterations to run the algorithm
 Initialization of K centroids c_k

for $l = 1..L$ **do**
 Compute d_i^k the matrix of distance of point x_i to each centroid c_k
 Compute m_i the distance of point x_i to the closest centroid c_k
 for $k = 1..K$ **do**
 Compute new centroid coordinates :
 $c_k = \frac{1}{N} \sum_i \mathbb{1}_{d_i^k \leq 3 \frac{p_i}{d_i^k} \frac{m_i}{d_i^k}} x_i$
 with $N = \sum_i \mathbb{1}_{d_i^k \leq 3 \frac{p_i}{d_i^k} \frac{m_i}{d_i^k}}$
 end
end

Вихідні дані алгоритму проілюстровано на рис. 2.16, де можна спостерігати, що центроїди зближуються один з одним, жертвуючи загальним покриттям, але наближаючись до областей з високою ймовірністю, щоб зменшити очікувану відстань. Результати цього компромісу детальніше проілюстровано на рис. 2.17, що кожна ітерація Algo.2 зменшує minFDE6 та збільшує MR6.

Експерименти з вибірки кінцевих точок.

Ми виділяємо результати нашої вибірки у табл. 2.1 і порівнюємо їх з іншими можливими стратегіями вибірки: ми намагаємося ранжувати пікселі за ймовірністю та вибирати їх у порядку спадання, видаляючи при цьому пікселі, що перекриваються, що знаходяться ближче, ніж радіус 1,8 м, використовуючи класичний метод немаксимального придушення. Ми також пробуємо KMeans.

Ми показуємо на рис. 2.17 результати нашого компромісу між MR6 та FDE6 на тестовому наборі Argoverse завдяки параметру L алгоритму Algo.2. Ми також включаємо пункти для інших 10 найкращих методів

Таблиця 2.1 - Дослідження абляції на траєкторії вибірки (набір для валідації Argoverse)

Bottleneck	K=1		K=6	
	minFDE	MR	minFDE	MR
Pixel ranking with NMS	3.07	51.0	1.21	10.7
KMeans	3.06	51.6	1.23	9.3
Ours (MR)	3.02	50.7	1.28	6.8
Ours (FDE L=6)	3.01	50.5	1.16	7.4

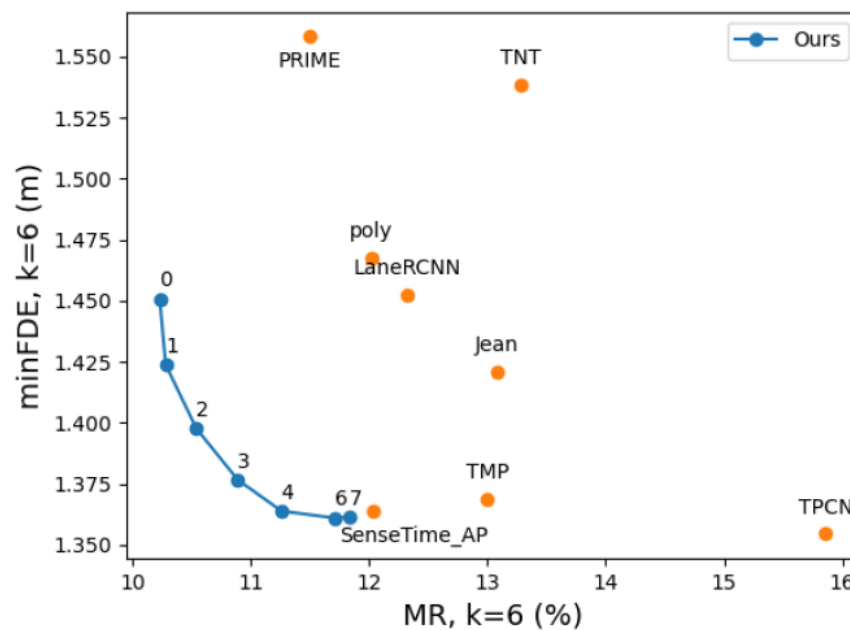


Рисунок 2.17 - Компроміс FDE6 - MR6. Внизу ліворуч показано краще. Точки кривої (сині) отримані зі збільшенням кількості ітерацій L алгоритму.2 від 0 до 7. Також включено бали за інші методи нарахування балів за топ-10 (помаранчевий колір).

Таблиця лідерів станом на 1 березня 2021 року для порівняння. Наш метод досягає найкращого можливого MR6 та дозволяє покращити FDE6 до другого найкращого, залишаючись першим у MR6 (четверта точка кривої, отримана з $L = 4$).

Повна регресія траєкторії.

Для кожної вибіркової цільової точки використовується окрема модель для генерації повних траєкторій, обумовлених кінцевою точкою. Будь-який метод планування, що не базується на навчанні, може виконати це завдання, але в нашому випадку ми використовуємо простий MLP, проілюстрований на рис. 2.18.

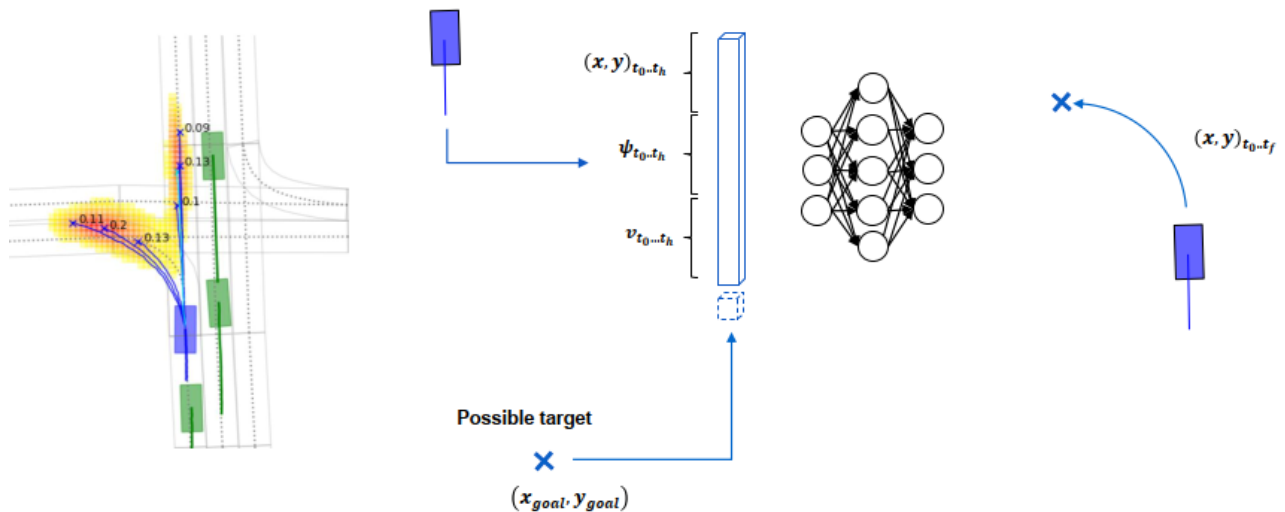


Рисунок 2.18 - Умовна регресія траєкторії досягнення мети.

Ця модель застосовує повністю зв'язний шар для кодування історії цільового агента у вектор із 32 ознак, який потім об'єднується з координатами (x, y) цільового майбутнього місцезнаходження. Потім застосовується інший повністю зв'язний шар для отримання вектора із 64 ознак, який потім перетворюється через останній повністю зв'язний шар до набору місць розташування, що представляють проміжне положення агента у часовому інтервалі $[1, T]$. Ймовірність траєкторії обчислюється як інтеграл від теплової карти ймовірностей під колом радіуса 2 м навколо кінцевої точки траєкторії.

Під час навчання ця регресійна модель навчається з кінцевою точкою ground-truth для регресії ground-повна траєкторія правди у спосіб, що примушує вчителя.

Як створити таку теплову карту?

Згортова нейронна мережа: НОМЕ (Вихід теплової карти для оцінки руху)

Найпростіший спосіб отримання теплових карт зображень – це використання моделі CNN, яка також приймає зображення як вхідні дані. Однак ми також хочемо використовувати детальну інформацію, присутню в минулій послідовності координат, тому ми розробляємо гібридну архітектуру, яка об'єднує растеризовану та скалярну інформацію.

Кодування карти та минулої траєкторії.

Локальний контекст доступний у вигляді карти високої чіткості, центрованої навколо цільового агента. Ми растеризуємо карту високої чіткості. Карта у 5 семантичних каналах: зона руху, межі смуг руху та спрямовані центральні лінії з їхніми напрямками, закодованими за допомогою HSV на 3 каналах. Ми також додаємо траєкторію цільового агента у вигляді рухомого прямокутника на 20 каналах історії та історію інших агентів на ще 20 каналах. Кінцевим вхідним даними є зображення (224, 224, 45) з роздільною здатністю 0,5 x 0,5 м² на піксель. Це зображення обробляється класичною моделлю CNN, що чергує згорткові шари та максимальне об'єднання для зменшення масштабу, щоб отримати кодування (14, 14, 512), як показано у верхній лівій частині рис. 2.19.

Скалярна історія агентів також береться на вхід до моделі у вигляді списку 2D координат. Відсутні кроки часу доповнюються нулями, а до траєкторії приєднується двійкова маска, яка вказує, чи було застосовано доповнення, а також позначки часу для кожного кроку, таким чином ми отримуємо вхідні дані (H , 4) для кожного агента. Кожна траєкторія агента проходить через 1D згортковий шар, за яким йде UGRU рекурентний шар. Ваги є спільними для всіх агентів, окрім цільового агента.

Міжагентна увага для взаємодії.

Для цільового агента генерується вектор запитів, а для інших акторів створюються вектори ключів та значень. Нормалізований скалярний добуток запитів та ключів створює карту уваги від цільового агента до іншого агента, яка потім використовується для об'єднання їхніх ознак цінностей у вектор контексту. Потім вектор контексту додається до вектора ознак цільового транспортного засобу через залишковий зв'язок, після чого виконується нормалізація шарів.

Отримане кодування траєкторії *Etra_jector* у потім повторюється, щоб відповідати контекстному кодуванню розмірів *Eraster*. Остаточне кодування *Econtext* є результатом об'єднання обох кодувань *Eraster* та *Etra_jector* у.

Збільшений вихідний розмір для більшої дальності.

Через високу швидкість деякі автомобілі можуть проходити більший діапазон у часовому горизонті T , який охоплюється вхідним діапазоном 56 м. Однак, просте збільшення розміру вхідних даних значно збільшить обчислювальне навантаження, не обов'язково приносячи корисну інформацію. Тому ми хочемо збільшити розмір вихідних даних, зберігаючи просторові відповідності між шарами. Для цього ми застосовуємо транспоновані згортки з кроком 1 та розміром ядра 3. Оскільки 1 вхідний піксель підключений до сітки з 3×3 вихідних пікселів, крайові пікселі генерують нову рамку пікселів навколо них, збільшуючи розмір кодування на 1 у кожному напрямку. Ми застосовуємо 2 таких шари, що призводить до доповненого кодування (18, 18, 512), так що після масштабування декодоване вихідне зображення матиме розмір (288, 288), що відповідає діапазону 72 м.

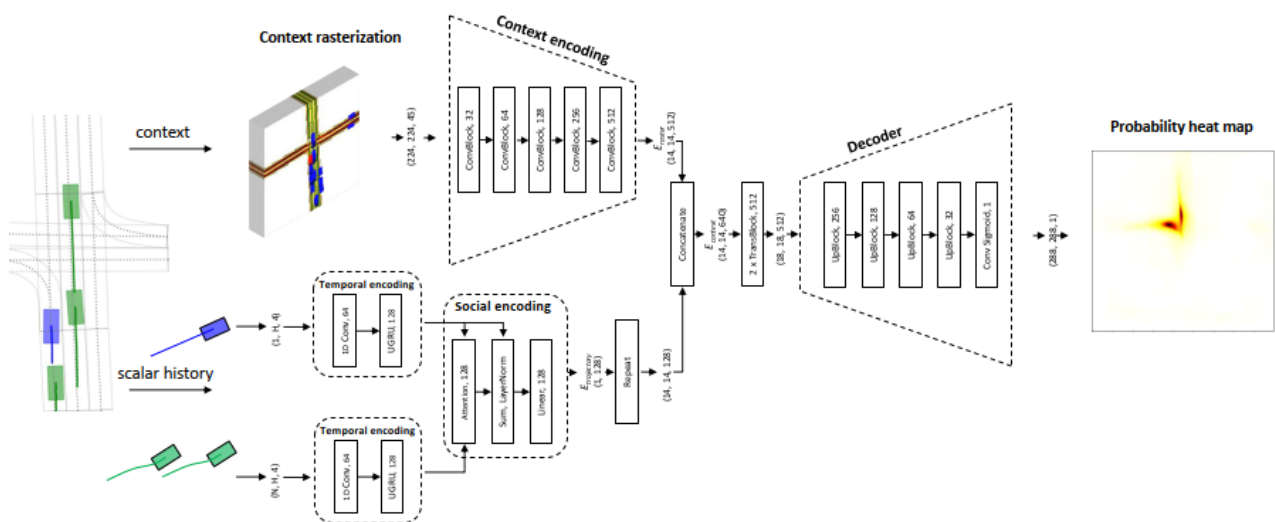


Рисунок 2.19 - Умовна регресія траєкторії досягнення мети

Декодування та навчання теплових карт.

Заключною частиною моделі є згортковий декодер, який чергує транспоновані згортки для масштабування та класичні згортки, доповнені сигмоподібною активацією. Ми виводимо зображення $\hat{Y}$ з роздільною здатністю, подібною до растрового входу ($0,5 \times 0,5$ м²/піксель). Вихідною ціллю є зображення Y з гауссовою графікою, центрованою навколо позиції істинної області. Це зображення навчається з попиксельною втратою фокусної відстані, натхненною, усереднено по загальній кількості P пікселів p теплової карти:

$$L = -\frac{1}{P} \sum_p (Y_p - \hat{Y}_p)^2 f(Y_p, \hat{Y}_p)$$

$$\text{with } f(Y_p, \hat{Y}_p) = \begin{cases} \log(\hat{Y}_p) & \text{if } Y_p=1 \\ (1 - Y_p)^4 \log(1 - \hat{Y}_p) & \text{else} \end{cases} \quad (2.6)$$

де ненульові пікселі навколо центру гаусової функції служать коефіцієнтами зменшення штрафу, а квадратичний коефіцієнт похибки дозволяє градієнту зосередитися на погано передбачуваних пікселях. Ми використовуємо стандартне відхилення 0,67 пікселя для гаусової функції.

Деталі впровадження

Було навчено всі моделі протягом 16 епох з розміром пакету 32, використовуючи оптимізатор Adam, ініціалізований зі швидкістю навчання 0,001. Кожен кадр вибірки центрований на цільовому агенті та вирівняний з його заголовком. Ми ділимо швидкість навчання навпіл на епохах 3, 6, 9 та 13. Ми доповнили навчальні дані, видаляючи кожен растровий канал з ймовірністю 0,1 та повертаючи кадр на рівномірний випадковий кут у $[-\pi/4, \pi/4]$ у 50% вибірок. Всі шари згортки є CoordConv з ядром розміром 3x3 (3 для 1D Convs), після чого виконується BatchNormalization та активація ReLU.

Порівняння з найсучаснішим станом техніки.

Як показано в табл. 2.2 наші результати порівняно з іншими методами на тестовому наборі прогнозування руху Argoverse. Орієнтовним показником є MR6, де ми посідаємо перше місце та значно покращуємо попередні результати,

демонструючи, що наявність виводу теплової карти забезпечує найкраще покриття порівняно з попереднім рівнем техніки. Ми також перевершуємо інші методи на $p\text{-minFDE}_6$, демонструючи чудове моделювання розподілу ймовірностей між прогнозами. Ще одне цікаве спостереження полягає в тому, що методи, такі як LaneGCN, дуже добре працюють на minFDE_6 , такі як LaneGCN та TPCN мають гірший MR_6 як недолік PRIME має найближчий до нашого MR_6 , але значно вищий minFDE_6 у порівнянні. Результати як нашої вибірки, оптимізованої для MR, так і minFDE з тією ж навченою моделлю. Наша вибірка FDE з $L = 4$ жертвує 1,1 бала MR_6 для 9 см minFDE_6 , що ставить нас на друге місце за minFDE_6 , водночас залишаючись достатньо хорошим для 1st позиції в таблиці лідерів.

Таблиця 2.2 - Результати в таблиці лідерів прогнозування руху Argoverse[ліа,] (тестовий набір)

	K=1			K=6			
	minADE	minFDE	MR	minADE	minFDE	p-minFDE	MR
WIMP [Khandelwal et al., 2020]	1.82	4.03	62.9	0.90	1.42	3.21	16.7
LaneGCN [Liang et al., 2020]	1.71	3.78	59.1	0.87	1.36	3.16	16.3
Alibaba-ADLab	1.97	4.35	63.4	0.92	1.48	3.23	15.9
TPCN [Ye et al., 2021]	1.64	3.64	58.6	0.85	1.35	3.11	15.9
HIKVISION-ADLab-HZ	1.94	3.90	58.2	1.21	1.83	3.62	13.8
TNT [Zhao et al., 2021]	1.78	3.91	59.7	0.94	1.54	3.33	13.3
Jean [Mercat et al., 2020]	1.74	4.24	68.6	1.00	1.42	3.21	13.1
TMP [Liu et al., 2021b]	1.70	3.78	58.4	0.87	1.37	3.16	13.0
LaneRCNN [Zeng et al., 2021]	1.69	3.69	56.9	0.90	1.45	3.24	12.3
SenseTime_AP	1.70	3.76	58.3	0.87	1.36	3.16	12.0
poly	1.70	3.82	58.8	0.87	1.47	3.28	12.0
PRIME [Song et al., 2022]	1.91	3.82	58.7	1.22	1.56	3.04	11.5
Ours-HOME (FDE L=4)	1.72	3.73	58.4	0.92	1.36	3.08	11.3
Ours-HOME (MR)	1.73	3.73	58.4	0.94	1.45	3.03	10.2

Дослідження абляції: вплив виводу теплової карти.

Вплив представлення виходу показано в табл. 2.3 використовуючи ту саму основу кодування та замінюючи декодер зображень за допомогою глобального пулінгу, а потім регресійної головки з 6 координатних модальностей. Ми навчаємо вихід регресії з використанням регресійної втрати l_1 за принципом "переможець отримує все" та втрата класифікації, де ціль отримується за

допомогою м'якого максимуму на відстанях між прогнозами та еталонними даними. Оскільки глобальне об'єднання призводить до втрати просторової інформації із зображення, для справедливого порівняння ми також включаємо модель зі "скалярним вузьким місцем", де об'єднання також застосовується до кодування зображення, а потім переформовується для формування зображення, до якого застосовується декодер теплової карти. Для вихідних даних теплової карти використовується MR-семплінг. Ми спостерігали, що вихідні дані теплової карти дають набагато кращий показник промахів, а наявність скалярного вузького місця об'єднання знижує продуктивність, оскільки створює втрату інформації, але не суттєво. Цікаво, що вихідний результат регресії досягає кращого значення minFDE у порівнянні з моделями вихідних зображень, оптимізованих для MR, але все ще гірший, ніж модель, оптимізована для FDE, оскільки цей вихідний результат скалярних координат не залишає місця для будь-якої оптимізації пост-обробки.

Таблиця 2.3 - Дослідження абляції на представленні вихідних даних (набір для валідації Argoverse)

Bottleneck	Output	K=1		K=6	
		minFDE	MR	minFDE	MR
Scalar	Regression	3.81	61.7	1.26	13.0
Scalar	Heatmap	3.07	51.9	1.30	8.0
Image	Heatmap	3.02	50.7	1.28	6.8

Якісні результати.

Додаткові якісні результати ми показуємо на рис. 2.20. Ми наводимо приклади руху по прямій, обгону, криві дороги, що виходять за межі карти та перехрестя. Результат нашої моделі теплової карти використовує та зазвичай відповідає попередньому сценарію з контекстної карти, але також може відхилятися від нього на основі взаємодій, реалістичних спостережень та натяків на розбіжності з історією.

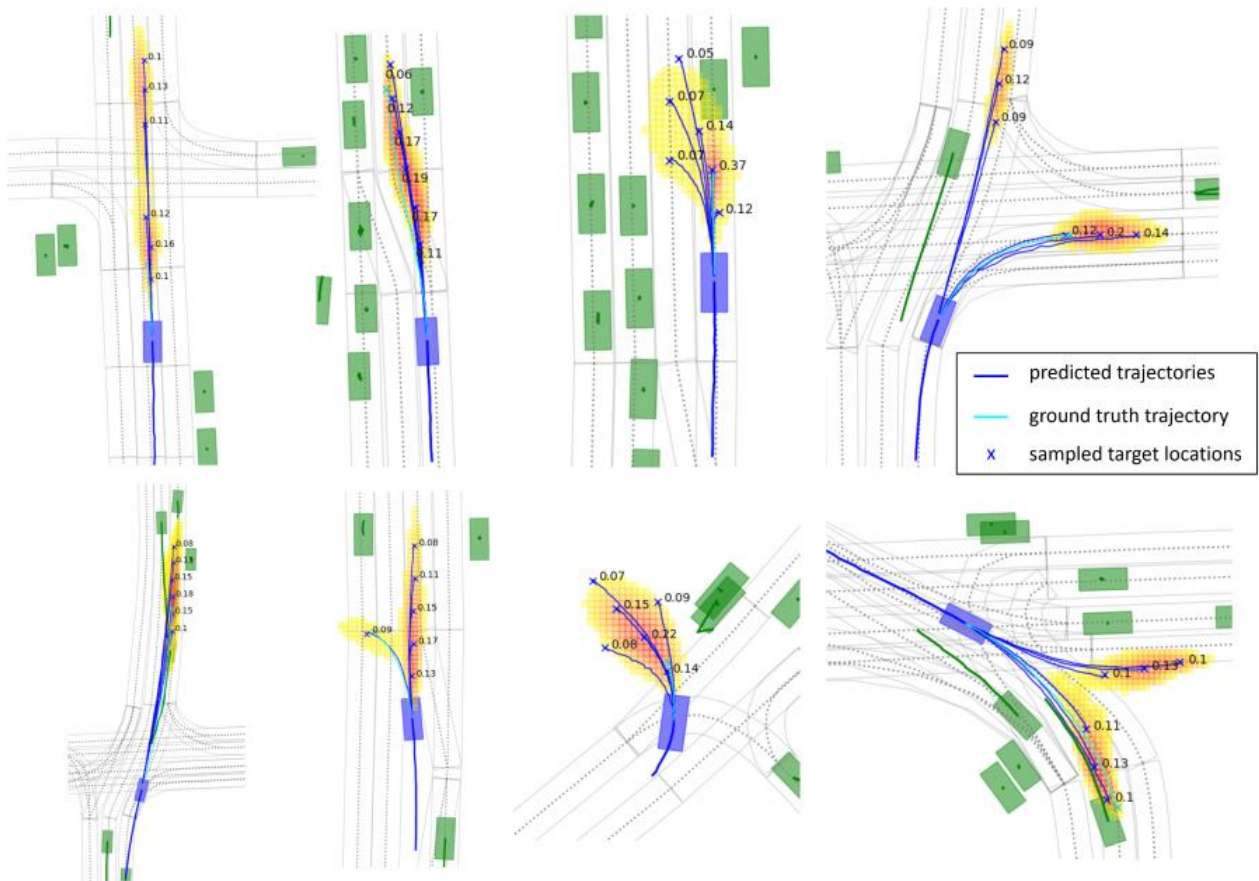


Рисунок 2.20 - Якісні приклади. Жовта/червона теплова карта — це наш прогнозований розподіл ймовірностей, а сині точки — це вибіркові прогнози кінцевих точок. Траєкторія наземного аналізу показана зеленим кольором.

Обмеження. Незважаючи на їх широке використання та оптимізацію в більшості фреймворків глибокого навчання, згорткові шари залишаються досить повільними порівняно з простими скалярними повністю зв'язаними шарами, особливо враховуючи відповідні розміри оброблюваних вхідних даних (повне зображення проти кількох координат). Як наслідок, ми спостерігаємо, що час навчання трохи повільніший, ніж оптимальний для швидкого та ефективного дослідження (кілька днів на навчання), і погано підходить для логічного висновку в реальному часі (~ 200 мс на висновок, тоді як <100 мс було б бажано для повноцінного самостійно керованого конвеєра в реальному часі).

Більше того, хоча оригінальні карти, що містяться в більшості наборів даних, надають інформацію про зв'язність (які смуги руху з'єднані між собою), що особливо зручно, коли йдеться про складні перехрестя з кількома можливими

змiнами смуг руху та поворотами, як у останньому прикладі внизу праворуч на рис.4.7, наш згортковий Кодер не враховує цю інформацію та представляє всю дорожню карту в одному шарі, де смуги руху, що перекриваються, можуть перекривати одна одну.

Теплові карти на основі смуг руху (GOHOME).

HOME використовує повністю згорткову модель та обмежений певним розміром зображення. GOHOME продовжує цю роботу та представляє оптимізовану структуру прогнозування руху, що базується виключно на операціях з графами, щоб ефективно створювати теплову карту з мірою невизначеності, використовуючи векторизовану форму HD-карти.

Система GOHOME зосереджена на операціях на рівні смуги руху, як показано на рис. 2.21 Локальна HD-карта надається в наборі даних у вигляді графіка L смуг руху. Смуга руху являє собою макроділянку дороги (в середньому від 10 до 20 метрів), оскільки наша мета — кодувати зв'язність на макрорівні (сегменти смуг руху), а не на мікрорівні (кожен метр). Кожна смуга руху визначається як послідовність точок центральної лінії та з'єднана зі своїм попередником, наступником, лівим та правим сусідами, якщо вони існують. Ми кодуємо кожен смугу руху в граф дороги, де представлена геометрична інформація та інформація про зв'язність.

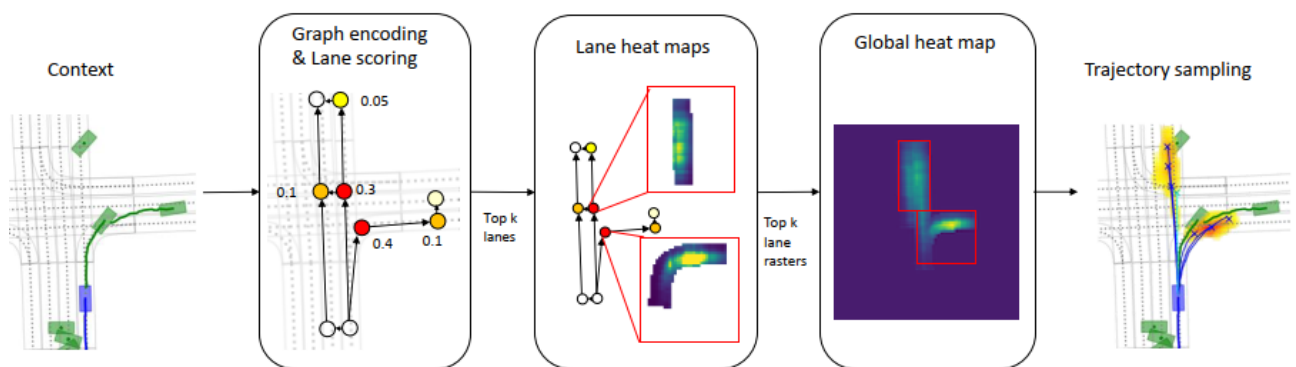


Рисунок 2.21 Конвеєр GOHOME. Графік доріжок, отриманий з HD-Мар, обробляється за допомогою кодера графів. Кожна доріжка потім генерує локальний криволінійний растр, який об'єднується в прогнозовану теплову карту розподілу ймовірностей.

Наша модель дає оцінку для кожної з цих смуг руху, яка використовується для визначення найімовірніших смуг руху. Потім для смуг руху з найвищим рейтингом генерується часткова теплова карта та проектується на глобальну теплову карту. Після цього ми вибірково вибираємо набір кінцевих точок з теплової карти та відтворюємо траєкторію для кожної з них.

3 СПЕЦІАЛЬНА ЧАСТИНА

3.1 Графова нейронна мережа для вхідних даних HD-Map

Архітектура моделі проілюстрована на рис. 3.1. Ми кодуємо кожну смугу за допомогою спільної 1D згортки та UGRU рекурентні *LaneEncoder* на ознаки F каналів C . Ознаки смуги F потім оновлюються за допомогою *GraphEncoder1*, складеної з послідовності чотирьох операцій згортки графів, для поширення інформації про зв'язок:

$$F \leftarrow FW + \sum_r A_r FW_r \quad (3.1)$$

де F – матриця ознак смуги руху (L, C);

W – вивчена вага (C, C) для кодування ознак ego;

A_r та W_r – відповідні матриці суміжності (L, L) та вивчена вага (C, C) для співвідношення $r \in \{predecessor, successor, left, right\}$ отримано з графіка смуг руху.

A_r фіксується, оскільки надходить з мапи, тоді як W_r дозволяє вивчати різні операції для кожного можливого відношення.

Паралельно, кожна траєкторія агента, визначена як послідовність положення, швидкості та рискання, кодується спільним шаром, що також складається зі спільної 1D згортки та шару UGRU. Потім кожна ознака агента оновлюється інформацією карти через шар перехресної уваги *Lanes2Agents* на ознаках lanelet.

Взаємодії між агентами потім враховуються за допомогою шару самоуваги *Agents2Agents* між агентами. Нарешті, ознака цільового агента об'єднується з усіма ознаками смуги за допомогою *Ego2Lanes*, а потім обробляється за допомогою останнього шару *GraphEncoder2*, також складається з 4 згорток графа, для отримання остаточного кодування графа, яке буде використовуватися для генерації різних прогнозів.

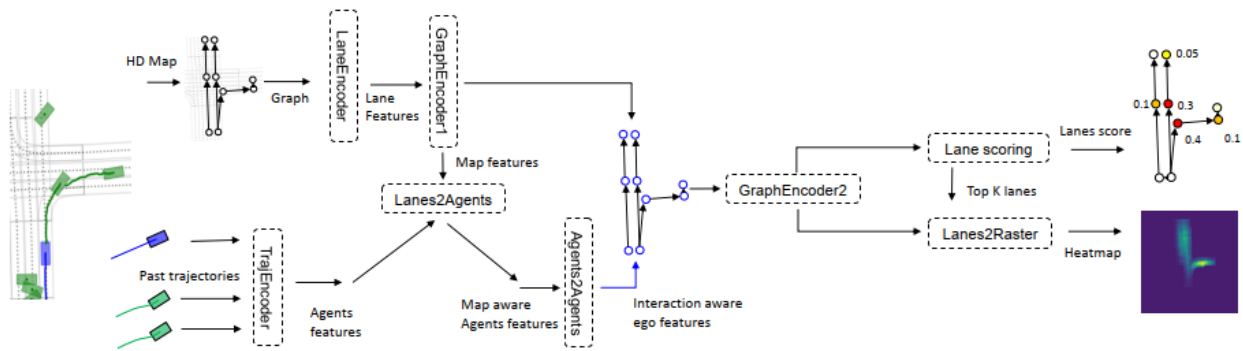


Рисунок 3.2 - Архітектура моделі GOHOME

Порівняно з іншими методами, що використовують графові нейронні мережі, наш метод використовує графові згортки, такі як LaneGCN та LaneRCNN, але застосовує їх до смугових ліній, а не до вузлів смуг (вузол смуги — це окрема точка в послідовності смугової лінії). VectorNet та TNT також використовують смугові лінії, які називаються полілініями, але з'єднують їх за допомогою глобальної уваги замість використання графової зв'язності. Ми вирішили використовувати GNN на смугових лініях, оскільки хотіли ефективніше...-функціональне та високорівневе представлення, що дозволяє легко поширювати інформацію через граф, використовуючи при цьому зв'язність.

3.2 Генерація теплових карт за допомогою растрових зображень на рівні смуги руху

Для виводу теплової карти ми хочемо мати щільне зображення в декартових координатах розмірів (H, W) . Щоб зробити це без використання згортки на повному зображенні, ми створюємо растр для кожної смуги в криволінійних координатах. Ми використовуємо ранжування смуг для генерації цих растрових зображень лише для верхніх k смуг, а не для всіх.

Генерація растра смуг Кожна з малих растрових схем смуг розміром (h, w) має поздовжню довжину 20 м та поперечну ширину 4 м. Ці растри смуг створюються як дискретизація референції Френе-Серре вздовж смуги, як

показано на рис. 3.3. Ми розкладаємо розподіл ймовірностей вздовж лейнелу на поздовжню складову $(h, 1, 8)$ та латеральну складову $(1, w, 8)$, передбачені на основі кодування лейнелу. Ці складові підсумовуються разом із широкомовленням для створення $(h, w, 8)$ розподілу.об'єм. Таким чином, складність створення об'єму $(h+w) \times 8$ замість $h \times w \times 8$. Отриманий об'єм потім об'єднується з попиксельними декартовими координатами, напрямком руху, заповненістю смуг руху та кривизною інформація перед останнім лінійним шаром, на який наноситься сигмоподібна кривина для отримання виходу *Rproba*.

Отримані теплові карти рівня смуги руху *Rproba* потім проектується на повну декартову теплову карту Y^* , використовуючи декартові координати кожного пікселя, як показано на рисунку 4.10. Якщо кілька пікселів рівня смуги руху потрапляють в один і той самий декартовий піксель, їхні значення усереднюються. Ширина растрової смуги руху встановлюється таким чином, щоб сусідні растри смуг перекривалися та могли охоплювати поведінку зміни смуги руху. Цільовим значенням Y для цього остаточного прогнозу Y^* є гаусова функція, центрована навколо позиції цільового агента наземної істини. Ми використовуємо ті ж попиксельні фокальні втрати, що й НОМЕ, визначені в рівнянні. 3.2.

Рейтинг смуг.

Класифікація лейнелетів отримана за допомогою лінійного шару на графовому кодуванні, наступним чином—знижується сигмоїдною активацією. Істинна базова позиція визначається як 1 для всіх смуг руху, де майбутнє положення автомобіля знаходиться всередині полігону смуги руху, та 0 в іншому випадку. Втрата - це бінарна перехресна ентропія, додана до попиксельних втрат рівняння. 4.4 з коефіцієнтом $1e-2$. Оскільки лише частина смуг руху буде фактично корисною для представлення майбутнього розташування автомобіля, ми можемо обчислити растри рівня смуги руху лише для підвибірки смуг руху, що заощадить більше обчислень.

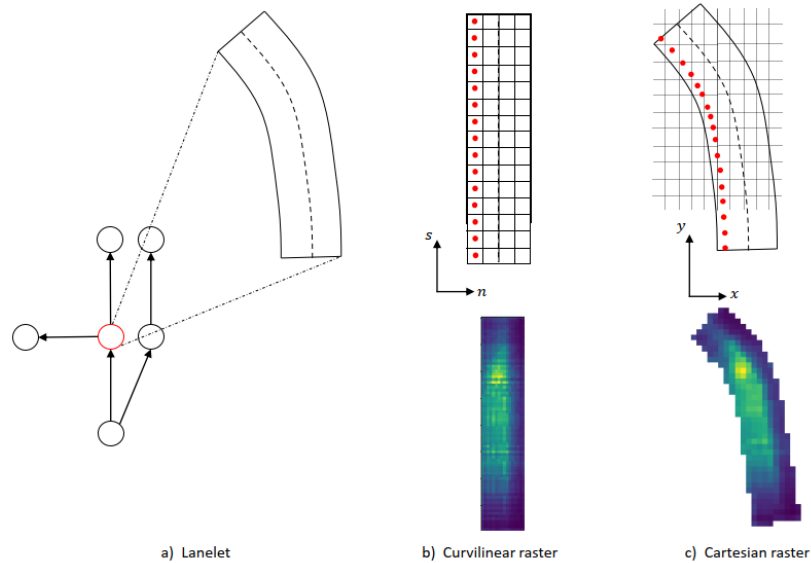


Рисунок 3.3 - Проекція сітки растра смуг руху на декартові координати. а) Один вузол графіка є смугою руху та описує сегмент дороги. б) Прямокутний растр генерується вздовж криволінійної координати. в) Координати ланелета. с) Координати ланелета потім використовуються для проектування передбачуваного растра назад у декартові координати для завершення остаточного виводу теплової карти.

Ми використовуємо оцінку класифікації *clane*, передбачену мережею, щоб вибрати лише смуги руху з найвищим рейтингом k , і обчислюємо та проектуємо растр смуги руху лише для них. Оскільки прогнози растрів для інших невибраних смуг руху були б дуже близькими до нуля в будь-якому випадку, це зовсім не знижує продуктивність.

Декартове з'єднання зображень.

Деякі растри смуг можуть проектуватися на ті самі пікселі та поверх них, про що моделі дуже важко знати заздалегідь. Щоб допомогти моделі дізнатися про перекриття та поширити інформацію про місцезнаходження через растри смуг руху, ми виконуємо першу проекцію растрів смуг руху на декартові координати перед остаточною оцінкою ймовірності. Характеристики растра смуг руху ($h, w, 8$)

Проектуються на $(H, W, 8)$ декартове зображення *I f eatures* за допомогою тієї ж операції раніше описано для теплових карт ймовірностей із усередненими перекриттями. Інформація про заповнюваність (кількість растрових пікселів, агрегованих у кожному декартовому пікселі) також об'єднується з об'ємом *I f eatures*. Потім лінійний шар наноситься на останній вимір *I f eatures*, який потім перепроєктується на криволинейну лінію. Координати вуха та об'єднані з початковими растрами смуг руху *R f eatures* перед остаточною оцінкою та прогнозуванням ймовірності. Таким чином, ознаки перекриваючих смуг руху розподіляються між ними, щоб вони могли поширювати інформацію та гомогенізувати ймовірність.

Масштабування розміру та роздільної здатності зображення.

Хоча діапазон зображення 192 м, що становить 88 м охоплення в кожному напрямку, може бути достатнім для більшості прогнозів міського руху з часовим горизонтом 3 секунди, інші набори даних можуть вимагати прогнозів до 6 або 8 секунд. Тому існує потреба збільшити цей діапазон вихідних даних, що можна зробити без необхідності збільшувати розмір вхідного сигналу, оскільки великі відстані досягаються за допомогою довгих прямих траєкторії, які можна легко екстраполювати на автомагістралі.

Ми порівнюємо масштабування нашої моделі GONOME на основі графів з масштабуванням моделі HOME на основі зображень щодо діапазону виводу та роздільної здатності. Рис. 4.11 підкреслює масштабування вихідного діапазону, де ми використовуємо фіксований вхідний діапазон 128 метрів та роздільну здатність 0,5 м на піксель. У той час як декодери CNN HOME та GNN-HOME призводять до квадратичного масштабування, растри смуг у поєднанні з рейтингом 20 найкращих смуг дозволяють масштабування, яке є навіть менш лінійним.

Ми показуємо на рис. 3.4 час виведення з урахуванням кількості пікселів на метр, що є оберненою величиною до роздільної здатності. Хоча ефективна оптимізація згортки та постійних витрат призводить до близьких часів виведення для початкових 2 пікселів на метр, більш ефективне масштабування чітко

проявляється для вхідних даних GNN та особливо для виходів растрів смуг. Оскільки роздільна здатність растрів смуг також масштабується із загальною роздільною здатністю, квадратична складність все ще застосовується, але зі значно меншим коефіцієнтом, що дозволяє забезпечити реалістичні часи навчання та виведення для точніших роздільних здатностей.

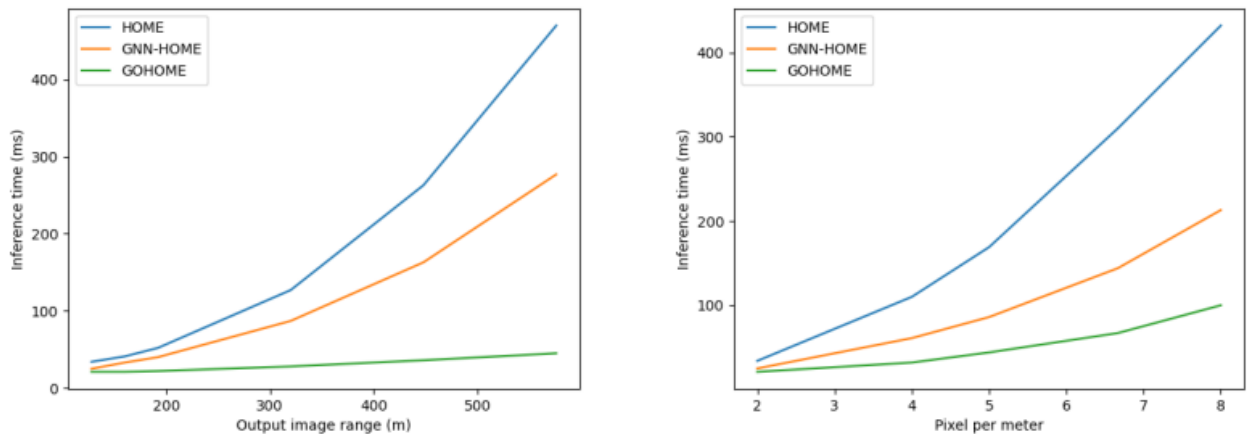


Рисунок 3.4 - Час виведення як функція від діапазону вихідного сигналу в метрах та пікселях на метр

Ми порівняли масштабування нашої моделі GOHOME на основі графів з масштабуванням моделі HOME на основі зображень щодо діапазону виводу та роздільної здатності. Рис. 3.4 підкреслює масштабування вихідного діапазону, де ми використовуємо фіксований вхідний діапазон 128 метрів та роздільну здатність 0,5 м на піксель. У той час як декодери CNN HOME та GNN-HOME призводять до квадратичного масштабування, растри смуг у поєднанні з рейтингом 20 найкращих смуг дозволяють масштабування, яке є навіть менш лінійним.

Час виведення з урахуванням кількості пікселів на метр, що є оберненою величиною до роздільної здатності. Хоча ефективна оптимізація згортки та постійних витрат призводить до близьких часів виведення для початкових 2 пікселів на метр, більш ефективне масштабування чітко проявляється для вхідних даних GNN та особливо для виходів растрів смуг. Оскільки роздільна здатність растрів смуг також масштабується із загальною роздільною здатністю, квадратична складність все ще застосовується, але зі значно меншим

коефіцієнтом, що дозволяє забезпечити реалістичні часи навчання та виведення для точніших роздільних здатностей.

На рис. 3.5 показано приклади багатоагентного прогнозування теплових карт з нашої моделі THOMAS.

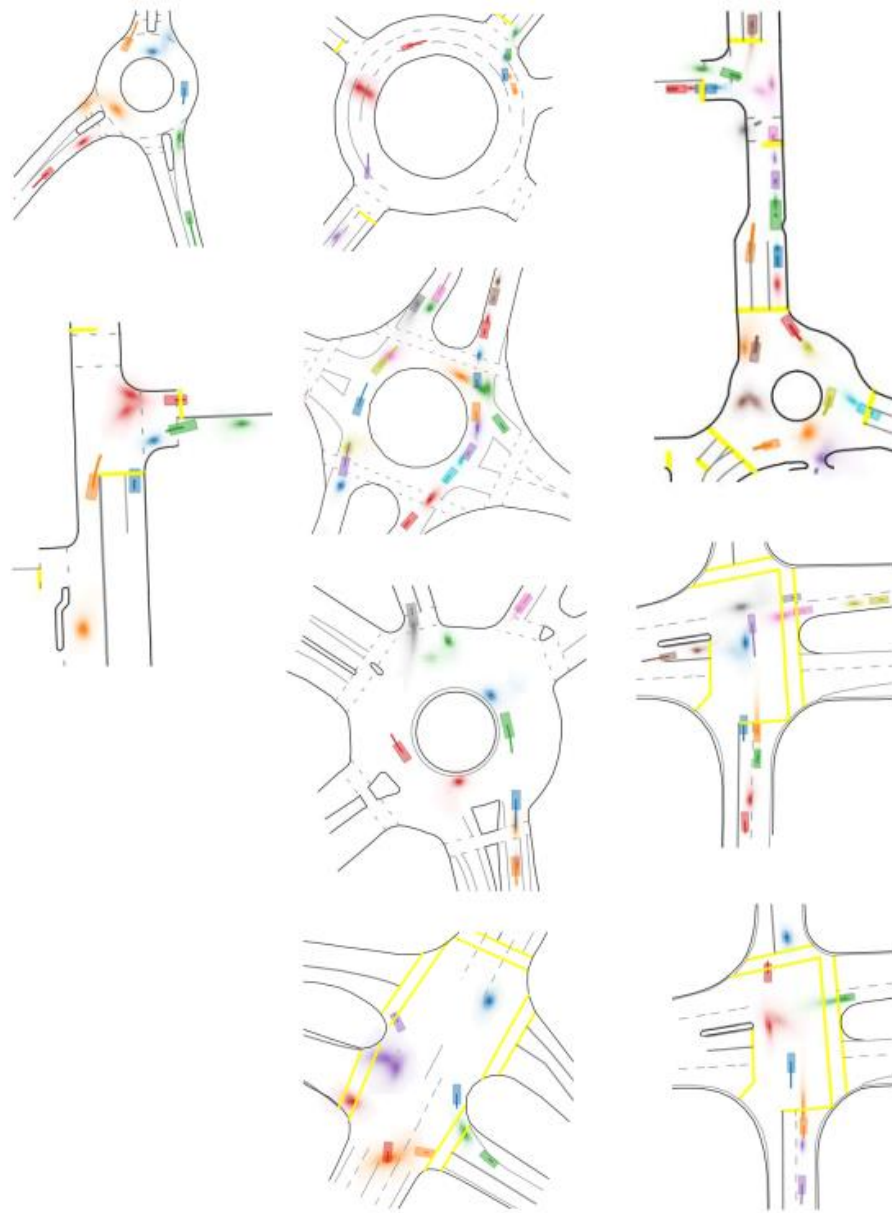


Рисунок 3.5 - Якісні приклади теплових карт, отриманих з нашої багатоагентної моделі. Усі теплові карти з однієї сцени є результатом одного прямого проходу в нашій моделі, яка передбачає всіх агентів одночасно. Ми використовуємо відповідні кольори для історії агента, поточного місцезнаходження та його прогнозованої майбутньої теплової карти (найкраще переглядати в кольорі).

Після представлення різних способів створення теплової карти, тепер ми розглянемо загальну концепцію прогнозування траєкторії на основі теплових карт та дослідження їхніх переваг порівняно з більш поширеними методами прогнозування руху.

Моделі теплових карт отримують найсучаснішу продуктивність, зокрема в minFDE1 та MR6, що можна інтерпретувати наступним чином:

Отримана мультимодальність для кожного пікселя забезпечує дуже хороший ранжування найімовірніших місць, куди потраплять агенти, що призводить до точного впорядкування прогнозованих режимів (низький minFDE1). Порівняно з більшістю режимів «Переможець отримує все» моделі, які мають проблеми з оцінкою того, який з їхніх кількох режимів є найбільш ймовірним, значення ймовірності, які наша модель приписує кожному ло-катіон перетворюється на глобально точну та порівнянну оцінку, демонструючи хороший прогнозований ймовірнісний розподіл.

Остаточні вибірккові моди забезпечують дуже добре покриття можливих досяжних позицій, досягаючи значень MR6, які не мають собі рівних у жодному іншому методі. Оскільки наше модельне прогнозування не обмежене та навчається в рамках, який наполегливо заохочує не ігнорувати жодного ймовірного пікселя, а також завдяки використанню алгоритму вибірки, розробленого з урахуванням продуктивності покриття, ми можемо забезпечити прогноз, який, хоча й не обов'язково є найточнішим з точністю до найближчого сантиметра, майже завжди матиме принаймні один передбачений мод поблизу еталонного значення.

Через мультимодальну природу прогнозів, ансамблювання моделей зазвичай є трудомістким при прогнозуванні траєкторії, оскільки неможливо визначити, яку модальність слід усереднювати, і навіть зіставлення за найкоротшою відстанню не гарантує, що два прогнози висвітлюють одне й те саме рішення і мають сенс усереднювати разом. З іншого боку, теплові карти ймовірностей є чудовим способом представлення інформації, що надходить з

різних джерел або моделей у спільній системі відліку, може бути усереднена разом без будь-яких припущень чи ризику збою режиму.

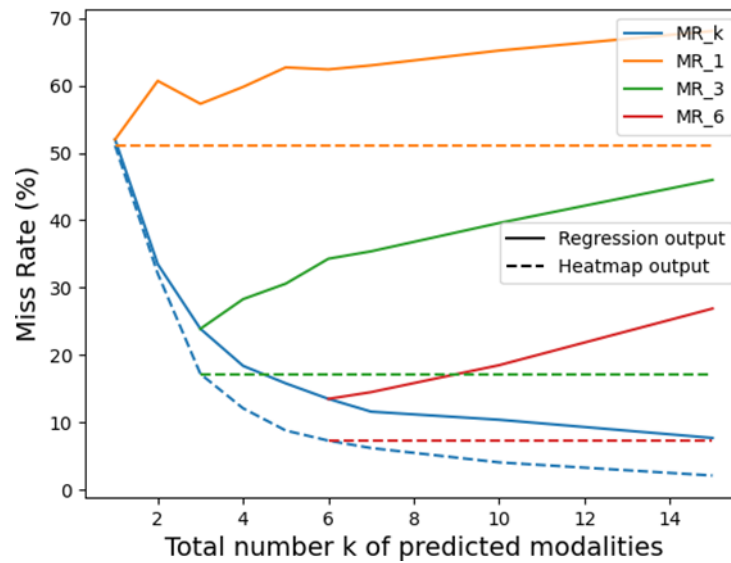


Рисунок 3.6 - Вплив максимальної кількості k прогнозованих модальностей, навчених на метрики з меншою фіксованою кількістю модальностей. Повні лінії – це результати регресійної вихідної моделі. Пунктирні лінії – це результат нашої моделі теплової карти. Ми показуємо коефіцієнт промахів для загальної кількості прогнозованих модальностей k (синій) та фіксованої кількості модальностей 1 (помаранчевий), 3 (зелений) та 6 (червоний). У випадку регресійного виходу, оскільки навчання кожної окремої модальності залежить від загальної кількості k навчених модальностей, метрики MR1, 3, 6 не є фіксованими та погіршуються, коли k збільшується.

4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ

4.1 Ергономічні проблеми безпеки життєдіяльності при роботі за комп'ютером

У сучасному світі все більше людей проводять значну частину свого робочого часу за комп'ютером, що призводить до ряду ергономічних проблем, які негативно впливають на безпеку та здоров'я людей. Неправильна позиція тіла, незручне розташування робочого місця, тривале сидіння та напружена робота з клавіатурою та мишею спричиняють м'язове напруження, біль у спині та напругу очей, що викликає дискомфорт та незручності.

Розташування робочого місця відіграє важливу роль у забезпеченні безпеки та здоров'я під час роботи за комп'ютером. Оптимальна висота столу та стільця є ключовим фактором для забезпечення комфорту і підтримки правильної позиції тіла. Стіл належної висоти, щоб лікті могли спокійно розташовуватися на клавіатурі, а стопи - на підлозі. Стілець обладнаний підтримкою для спини та належними регульованими підлокітниками. Клавіатура розташована на рівні ліктів, монітор - належним чином вирівняний перед очима, а миша - в зоні доступу для зап'ястя.

Правильне розташування робочого місця сприяє уникненню незручностей та проблем зі здоров'ям. Ергономічні пристрої підтримки є корисними для забезпечення комфорту та запобігання напрузі м'язів. Використання регульованих підлокітників та підставок для зап'ястя допомагає знизити напругу на м'язах і запобігти незручностям. Оптимальне розташування робочого місця повинне враховувати індивідуальні особливості користувача та забезпечувати комфортні умови для роботи.

Правильна позиція тіла та рухи є ключовими факторами для забезпечення комфорту та запобігання напрузі м'язів та незручностям.

Важливо пам'ятати про правильну позицію спини та уникати підгорблення або надмірного нахилу голови під час тривалої роботи за комп'ютером. Регулярні перерви для розтяжки та руху є важливими для підтримання здоров'я. Прості фізичні вправи, такі як розтягування шиї, плечей, рук і ніг, допомагають зняти напругу з м'язів та покращити кровообіг. Регулярні перерви дозволяють розслабитися і зберегти енергію для продуктивної роботи. Очі є одними з найбільш вразливих органів під час роботи за комп'ютером. Постійне спрямування погляду на екран призводить до напруження та втому очей.

Для зменшення негативного впливу необхідно використовувати екрани з антиблисковим покриттям, яке знижує відблиск та рефлексію світла. Також важливо налаштовувати яскравість та контрастність екрану для комфортного сприйняття. Щоб зменшити напругу на очі, необхідно робити перерви для відпочинку, фокусуючись на далеких предметах або виконуючи вправи для розслаблення очей.

Використання неправильної клавіатури та миші призводить до м'язового напруження і тунельного синдрому. Важливо використовувати ергономічні клавіатури та миші з додатковою підтримкою для зап'ястя та комфортною формою. Правильна позиція рук та зап'ястя під час роботи з ними має велике значення. Регулярні перерви для розтяжки рук та масажу зап'ястя допомагають уникнути негативних наслідків від тривалого використання клавіатури та миші. Розглядаючи ергономічні проблеми безпеки життєдіяльності при роботі за комп'ютером, варто зазначити, що їх вирішення є ключовим для забезпечення безпеки та здоров'я під час роботи.

Дотримання принципів ергономіки, налагодження робочого місця, правильна позиція тіла та виконання фізичних вправ сприяють покращенню безпеки та створенню здорового робочого середовища. Застосування ергономічних пристроїв підтримки, таких як регульовані підлокітники та підставки для зап'ястя, також сприяє комфорту та попередженню

незручностей. Загальною метою є створення безпечного та здорового робочого середовища для людей, які працюють за комп'ютером та виконують дослідження захищеності веб сервісу електронного навчання Atutor, що забезпечує збереження здоров'я та підвищення продуктивності працівників, сприяє запобіганню травмам та ергономічним проблемам, а також сприяє загальному комфорту та задоволенню від роботи.

Додатково, важливо зазначити, що належна освітленість приміщення також є важливим фактором, який впливає на комфорт та здоров'я під час роботи за комп'ютером. Потрібно забезпечити достатнє природне або штучне освітлення, яке не перевантажує очі. Розміщення робочого місця біля вікна або використання належної освітлювальної техніки допомагає забезпечити оптимальні умови освітлення.

Важливо також уникати відблисків на екрані, розташовуючи монітор під правильним кутом до джерел світла. Безпека та конфіденційність інформації також є важливим аспектом при роботі за комп'ютером. Важливо зберігати конфіденційні дані та захищати їх від несанкціонованого доступу. Використання паролів, шифрування даних та регулярне оновлення програмного забезпечення допомагає забезпечити безпеку інформації.

Крім того, необхідно усвідомлювати потенційні загрози з боку шкідливих програм та фішингових атак і приймати заходи для їх запобігання, такі як використання антивірусного програмного забезпечення та обережне відкривання електронних листів та посилань. Регулярне навчання та свідомість про важливість ергономіки та безпеки при роботі за комп'ютером є важливими.

Працівники повинні мати бути проінформовані про правильні методи роботи, виконання пауз і фізичних вправ, а також процедур безпеки. Організації можуть проводити навчання та інформаційні тренінги, щоб підвищити свідомість та забезпечити правильну поведінку під час роботи за комп'ютером.

Враховуючи всі ці аспекти, створення комфортного, безпечного та здорового робочого середовища є важливим завданням як для працівників, так і для роботодавців. Захист здоров'я та добробуту працівників при роботі за комп'ютером не тільки покращує їхню якість життя, але й сприяє збільшенню продуктивності та задоволення від роботи, що має позитивний вплив на всю організацію.

4.2 Організація безпечної роботи електроустаткування задіяного при роботі системи електронного навчання

Безпечна робота електроустаткування є ключовим елементом для забезпечення стабільної та безперебійної роботи системи електронного навчання. Це включає в себе правильне проектування, встановлення, експлуатацію та обслуговування електроустаткування. Дотримання стандартів і правил охорони праці допомагає мінімізувати ризики електричних ударів, пожеж та інших небезпек.

Перед встановленням електроустаткування необхідно ретельно оцінити його відповідність вимогам системи електронного навчання. Вибір обладнання повинен базуватися на таких критеріях: надійність та безпека: обладнання має відповідати стандартам якості та безпеки, мати сертифікати відповідності та бути розрахованим на довготривалу експлуатацію; відповідність технічним вимогам: обладнання повинно підтримувати необхідні технічні параметри, такі як напруга, потужність, тип з'єднання тощо; сумісність з іншими компонентами: всі компоненти системи повинні бути сумісні між собою, щоб уникнути збоїв та небезпек при експлуатації/

Під час встановлення необхідно дотримуватися таких заходів безпеки:
правильне заземлення: всі пристрої повинні бути заземлені відповідно до норм, щоб уникнути накопичення статичної електрики та можливості удару струмом; використання захисних пристроїв: встановлення автоматичних

вимикачів, пристроїв захисного вимкнення (ПЗВ) та інших захисних засобів є обов'язковим для забезпечення безпеки; професійний монтаж: монтаж повинен виконуватися кваліфікованими спеціалістами з дотриманням всіх норм і правил/

Для забезпечення безпечної експлуатації електроустаткування слід дотримуватися таких рекомендацій: регулярний контроль та технічне обслуговування: обладнання повинно регулярно перевірятися на наявність зношення, перегріву, пошкоджень проводів та інших дефектів. Технічне обслуговування повинно проводитися відповідно до інструкцій виробника; контроль температурного режиму: устаткування не повинно перегріватися. Необхідно забезпечити достатню вентиляцію та уникати розташування пристроїв поблизу джерел тепла; правильне використання: обладнання має використовуватися тільки за призначенням, з дотриманням інструкцій та рекомендацій виробника.

При обслуговуванні та ремонті електроустаткування необхідно дотримуватися таких заходів безпеки: відключення живлення: перед проведенням будь-яких робіт обладнання має бути відключене від джерела живлення; використання засобів індивідуального захисту (ЗІЗ): спеціалісти повинні використовувати відповідні ЗІЗ, такі як діелектричні рукавички, килимки, інструменти з ізоляційними покриттями; дотримання процедур: всі ремонтні роботи повинні проводитися відповідно до технічних інструкцій та нормативних документів;

Управління ризиками та навчання персоналу є невід'ємною частиною забезпечення безпеки електроустаткування: ідентифікація ризиків: постійний аналіз можливих ризиків та їх мінімізація є ключовим аспектом безпеки. Для цього проводяться регулярні оцінки стану обладнання та аналіз умов експлуатації; навчання та інструктажі: персонал повинен проходити регулярні навчання та інструктажі з питань безпечної роботи з електроустаткуванням. Це включає як базові знання, так і спеціальні навички для роботи з

конкретними видами обладнання; розробка та впровадження процедур безпеки: необхідно розробити детальні інструкції та процедури з безпеки, які повинні бути доступними для всіх працівників та регулярно оновлюватися. Таким чином, організація безпечної роботи електроустаткування задіяного при роботі системи електронного навчання вимагає комплексного підходу, що включає вибір надійного обладнання, правильне його встановлення, регулярне технічне обслуговування, навчання персоналу та управління ризиками. Дотримання цих заходів дозволяє забезпечити стабільну та безпечну роботу системи, що є критично важливим для ефективного функціонування освітнього процесу.

ВИСНОВКИ

У роботі, після огляду існуючих методів, що не пов'язані з навчанням, у рамках прогнозування траєкторій вивчено різні існуючі представлення та підходи до прогнозування руху на основі машинного навчання. Запропоновано вирішити проблему прогнозування траєкторії за допомогою теплових карт ймовірностей для полегшення мультимодальності та розроблено три різні способи створення цих теплових карт, оцінено їх продуктивність одна відносно одної та існуючих аналогів. Також запропоновано повний метод вибірки для вилучення фактичних траєкторій з цих теплових карт та вивчено переваги та недоліки цих методів теплових карт порівняно з іншими поширеними фреймворками. Розглянуто багатоагентне прогнозування для цих типів моделей теплових карт за допомогою методу самопланування.

ПЕРЕЛІК ПОСИЛАНЬ

1. Amirloo, E., Rasouli, A., Lakner, P., Rohani, M., Luo, J. Latentformer: Multi-agent transformer-based interaction modeling and trajectory prediction. arXiv preprint arXiv:2203.01880, 2022.
2. Azevedo, C., Gilles, T., Sabatini, S., Tsishkou, D. Exploiting map information for self-supervised learning in motion forecasting. arXiv preprint arXiv:2210.04672, 2022.
3. Bahari, M., Saadatnejad, S., Rahimi, A., Shaverdikondori, M., Shahidzadeh, A. H., Moosavi-Dezfooli, S.-M., Alahi, A. Vehicle trajectory prediction works, but not everywhere. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022. C. 17123–17133.
4. Gesnouin, J. Analysis of pedestrian movements and gestures using an on-board camera to predict their intentions. PhD thesis. Université Paris sciences et lettres, 2022.
5. Gilles, T., Sabatini, S., Tsishkou, D., Stanciulescu, B., Moutarde, F. Thomas: Trajectory heatmap output with learned multi-agent sampling. In: International Conference on Learning Representations (ICLR), 2022.
6. Gomes, I., Wolf, D. A review on intention-aware and interaction-aware trajectory prediction for autonomous vehicles. 2022.
7. Hasan, I., Liao, S., Li, J., Akram, S. U., Shao, L. Pedestrian detection: Domain generalization, CNNs, transformers and beyond. arXiv preprint arXiv:2201.03176, 2022.
8. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R. Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022. C. 16000–16009.
9. Horváth, D., Erdős, G., Istenes, Z., Horváth, T., Földi, S. Object detection using sim2real domain randomization for robotic applications. IEEE Transactions on Robotics, 2022.

10. Huang, Z., Mo, X., Lv, C. Multi-modal motion prediction with transformer-based neural network for autonomous driving. In: 2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022. С. 2605–2611.
11. Ivanovic, B., Lee, K.-H., Tokmakov, P., Wulfe, B., McIlister, R., Gaidon, A., Pavone, M. Heterogeneous-agent trajectory forecasting incorporating class uncertainty. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2022. С. 12196–12203.
12. Karle, P., Geisslinger, M., Betz, J., Lienkamp, M. Scenario understanding and motion prediction for autonomous vehicles – review and comparison. IEEE Transactions on Intelligent Transportation Systems, 2022.
13. Mahjourian, R., Kim, J., Chai, Y., Tan, M., Sapp, B., Anguelov, D. Occupancy flow fields for motion forecasting in autonomous driving. IEEE Robotics and Automation Letters, 2022.
14. А.Г. Микитишин, М.М. Митник, П.Д. Стухляк, В.В. Пасічник Комп'ютерні мережі. Книга 1. [навчальний посібник] (Лист МОНУ №1/11-8052 від 28.05.12р.) - Львів, "Магнолія 2006", 2013. – 256 с.
15. А.Г. Микитишин, М.М. Митник, П.Д. Стухляк, В.В. Пасічник Комп'ютерні мережі. Книга 2. [навчальний посібник] (Лист МОНУ №1/11-11650 від 16.07.12р.) - Львів, "Магнолія 2006", 2014. – 312 с.
16. Микитишин А.Г., Митник, П.Д. Стухляк. Комплексна безпека інформаційних мережевих систем: навчальний посібник – Тернопіль: Вид-во ТНТУ імені Івана Пулюя, 2016. – 256 с.
17. Микитишин А.Г., Митник М.М., Стухляк П.Д. Телекомунікаційні системи та мережі : навчальний посібник для студентів спеціальності 151 «Автоматизація та комп'ютерно-інтегровані технології» – Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2017 – 384 с.
18. Введення в комп'ютерну графіку та дизайн: Навчальний посібник для студентів спеціальності 174 "Автоматизація, комп'ютерно-інтегровані

технології та робототехніка"/Укладачі: О.В. Тотосько, П.Д. Стухляк, А.Г. Микитишин, В.В. Левицький, Р.З. Золотий - Тернопіль: ФОП Паляниця В.А., 2023 - 304с. <http://elartu.tntu.edu.ua/handle/lib/41166>.

19. Пилипець М. І. Правила заповнення основних форм технологічних документів : навч.-метод. посіб. / Уклад. Пилипець М. І., Ткаченко І. Г., Левкович М. Г., Васильків В. В., Радик Д. Л. Тернопіль : ТДТУ, 2009. 108 с. <https://elartu.tntu.edu.ua/handle/lib/42995>.