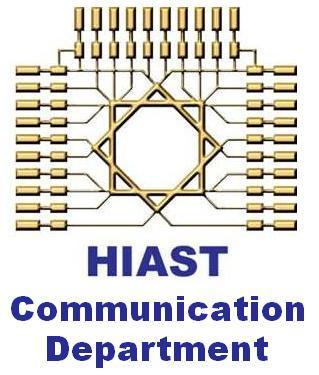
**الجمهـوريـة العربيــة الســـورية**

**المعهد العالي للعلوم التطبيقية والتكنولوجيا**

**قسـم المعلوميات**

**العام الدراسي 2024/2025**

**مشروع تخرج**

أعد لنيل درجة الإجازة في هندسة الاتصالات

نظام قيادة يعتمد على النماذج اللغوية الكبيرة

تقديم

**أحمد علي**

إشراف

**د. عمر حمدون**

**م.مرح حسن**

**ما.نديم الرباعي**

30/8/2025

**أهدي هذا العمل إلى طلاب المعهد العالي للعلوم التطبيقية والتكنولوجيا.**

**ضع حكمتك المفضلة هنا (اختياري) مع ذكر قائلها.**

**كلمة شكر**

**أتقدم بالشكر إلى رئيس قسم الاتصالات الدكتورة أميمة دكاك على دعمها لهذه المبادرة وعلى ملاحظاتها القيمة.**

**أيمن السـواح**

**الخلاصة**

(وتتضمن الهدف من المشروع) يهدف هذا المشروع إلى تصميم نموذج لتقارير مشاريع طلاب المعهد العالي للعلوم التطبيقية والتكنولوجيا وتقديم بعض النصائح والتوجيهات حول كيفية كتابة التقرير بطريقة علمية منهجية. إن التزام الطلاب بنموذج موحد للتقارير يوفر عليهم عناء البحث عن التنسيق المناسب كما أنه يسهل عملية الرجوع إلى التقارير وتوثيقها ورقياً وإلكترونياً. أما احترام قواعد الكتابة العلمية الصحيحة فهو ينعكس إيجاباً على إبراز القيمة الحقيقية للعمل الهندسي المنجز وإفساح الطريق أمام الاستفادة منه لاحقاً.

**Abstract**

Translate your abstract here.

**المحتويات**

[**الفصل الأول 1**](#_o4r8m0lsb22p)

[**التعريف بالمشروع 1**](#_lifq2sdd0sxe)

[1.1- مقدمة 1](#_kkpc99tsuoqi)

[2.1- هدف المشروع 1](#_69lco52po1b1)

[3.1- المتطلبات الوظيفية 1](#_t2tht9c2q33i)

[4.1- المتطلبات الغير وظيفية 2](#_twdxl2d7lf7b)

[**الفصل الثاني 2**](#_85sh382ezken)

[**الدراسة النظرية 3**](#_uzuk88mytfmn)

[1.2- الإسقاط متعدد الطبقات (Multilayer perceptron) 3](#_qd4lgmwrbelw)

[2.2- آليات الانتباه (Attention mechanisms) 3](#_yrpyb3aqn7wx)

[1.2.2- مزايا الانتباه 3](#_7uw4b4ia11xc)

[3.2- المحولات (Transformer) 4](#_dnwdhjd26dy4)

[4.2- النماذج اللغوية الكبيرة (Large language model) 4](#_2xtmpv6pqz8s)

[5.2- التحسين الفعال للمعلمات (Parameter efficient fine tuning) 4](#_teftx5igu099)

[6.2- التكيف منخفض الرتبة (Low rank adaption) 4](#_h3zgso7r3eu6)

[7.2- الحاوية (Container) 5](#_amvz7aqicbm9)

[8.2- بيئة المحاكاة (Simulator) 5](#_iyqu3trdwomu)

[**الفصل الثالث 6**](#_wmvbgcb580jm)

[**الدراسة المرجعية 6**](#_qvi24zjpfvdv)

[1.3- مقدمة 6](#_fis2r3o1fchp)

[2.3- إطار عمل OmniDrive للقيادة الذاتية 6](#_6cqd6v7ytohq)

[1.2.3- بيانات التدريب 7](#_s7s47ewq8f32)

[2.2.3- التدريب 8](#_liik8t1q5dxo)

[3.2.3- معيار التقييم 9](#_m2wr9jtqoryx)

[4.2.3- النتائج 9](#_xye9keve4jei)

[3.3- تعزيز وكلاء القيادة الذاتية المعتمدة على النماذج اللغوية الكبيرة من خلال تجارب مجسدة واجتماعية DriVLMe 10](#)

[1.3.3- مُرمِّز الفيديو 10](#_3ta77f8sp6yj)

[2.3.3- العمود الفقري لنظام إدارة التعلم (LLM) 11](#_8h7kw5z24x60)

[3.3.3- وحدة تخطيط المسار (ملاح GPS) 11](#_kwklo2z28j8s)

[كيفية عمل الوحدات الثلاثة 12](#)

[4.3.3- بيانات التدريب 12](#_1hfnq4xv2yc4)

[5.3.3- التدريب 12](#_itkhd184nha4)

[6.3.3 - معايير التقييم 13](#_6tg9onxhdcoi)

[7.3.3- النتائج 13](#_kos7e71o1r02)

[التحديات والمشاكل 14](#_inrt2rcfxk7f)

[4.3- نموذج تتبع سيارة ذاتية القيادة شبيه بالإنسان باستخدام التعلم المعزز العميق 14](#_5uhz8za02to7)

[1.4.3- بنية النموذج المقترح 15](#_hldmqzljmf3r)

[2.4.3- آلية عمل الخوارزمية 15](#_qfly6nrjrfo)

[3.4.3- البيانات المستخدمة 17](#_rf8jhmgacc6h)

[4.4.3- التدريب 17](#_3l2l5aeqseuh)

[5.4.3- معايير التقييم 17](#_kcwa2zcydt3x)

[6.4.3- النتائج 18](#_rl1drfd0nyko)

[المشاكل التي واجهها هذا البحث 18](#_kwlbadwbu8dt)

[2. السلامة (Safety) لم يتم التعامل معها بشكل مباشر 18](#_fvs8zaaq39nr)

[5.3- القيادة باستخدام النماذج اللغوية الكبيرة: دمج متجه مستوى الكائن في القيادة الذاتية القابلة للتفسير 18](#_gnpm9rrv8w5q)

[1.5.3 مرمّز الأشعة (Vector Encoder) 18](#_kv78z7c3p0z0)

[2.5.3 مشكّل الأشعة (Vector-Former) 19](#_5vwnet4xkgax)

[3.5.3 النموذج اللغوي مع التكيف المنخفض (LLM with LoRA) 19](#_n66m4e6v1uw4)

[4.5.3- بيانات التدريب 19](#_rwi1hsakfk0i)

[5.5.3- تدريب النموذج 20](#_6ydpb1oj74lj)

[6.5.3- معايير التقييم 21](#_eh45j9xjnx7j)

[7.5.3- النتائج 21](#_ygrkrfedj3sa)

[8.5.3- المشاكل 22](#_i6q8hb3tf1sr)

[6.3- الخاتمة 22](#_ny34duupbsye)

[**الفصل الرابع 23**](#_4gmx9ig0xu2a)

[**المنهجية المقترحة 23**](#_pppfvretxjo2)

[1.4- مقدمة 23](#_5ravpcnh1oo3)

[2.4- النموذج المقترح 23](#_r5rctz307wnh)

[3.4- مكونات النموذج 24](#_7surukj3ymi1)

[1.3.4- مرمّز الأشعة 24](#_y1vak5i5o6zv)

[2.3.4- مشكل الأشعة 26](#_ofplvl3u22qr)

[3.3.4-النموذج اللغوي مع التكيف المنخفض الرتبة 28](#_36sqmtug41gy)

[4.4- معايير التقييم المتبعة 29](#_nu9fbmyjyuiu)

[**الفصل الخامس 31**](#_og62dyay9ekx)

[**تصميم النظام 31**](#_5t0281nnmghs)

[1.5- مقدمة 31](#_7w7oar29a52q)

[2.5- بيئة المحاكاة 31](#_7pz3ku3x1jfb)

[3.5- النموذج اللغوي 31](#_sjw3t923z80g)

[4.5- الواجهات الخلفية Back-End 31](#_u1aznkroh24a)

[**الفصل السادس 32**](#_2lp8o7c3irqe)

[**الأدوات المستخدمة 32**](#_2zhy9zwjspnp)

[1.6-Docker 32](#_c34lpapo44sr)

[2.6-MongoDB 32](#_spnskt5yjadn)

[3.6-GitHub 32](#_6h6qunkjelk4)

[4.6-Hugging Face 32](#_8eddld34smn1)

[5.6- Pytorch 33](#_z94yvfvntrzs)

[6.6- FastAPI 33](#_844ebvtgkrv1)

[7.6-React 33](#_7y7vpg83hyh6)

[8.6- محاكي CARLA 34](#_30x33l665r8j)

[9.6- DeepSeek 34](#_danvcwyl3bl2)

**قائمة الأشكال**

[الشكل 1- مشكلة طريقة الترقيم الآلية المعكوسة للقوائم المتعددة المستويات. 7](#_e1gsovagsqjg)

[الشكل 2- أيقونة البرنامج المستخدم في كتابة هذا التقرير (شكل غير مفيد!). 11](#_a8l339sfki3w)

[الشكل 3- مخطط صندوقي لنظام اتصالات. 11](#_w3phzquas426)

[الشكل 4- معدل البيانات. 12](#_icpmnczhdzws)

[الشكل 5- مثال على شكل مبالغ في حجمه. 13](#_y0mt9r8vmd3b)

[الشكل 6- مثال على شكل صغير جداً بحيث تصعب قراءته. 13](#_1o0ctjoxeq3k)

[الشكل 7- مثال على شكلين غير متجانسين. 14](#_mg1g5ey6o7gz)

[الشكل 8- مثال على شكلين متجانسين. 15](#_qeh73kep0gbp)

[الشكل 9- مثال على إدخال تعديلات على شكل جاهز (الأيسر) للحصول على شكل مناسب أكثر (الأيمن). 15](#_z7hqja9um48o)

**قائمة الجداول**

[الجدول 1- رموز وقيم العناصر المستخدمة في الدارة. 16](#_ss67cb34ebh4)

**الاختصارات**

|  |
| --- |

| **CIDEr** | مقياس أكثر تطورًا يعتمد على الدلالة اللغوية. |
| --- | --- |

| **SPICE** | يقيس التطابق في المعنى والعلاقات بين الكائنات في النص المولد والمرجعي |
| --- | --- |

Root Mean Square Percentage Error (RMSPE)**الرموز**

d المسافة بالأمتار.

α فقد المسار Path-loss.

# الفصل الأول

# التعريف بالمشروع

*يتضمن هذا الفصل التعريف بالمشروع ومتطلباته*

## 1.1- مقدمة

شهد العالم في السنوات الأخيرة تطوراً تقنيًا هائلًا في مجالات الذكاء الاصطناعي، وكان من أحد أهم نتائجه ظهور السيارات ذاتية القيادة، حيث تعتمد على أنظمة متقدمة من الحساسات، والكاميرات، والخوارزميات الذكية، لقيادة السيارة دون تدخل بشري مباشر. تهدف هذه السيارات إلى تحسين السلامة المرورية، وتخفيف الازدحام، وتقليل الحوادث الناتجة عن الخطأ البشري. ومع استمرار الأبحاث والتجارب، تقترب السيارات ذاتية القيادة من أن تصبح جزءاً لا يتجزأ من حياتنا اليومية، ولكنها تواجه مشاكل التعامل مع المواقف الجديدة واحتمال تعرضها للاختراق بالإضافة إلى استمرار تشكيك الناس في أمان هذه السيارة بعد وقوع العديد من الحوادث.

## 2.1- هدف المشروع

يهدف مشروعنا إلى توظيف التطورات الأخيرة في مجال نماذج اللغات الكبيرة لبناء نموذج قادر على اتخاذ قرارات القيادة انطلاقاً من بيانات مهيكلة رقمية، و توليد أسئلة وأجوبة لتبرير القرارات المتخذة ، وتقديم هذا المشروع من خلال بيئة محاكاة مخصصة لمهام القيادة الذاتية تقدم لنا ميزات جمع البيانات الخاصة من أجل التدريب والاختبار.

## 3.1- المتطلبات الوظيفية

1. استعراض
2. التنبؤ بالإجراء المناسب للقيادة
3. توصيف مشهد القيادة بشكل تفصيلي
4. تسجيل جميع المشاهدات التي يقوم بها النموذج
5. تتبع مسار السيارة من خلال خريطة لمسار السيارة
6. عرض إحصائيات زمن الاستجابة للنموذج
7. تحليل وعرض مؤشرات الأداء الرئيسية
8. عرض إحصائيات السرعة والمكابح والالتفاف
9. الفلترة الزمنية للبيانات والإحصائيات

## 4.1- المتطلبات الغير وظيفية

1. النظام المقدم سيكون آمن:
   1. غير قابل للاختراق فالنظام مثبت لكل سيارة على حدة أي النظام مكون من مستخدم واحد
   2. يقدم تسجيل لجميع العمليات المتخذ
2. قابلية الصيانة:
   1. سهولة تحديث وتطوير النظا

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# الدراسة النظرية

*سنقدم في هذا الفصل المفاهيم النظرية التي تم استخدامها والاعتماد عليها لإتمام هذا العمل*

## 1.2- الإسقاط متعدد الطبقات (Multilayer perceptron)

تتكون من طبقات كثيفة متصلة بالكامل، تحول بيانات الإدخال من بعد إلى آخر ، وتتألف من ثلاث طبقات على الأقل، طبقة دخل وطبقة خرج وطبقة مخفية واحدة على الأقل، وتستخدم من أجل تبسيط العلاقات المعقدة واللاخطية بالإضافة إلى استخدامها في حل مسائل التصنيف والانحدار

## 2.2- آليات الانتباه (Attention mechanisms)

هي نوع من الشبكات العصبية التي تساعد النموذج في التركيز على أجزاء محددة من بيانات الإدخال، ويتم ذلك من خلال تعيين أوزان لعناصر مختلفة في المدخلات، مما يساعد النموذج في تحديد الأجزاء الأكثر أهمية.

### 1.2.2- مزايا الانتباه

* تسمح بالحساب المتوازي (Parallelization) : على عكس الشبكات العصبونية المتكررة (RNN) ، التي تعالج البيانات بشكل متسلسل، تسمح آليات الانتباه بالمعالجة المتوازية للدخل . وهذا يسرع بشكل كبير من الحوسبة ويجعلها أكثر كفاءة، وخاصة بالنسبة لمجموعات البيانات الكبيرة.
* التعامل مع التبعيات طويلة المدى (Long-Range Dependencies) : تسمح للنماذج بالتقاط العلاقات بين العناصر البعيدة ضمن سلسلة الدخل بشكل أكثر فعالية من النماذج التي تعتمد فقط على الهياكل المتكررة (RNN) .وهذا أمر بالغ الأهمية للمهام حيث يكون السياق من الأجزاء السابقة من التسلسل مهمًا لفهم الأجزاء اللاحقة.

## 3.2- المحولات (Transformer)

هي بنية شبكات عصبونية تستخدم في معالجة البيانات المتسلسلة، مثل معالجة اللغات الطبيعية وفصل الكلام، وتستفيد من آلية الانتباه لتقييم أهمية الأجزاء المختلفة من تسلسل الإدخال عند معالجته، مما يتيح فهم السياق والعلاقات بشكل أكثر فعالية من النماذج السابقة مثل الشبكات العصبية المتكررة (RNN) لذلك تستخدم في الترجمة والتلخيص والإجابة عن الأسئلة.

## 4.2- النماذج اللغوية الكبيرة (Large language model)

هي أنظمة ذكاء اصطناعي متقدمة تم تدريبها على كميات هائلة من النصوص المكتوبة بلغات مختلفة، وهي تعتمد على المحولات، والتي تُدرب بطريقة التعلم ذاتي الإشراف على مليارات الرموز، حيث يستخدم هذا التدريب لتعلم تمثيلات دقيقة النحو والدلالة والسياقات المعقدة، مما يمكن النموذج من أداء مهام لغوية متنوعة مثل الترجمة، والتلخيص، والإجابة عن الأسئلة.

أمثلة على نماذج لغوية كبيرة:

GPT(OpenAI),LLaMA(Meta),PaLM(Google)

## 5.2- التحسين الفعال للمعلمات (Parameter efficient fine tuning)

هي مجموعة من التقنيات التي تسمح بتخصيص النماذج اللغوية لمهام محددة عن طريق تعديل جزء صغير فقط من المعلمات، مع إبقاء معظم النموذج الأصلي مجمداً، مما يقلل التكلفة الحسابية والتخزينية مع الحفاظ على الأداء، وتستخدم عندما يكون حجم النموذج كبير جداً ويصعب إعادة تدريبه بشكل كامل وذلك لأسباب مختلفة:

* محدودية الموارد
* تخصيص النموذج مع الحفاظ على الأداء السابق
* تجريب سريع لتقييم أداء النموذج على بيانات جديدة دون تكلفة عالية

## 6.2- التكيف منخفض الرتبة (Low rank adaption)

هي تقنية من ضمن التحسين الفعال للمعلمات (PEFT) وتهدف إلى تعديل النماذج اللغوية الكبيرة بطريقة فعالة من حيث الحساب والتخزين، وتستخدم لتخصيص مهام النموذج اللغوي ضمن مجال معين أو من أجل المشاريع الصغيرة بسبب محدودية الموارد.

**فوائد LoRA**

* تقليل عدد المعلمات القابلة للتدريب بنسبة تصل إلى 10000 مقارنة بالتدريب الكامل
* المحافظة على المعرفة الأساسية من خلال تجميد الأوزان
* خفض تكلفة التخزين
* لا يحتاج إلى موارد كبيرة

## 7.2- الحاوية (Container)

هي عبارة عن مجموعة من الحزم المعزولة عن الجهاز المحلي الخاص بك، يمكن تشغيل واستخدام البرامج الموجودة داخل الحاوية دون أن تتعارض الحزم داخل الحاوية مع الحزم المنصبة بالفعل على جهازك

## 8.2- بيئة المحاكاة (Simulator)

هي نظام برمجي متكامل يهدف إلى محاكاة الظروف الواقعية بشكل افتراضي، مما يتيح للمستخدمين اختبار وتحليل الأنظمة والعمليات في بيئة آمنة ومنخفضة التكلفة. تُستخدم المحاكيات على نطاق واسع في مجالات مختلفة مثل القيادة الذاتية، والروبوتات، والطيران، وتطوير الألعاب، والهندسة، والتدريب والتعليم. تتيح بيئة المحاكاة اختبار سيناريوهات متنوعة قد يصعب تنفيذها أو تكون خطيرة أو باهظة التكلفة في الواقع، مما يساهم في توفير الوقت والجهد والموارد، بالإضافة إلى زيادة الدقة والفعالية في عملية تطوير الأنظمة واختبارها. من أبرز فوائد المحاكاة هي إمكانية تحديد الأخطاء والمشاكل مبكرًا، تحسين التصميمات بشكل سريع، وتوفير تجربة تعليمية وتدريبية غنية دون المخاطرة أو التعرض لأي ضرر فعلي.

# الفصل الثالث

# الدراسة المرجعية

*يعرض هذا الفصل الأبحاث المرتبطة بالعمل المقدم*

## 1.3- مقدمة

إن مجال أنظمة القيادة المعتمدة على النماذج اللغوية الكبيرة مجال جديد، وذلك بسبب التطور الكبير في النماذج اللغوية الكبيرة التي أظهرت آفاقاً جديدة والتي توفر لنا التفكير المنطقي في المسائل ومعالجة أنواع مختلفة من المعلومات كالصور والنصوص وبالإضافة إلى أهمية التواصل مع المركبات ذاتية القيادة حيث أن التحدث إلى السيارة يساعد في ضمان سلامة الجميع وتزيد من ثقة الناس بها نتيجة إعلام الراكب بالتبرير المنطقي للفعل الذي اتخذته المركبة .

سنستعرض في الفقرات التالية الأفكار التي تم الاعتماد عليها ضمن هذا المجال

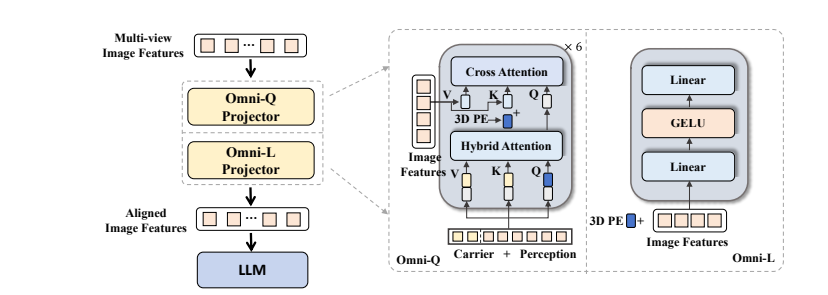
## 2.3- إطار عمل OmniDrive للقيادة الذاتية

يهدف هذا الإطار إلى تعزيز عملية اتخاذ القرار من خلال تقييم سيناريوهات القيادة المحتملة ونتائجها وذلك من خلال سؤال النموذج اللغوي عن الفعل الواجب اتخاذه ويقوم النموذج بالإجابة والتي هي عبارة عن الفعل الواجب اتخاذه والسبب هذا الفعل.

بدايةً تم فرز البيانات إلى مجموعات بما يضمن اختيار البيانات الأكثر تمثيلاً دلالياً والتي تغطي عناصر حركة المرور الديناميكية والثابتة وكل مجموعة تعكس سلوك قيادة معين مثل الانعطاف يساراً أو التباطؤ....

يتكون هذا الإطار من مكونين أساسيين وهما Omni-Q و Omni-L.

Omni-Q: هو نموذج VLM مع تعديل على Q-Former لتعزيز قدرات تحديد المواقع في نماذج VLM من خلال توظيف أفكار من نماذج الإدراك ثلاثية الأبعاد مثل Stream PETR في بنية Q-Former. مما يحسن الفهم المكاني والربط بشكل أفضل بتوليد اللغة.

Omni-L: هذا المكون مسؤول عن محاذاة اللغة البصرية في النموذج اللغوي حيث يقوم بربط الميزات البصرية متعددة المشاهد بنموذج اللغة الكبير من خلال MLP بسيط لإسقاط الميزات البصرية في مساحة تضمين نموذج اللغة الكبير أي يمكن النموذج من معالجة المعلومات البصرية الواردة من وجهات نظر متعددة في وقت واحد.

الشكل (3-1): طبقات إطار OmniDrive

### 1.2.3- بيانات التدريب

تم استخدام مجموعة بيانات NuScences وهي عبارة عن:

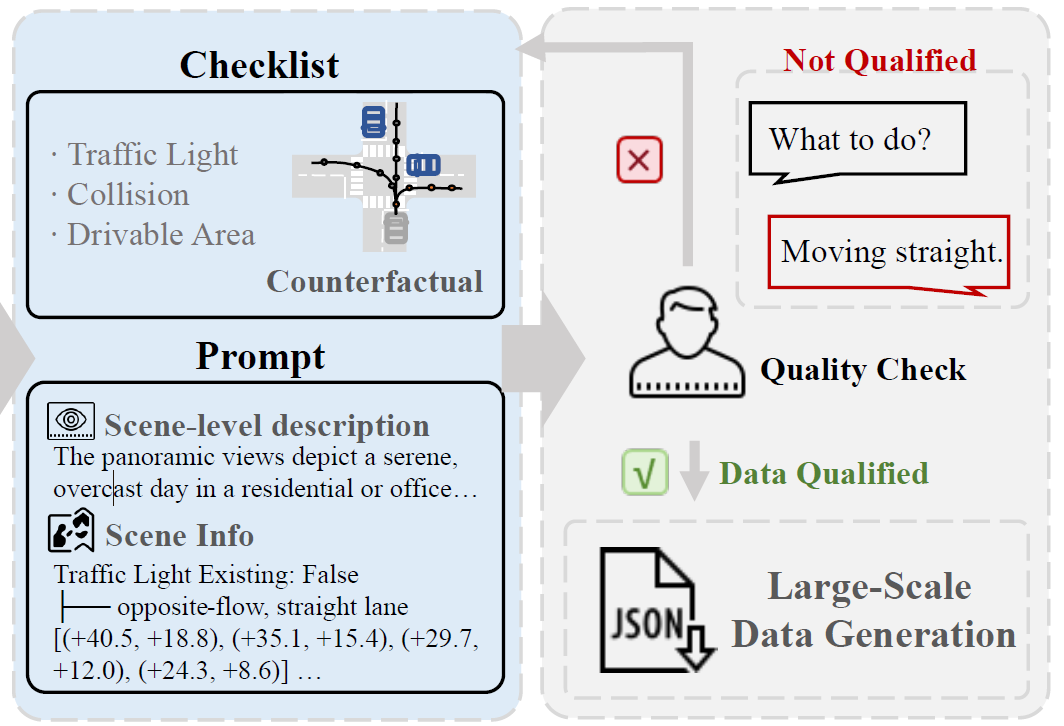
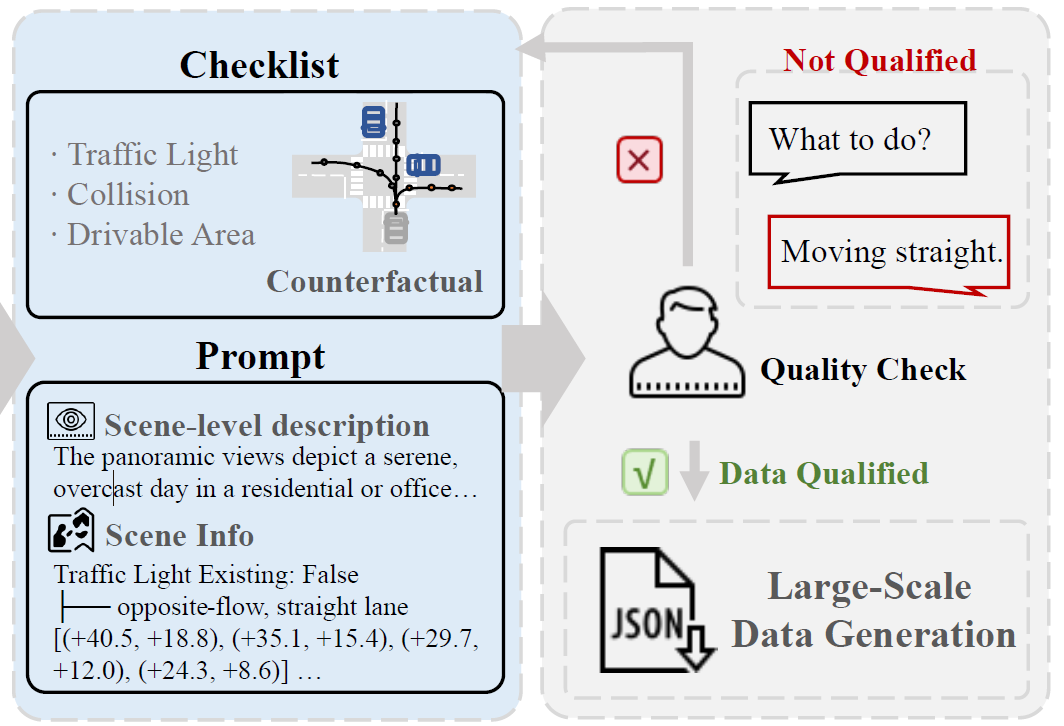
* صور من عدة كاميرات: أمام، خلف، يمين، يسار
* كائنات ثلاثية الأبعاد: سيارات، مشاة، إشارات مرور، مخاريط "أقماع"
* حالة السيارة: السرعة، الموقع، التوجه

بالإضافة إلى البيانات المولدة بالاستدلال والتي هي عبارة عن "ماذا سيحدث لو سلكت السيارة طريقاً مخلتفاً" أو "ماذا لو تجاهلت الإشارة الحمراء"، ويتم إضافة النتيجة لهذا السؤال المقترح مثلاً:

* الاستمرار في التوقف => آمن
* الانعطاف يساراً => خطر
* قطع الإشارة الحمراء => خطر

بالإضافة إلى ذلك تم إضافة الأسئلة والأجوبة باستخدام GPT-4، وأنواع الأسئلة هي:

* أسئلة وصفية: ما الذي يظهر في المشهد؟
* أسئلة الانتباه: ما الكائنات التي يجب الانتباه عليها؟
* أسئلة الاستدلال: ماذا سيحدث إذا تسارعت السيارة؟
* أسئلة تخطيط: ما هو أفضل مسار يمكن اتباعه؟

الشكل(3-2): توضيح لـ تدريب النظام

### 2.2.3- التدريب

تم التدريب على مرحلتين:

* التدريب المسبق ثنائي الأبعاد: يتم تزويد مكونات ربط الصورة بالنموذج اللغوي بفهم أساسي لكيفية ربط المعلومات البصرية بالأوصاف والمفاهيم النصية

البيانات: مجموعات كبيرة من الصور الثنائية الأبعاد المقترنة بنص

الهدف: توليد النصوص استناداً إلى الصور

النتيجة: تهيئة الأوزان للمكونين Omni-L وOmni-Q

* الضبط الدقيق ثلاثي الأبعاد: يتم تكييف النموذج المدرب مسبقاً لفهم تعقيدات البيئات ثلاثية الأبعاد

البيانات: مجموعات بيانات تحتوي على صور متعددة العرض من سيناريوهات القيادة مرفقة بنصوص

الهدف: الإجابة على أسئلة "ماذا لو" أو "ماذا سيحدث" حول مشهد القيادة والتنبؤ بمسار الحركة

النتيجة: يصبح النموذج متخصصاً في فهم المشاهد القيادة ثلاثية الأبعاد والتفكير فيها

### 3.2.3- معيار التقييم

تم تقييم النموذج في ثلاث جوانب رئيسية:

* الاستدلال: لتحديد ما إذا كان النموذج قادراً على تحديد النتائج الخطيرة لمسارات القيادة التي تم توليدها بالاستدلال، وتم استخدام Precision و Recall
* تقييم التخطيط المفتوح: تقييم قدرة النموذج على توليد مسارات قيادة قريبة من المسارات الصحيحة ،وتم استخدام:
  + L2 لقياس الفرق بين المسار الحقيقي والمسار المتوقع
  + معدل اصطدام من أجل حساب نسبة المسارات التي قد تؤدي إلى اصطدام
  + معدل التقاطع من أجل حساب نسبة المسارات التي تؤدي إلى الخروج عن الطريق الحقيقي مثلاً القيادة على الرصيف
* تقييم جودة الأسئلة و الأجوبة المقترحة وتم استخدام
  + تقييم GPT والذي يحسب جودة الإجابة
  + تقييم اللغة ويقيس دقة التوليد اللغوي
  + تقييم الترابط ويقيس نسبة الكائنات المتوقعة الصحيحة المذكورة في الإجابة

### 4.2.3- النتائج

| المعيار | Precision | Recall | L2 | Collision | Language | Final score |
| --- | --- | --- | --- | --- | --- | --- |
| النتيجة | 72.1% | 71.3% | 0.7 | 0.06% | 73.2 | 58% |

الجدول (3-1): نتائج معيار القسم

## 3.3- تعزيز وكلاء القيادة الذاتية المعتمدة على النماذج اللغوية الكبيرة من خلال تجارب مجسدة واجتماعية DriVLMe

ظهر هذا البحث لمعالجة مشاكل التعقيد الحقيقي للقيادة في العالم الواقعي الذي لم تعالجه الدراسات السابقة والتي تعالج مواقف بسيطة ومحدودة فقط ولم تعالج مشكلة التعامل مع الرحلات الطويلة أو الانخراط في محادثات طبيعية ومرنة مع الركاب أو التعامل مع الأحداث المفاجئة مثل تغيير الراكب لوجهته في منتصف الرحلة.

طور الباحثون هذا النموذج DriVLMe لفهم كل ما تراه السيارة وما يقوله البشر بحيث يمكن التواصل الفعال بين البشر والمركبة مما يساعد المركبة على إدراك محيطها والتنقل

يتكون وكيل DriVLMe من ثلاثة أجزاء رئيسية تعمل معًا:

مُرمِّز الفيديو: يُشبه "عيني" الوكيل. يأخذ لقطات فيديو خام لما يراه الوكيل ويحولها إلى تمثيل رقمي مُدمج يفهمه "الدماغ".

العمود الفقري لـ LLM (نموذج اللغة الكبير): يُشبه "دماغ" الوكيل يُعالج المعلومات المرئية (من مُرمِّز الفيديو) وأي تعليمات حوارية (مثل "اذهب إلى المكتبة") لاتخاذ القرارات.

وحدة تخطيط المسار: تُشبه "ملّاح GPS" للوكيل. عندما يحتاج الوكيل إلى الوصول إلى معلم مُحدد، تُحدد هذه الوحدة أفضل مسار باستخدام خريطة.

لنتعمق في كل جزء:

## 1.3.3- مُرمِّز الفيديو

الغرض: أخذ سجل لما شاهده الوكيل (سلسلة من إطارات الفيديو) وتحويله إلى صيغة يستطيع النموذج اللغوي معالجتها.

خطوات المعالجة:

أخذ العينات: لا يستخدم كل إطار على حدة. بل يأخذ عينات آخر 40 إطارًا من السجل الحديث، هذا يُنشئ مقطع فيديو V سهل الاستخدام، ومن ثم يتم استخدام CLIPViT-L من أجل استخراج خريطة الميزات

يُحوّل هذا المرمز كل إطار إلى شبكة من الأرقام ، حيث تحتوي كل نقطة في الشبكة على D رقم يصف ذلك الجزء من الصورة، ثم يتم إنشاء نوعين من التضمينات:

* تضمين مُركّز مكانيًا Vs: يأخذ خريطة الميزات لكل إطار ويحسب متوسط ​​الميزات عبر الزمن لكل موقع مكاني. هذا يُعطي تمثيلًا يُبرز مكان وجود الأشياء، بمتوسط ​​40 إطارًا.
* ترميز مُركّز زمنيًا Vt: يأخذ خريطة الميزات ويحسب متوسط ​​المعالم عبر الفضاء لكل إطار. هذا يُعطي تمثيلًا يُبرز كيفية تغير الأشياء بمرور الوقت، بمتوسط ​​المواقع المكانية لكل إطار من إطارات.

يتم دمج التمثيلين للحصول على فيديو مدمج V=(Vs,Vt) ليصبح قادراً على التقاط الديناميكيات الزمنية والتفاصيل المكانية ويتم استخدام طبقة إسقاط بسيطة لتحويل بعد الفيديو إلى البعد الذي يمكن للنموذج اللغوي استقباله

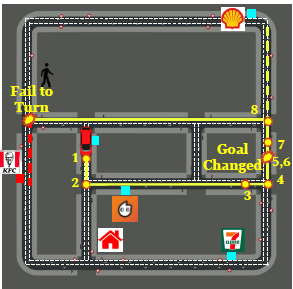
## 2.3.3- العمود الفقري لنظام إدارة التعلم (LLM)

الغرض: هذه هي وحدة صنع القرار المركزية. تستقبل معلومات الفيديو المُعالجة، وتعليمات الحوار (مثل أهداف المستخدم)، وتُقرر الإجراءات البسيطة التي يجب اتخاذها (مثل "الانعطاف يسارًا"، "التسارع").

تم استخدام Vicuna-7B وهو نموذج لغة مفتوح المصدر ومُحدد.

## 3.3.3- وحدة تخطيط المسار (ملاح GPS)

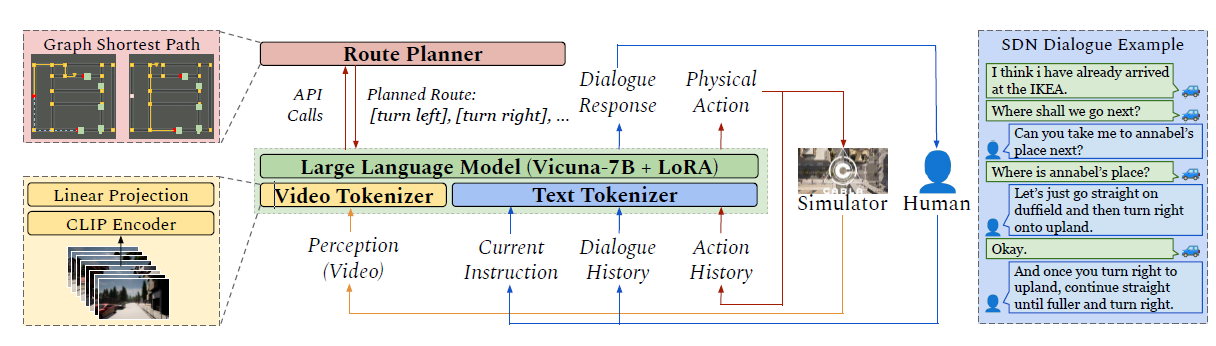
الغرض: تزويد العميل بخطة تفصيلية (مسار) للوصول إلى الوجهة، باستخدام معرفة الخريطة، يُعد هذا أمرًا بالغ الأهمية لتحقيق الأهداف بعيدة المدى.

حيث تأخذ هذه الوحدة الوجهة كدخل والموقع الحالي والخريطة وتقوم بحساب أقصر مسار ثم يُترجم هذا المسار إلى قائمة اتجاهات الانعطاف بلغة طبيعية (مثل: "انعطف يسارًا عند التقاطع التالي"، أو "استمر مستقيمًا لمسافة مبنيين ثم انعطف يمينًا").

الشكل (3-3): وحدة تخطيط المسار

### كيفية عمل الوحدات الثلاثة

يراقب مُرمّز الفيديو البيئة ويُحوّل الفيديو إلى خصائص رقمية، والتي بدورها تُنقل إلى النموذج اللغوي ليقوم بمعالجتها بالإضافة إلى التعليمات الحوارية وفي حال كانت الوجهة بعيدة فيقوم النموذج باستخدام GPS واستدعاء وحدة تخطيط المسار والذي يقوم بتحديد أقصر طريق ويعيد قائمة بالاتجاهات المبسطة مثل مستقيم، يسار، مستقيم، يمين... ويقوم النموذج باستخدام هذه التوجيهات بالإضافة إلى المدخلات المرئية اللحظية من مرمز الفيديو لزيادة دقة القيادة مثل أسرع قليلاً أو أدر عجلة القيادة يساراً الآن

يستخدم نظام إدارة المركبات الآن هذه الاتجاهات، بالإضافة إلى المُدخلات المرئية اللحظية من مُرمز الفيديو، لاتخاذ قرارات قيادة دقيقة وبسيطة (مثل "أدر عجلة القيادة يسارًا الآن"، "سرّع قليلاً")، مما يساعد في التنقل الصحيح حتى لو كانت الوجهة بعيدة.

الشكل (3-4): تكامل الوحدات الثلاثة

### 4.3.3- بيانات التدريب

تم استخدام بيانات NuScences بالإضافة إلى صور من بيئة محاكاة CARLA وإحداثيات دقيقة من هذا البيئة بالإضافة إلى توصيف نصي لمشهد القيادة بالإضافة إلى أسئلة وأجوبة حول مشهد القيادة، وهذه الأسئلة إما للاستدلال المكاني أو وصفية أو لاتخاذ القرار



### 5.3.3

### - التدريب

تم التدريب على مرحلتين

المرحلة الأولى: تم إنشاء الأوزان لطبقة الإسقاط بين النموذج اللغوي و مرمز الفيديو وذلك من خلال 50000 فيديو- نص للقيادة مع توصيف الأفعال المتخذة مثل التوقف والانعطاف... .

المرحلة الثانية:

* ضبط التعليمات الاجتماعية: تم ضبط النموذج اللغوي باستخدام LoRA وذلك من خلال 13000 فيديو- حوار والتي تتضمن:
  + فيديو لسيناريو قيادة وحوار مرتبط به.
  + حوارات بين الإنسان والسيارة: حوارات بين الإنسان والسيارة/الوكيل ، على سبيل المثال: "انعطف يسارًا عند التقاطع التالي"، أو "ما هو حد السرعة هنا؟"
  + أهداف طويلة المدى للمخططين: تعليمات تُحدد هدفًا عامًا، مثل "القيادة إلى السوبر ماركت". يحتاج النموذج إلى تعلم كيفية تقسيم هذا الهدف إلى أفعال أصغر.

مما يجعل النموذج اللغوي يفهم ما يريده الإنسان من المحادثة والهدف العام، ثم تحديد الإجراءات التدريجية (الإجراءات الأولية) اللازمة لتحقيقه.

* ضبط التعليمات المجسد :تم إنشاء بيانات تدريب جديدة من خلال محاكي لإدراك ما تراه السيارة عن محيطها المباشر حيث تم تسجيل البيانات مثل الطقس والأجسام القريبة وبعدها والبعد عن نهاية الطريق الحالي وبعد إشارات المرور وحالتها مثل حمراء أو خضراء و وضعوها في قوالب نصية.

أي في المرحلة الأولى تم تعلم كيفية تحويل الفيديو إلى نص وفي الجزء الأول من المرحلة الثانية يتعلم النموذج اللغوي كيفية فهم النوايا البشرية من خلال المحادثة وفي الجزء الثاني يجعل النموذج اللغوي يرسخ فهمه في الواقع الفعلي للقيادة من منظور السيارة.

### 6.3.3 - معايير التقييم

تم تقييم النموذج من النواحي التالية:

* دقة الأسئلة والأجوبة والذي تقيس نسبة الإجابات الصحيحة ومدى فهم النموذج للمشهد
* دقة التوصيف النصي للمشهد وتم استخدام CIDEr و SPICE
* دقة تحديد الكائنات في الصور
* دقة النموذج في استنتاج الإجراءات الممكنة

### 7.3.3- النتائج

| النموذج | QA Accuracy | CIDEr | SPICE | Grounding Accuracy |
| --- | --- | --- | --- | --- |
| LLaVA | 52.4% | 59.3 | 23.1 | 48.2% |
| RoadDreamer | 66.9% | 83.5 | 29.4 | 70.6% |

الجدول (3-2): نتائج معيار التقييم

### التحديات والمشاكل

على الرغم من الأداء الجيد لكل من البحثين السابقين لكن لم يكن قادراً على حل المشاكل التالية:

* وقت استدلال غير مقبول: يستغرق الذكاء الاصطناعي وقتًا طويلاً لمعالجة المعلومات واتخاذ القرارات/الاستجابة. وهذا يضر بالقيادة الآنية.
* دقة ضعيفة تجعل هذه الأبحاث غير قابلة للتطبيق فعلياً
* بيانات تدريب غير متوازنة: قد لا تغطي البيانات المستخدمة لتعليم الذكاء الاصطناعي جميع المواقف بالتساوي، مما يؤدي إلى تحيزات أو ضعف الأداء في السيناريوهات غير الممثلة تمثيلاً كافيًا.
* فهم بصري محدود: لا "يرى" الذكاء الاصطناعي العالم المرئي أو يفسره بشكل مثالي؛ وقد يغفل أو يسيء فهم الإشارات البصرية المهمة.
* تحديات التفاعلات متعددة الأدوار: يواجه صعوبة في المحادثات الطويلة والمتبادلة، مما قد يؤدي إلى فقدان السياق أو الاتساق.

## 4.3- نموذج تتبع سيارة ذاتية القيادة شبيه بالإنسان باستخدام التعلم المعزز العميق

يهدف هذا البحث إلى تطوير إطار قيادة ذكي لتخطيط القيادة الذاتية، وذلك من خلال التعلم من بيانات القيادة الطبيعية.

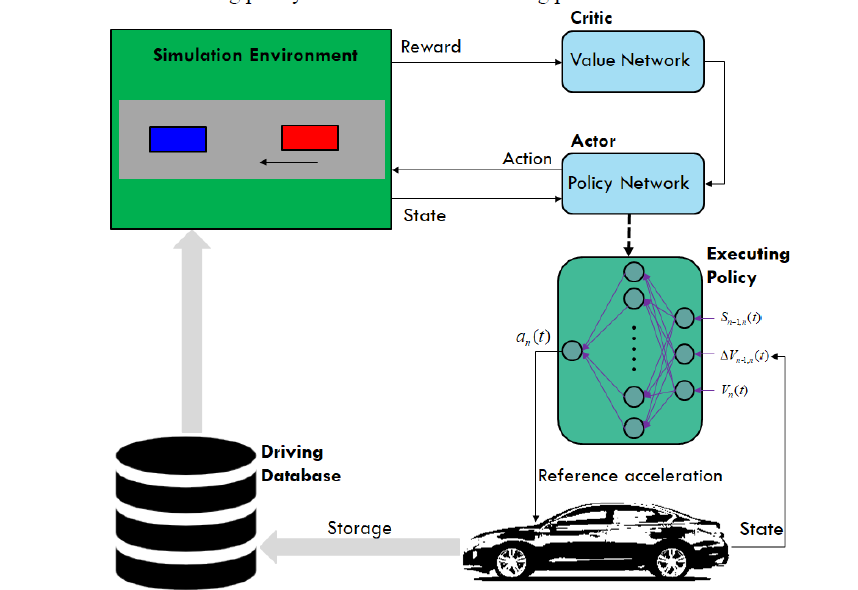
يعتمد الإطار على التعلم المعزز العميق لتعليم النموذج كيفية اتخاذ القرار بناءً على السياق الزمني لعناصر المشهد، مثل السرعة الحالية، الفجوة بين المركبات، والسرعة النسبية مع المركبة الأمامية.

### 1.4.3- بنية النموذج المقترح

يتكون الإطار من وكيل تعلم معزز يستخدم خوارزمية DDPG، وهي خوارزمية تنتمي إلى فئة Actor-Critic، وتهدف إلى تعلم السياسة المثلى لإنتاج أفعال مستمرة لتحقيق أكبر مكافأة تراكمية ممكنة وتتكون من:

* Actor Network:تتنبأ بالتسارع المطلوب للمركبة
* Critic Network: تقيم جودة الفعل بناءً على الحالة والمكافأة المتوقعة
* Replay Buffer: ذاكرة تخزن النتائج السابقة (state, action, reward, next\_state)
* Target Network: وهي نسخ مكررة من actor و critic يتم تحديثها تدريجياً لتحقيق استقرار التدريب

### 2.4.3- آلية عمل الخوارزمية

* مرحلة التهيئة
  + تهيئة شبكتي **Actor** و **Critic** الرئيسيتين.
  + تهيئة شبكتي **Target Actor** و**Target Critic** بنسخ من الشبكات الأصلية.
  + إنشاء **Replay Buffer**.
* التفاعل مع البيئة من أجل كل حلقة تدريبية
  + الحصول على الحالة الحالية
  + اختيار الفعل المطلوب
  + تنفيذ الفعل والحصول على المكافأة
  + الانتقال إلى الحالة التالية
  + تخزين التجربة في replay buffer
* التدريب باستخدام عينات من replay buffer
  + اختيار mini-batch من buffer
  + حساب critic
  + تحديث شبكة critic بناءً على الخطأ
  + تحديث شبكة actor باستخدام التدرج العكسي

الشكل(3-5): توضيح لآلية عمل الخوارزمية

تم اقتراح ثلاث نسخ منه

* **DDPGv**
  + يعتمد على الفرق بين السرعة المحاكاة والسرعة الملاحظة
  + Actor: تتنبأ بالتسارع المطلوب للمركبة التابعة.
  + Critic: تقييم جودة الفعل بناءً على الحالة والمكافأة المتوقعة.
* **DDPGvRT**
  + نسخة محسنة من DDPGv تأخذ زمن الاستجابة البشري في الحسبان
  + تستخدم تسلسل زمني للحالات السابقة والحالية كمدخلات مما يمكن النموذج من فهم السياق الديناميكي
* **DDPGs**
  + تعتمد على الفرق في المسافة بدلًا من السرعة كمصدر للمكافأة.

### 3.4.3- البيانات المستخدمة

تم جمع البيانات باستخدام خمس سيارات قادت على مدى 161000 كم، وهذه البيانات تتألف من:

* موقع GPS بدقة عالية
* السرعة والتسارع على المحورين x,y
* بيانات الزاوية التوجيهية للمقود، ودواسة الوقود
* رادار لقياس المسافات والسرعات للمركبات المحيطة
* 4 كاميرات على الجوانب الأربعة
* تصنيف البيانات إلى بيانات من سائقين عدائين وبيانات من سائقين محافظين

### 4.4.3- التدريب

تم التدريب على مرحلتين

* **التدريب المسبق (Pre-training):**
  + **البيانات**: فترات قيادة تتابعية من 20 سائقاً تم اختيارها عشوائيًا من قاعدة بيانات.
  + **الهدف**: تعلّم العلاقات بين المسافة، السرعة، والفجوة الزمنية، وتوليد سياسات تسارع واقعية.
  + **النتيجة**: نموذج مهيأ بدقة عالية في تتبع المسارات الحركية المُلاحظة.
* **الضبط الدقيق (Fine-tuning) لكل سائق:**
  + **البيانات**: إعادة تدريب النموذج نفسه باستخدام بيانات سائق جديد.
  + **الهدف**: اختبار قدرة النموذج على **التكيف** مع أنماط قيادة فردية.
  + **النتيجة**: انخفاض ملحوظ في الخطأ بعد عدة دورات تدريبية، مما يثبت قابلية التكيف للنموذج.

### 5.4.3- معايير التقييم

تم استخدام RMSPE للمقارنة بين البيانات الحقيقية والمحاكاة، ويتم حسابه لكل من المسافة والسرعة، وتكون المكافأة بهدف تقليص الفروقات بين القيم المحاكاة والواقعية وكانت عبارة عن:

* DDPGS:تعتمد على الفرق في المسافة
* DDPGv:تعتمد على الفرق في السرعة
* DDPGvRT:تعتمد على السرعة مع تأخير زمني قدره ثانية واحدة

### 6.4.3- النتائج

النموذج DDPGvRT حقق أفضل أداء بخطأ 18% في المسافة وخطأ 5% في السرعة.

## المشاكل التي واجهها هذا البحث

1. الاعتماد على بيانات القيادة البشرية (الواقعية) قد يؤدي إلى تعلم سلوكيات خاطئة

### 2. السلامة (Safety) لم يتم التعامل معها بشكل مباشر

* النموذج يركّز على محاكاة السلوك البشري وليس على تجنّب الحوادث.

3. ضعف التعميم عبر أنماط قيادة مختلفة

* بالرغم من أن النموذج يتفوق في تعميمه على بيانات جديدة، إلا أن الاختلاف الكبير بين السائقين أدى إلى زيادة الخطأ في بعض الحالات.

## 5.3- القيادة باستخدام النماذج اللغوية الكبيرة: دمج متجه مستوى الكائن في القيادة الذاتية القابلة للتفسير

يهدف هذا الإطار إلى تمكين النماذج اللغوية الكبيرة (LLMs) من اتخاذ قرارات قيادة قابلة للتفسير، وذلك من خلال دمج بيانات المشهد الرقمية (على شكل متجهات) مباشرةً في النموذج، ويتم تحقيق ذلك عبر طرح أسئلة على النموذج حول الموقف الحالي، ليقوم بدوره بتوليد الإجراءات الواجب اتخاذها (مثل التسارع والفرملة) مع شرح سبب القرار، مما يعزز الثقة بين الراكب والسيارة ذاتية القيادة

### 1.5.3 مرمّز الأشعة (Vector Encoder)

في هذه المرحلة الأولية، يستقبل النموذج أربعة أنواع من البيانات الشعاعية التي تصف مشهد القيادة بشكل شامل: بيانات المسار (route)، بيانات المركبات المجاورة (vehicles)، بيانات المشاة (pedestrians)، وبيانات الحالة الذاتية للمركبة الذاتية القيادة (ego state).

### 2.5.3 مشكّل الأشعة (Vector-Former)

بعد الحصول على التمثيل الكامن، يتم إرساله إلى وحدة تُدعى (Vector-Former)، وهي عبارة عن بنية قائمة على معمارية المحولات (Transformer). وظيفة هذه الوحدة هي فهم التمثيل الكامن المدخل إليها وتحويله إلى صيغة نهائية متوافقة مع النموذج اللغوي.

### 3.5.3 النموذج اللغوي مع التكيف المنخفض (LLM with LoRA)

في هذه المرحلة الحاسمة، يتم دمج التضمين الشعاعي الناتج عن الـ (Vector-Former) مع تضمين الأمر النصي (prompt embedding) الذي أدخله المستخدم. يتم إدخال هذين التضمينين إلى نموذج لغوي كبير مدرب مسبقاً، والذي تم تكييفه لأداء مهمة القيادة باستخدام تقنية (LoRA - Low-Rank Adaptation). تسمح هذه التقنية بتعديل سلوك النموذج اللغوي دون الحاجة إلى إعادة تدريب كامل أوزانه، مما يوفر موارد حاسوبية هائلة.

يقوم النموذج اللغوي بتحليل السياق الكامل، ثم يقوم بتوليد استجابة نصية متكاملة. يتم بعد ذلك تحليل هذه الاستجابة لاستخلاص:

* أوامر التحكم بالمركبة: قيم رقمية دقيقة مثل نسبة التسارع، وقوة المكابح، و زاوية عجلة القيادة.
* شرح منطقي للقرار: نص باللغة الطبيعية يوضح سبب اتخاذ هذه الأفعال (مثال: "سأقوم بالفرملة لأن هناك مشاة يعبرون الطريق أمامي").
* توليد عينة من الأسئلة المرتبط بمشهد القيادة والإجابة عليها

وبهذا، لا يكتفي النموذج باتخاذ القرار فحسب، بل يقدم أيضاً تفسيراً منطقياً له، مما يحقق الهدف الأساسي للبحث وهو القيادة الذاتية بالإضافة إلى التمتع بموثوقية عالية.

### 4.5.3- بيانات التدريب

بيانات التدريب عبارة عن لائحة من:

* لائحة من أشعة المركبات المحيطة
* شعاع يمثل السيارة الحالية
* لائحة من المشاة
* الطريق الواجب سلوكه
* التوصيف النصي الذي يعبر عن حالة مشهد القيادة بناءً على البيانات الشعاعية، والذي يحتوي على الفعل الواجب اتخاذه
* أسئلة وأجوبة مرتبطة بمشهد القيادة

### 5.5.3- تدريب النموذج

تم تدريب النموذج على مرحلتين

المرحلة الأولى:

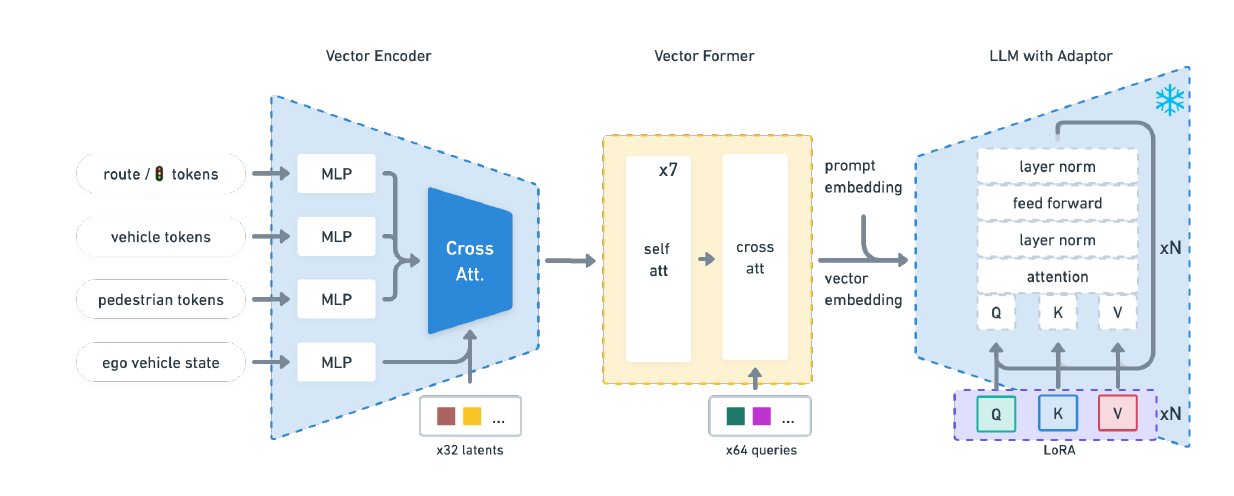
التدريب الأولي لفهم البيانات الشعاعية (Vector Representation Pre-training):

الهدف الأساسي من هذه المرحلة هو تعليم النموذج كيفية فهم المدخلات الرقمية (المتجهات) التي تصف حالة القيادة (مثل موقع السيارات والمشاة وسرعتهم)،يتم ذلك عن طريق ربط هذه البيانات الرقمية بأوصاف نصية تصف المشهد.

في هذه المرحلة، يتم "تجميد" النموذج اللغوي الكبير، ويتم تدريب الأجزاء المسؤولة عن تحويل الأشعة فقط، مما يجبر النموذج على تعلم كيفية ترجمة الأرقام إلى تمثيل يمكن للنموذج اللغوي فهمه.

الضبط الدقيق للإجابة عن أسئلة القيادة (Driving QA Finetuning):

بعد أن أصبح النموذج قادرًا على فهم المدخلات الرقمية من المرحلة الأولى، يتم في هذه المرحلة تدريبه على أداء المهام النهائية. يُستخدم هنا مجموعة بيانات غنية تحتوي على أسئلة حول سيناريوهات القيادة (مثل "ماذا ستفعل الآن ولماذا؟") وإجاباتها النموذجية، بالإضافة إلى أوامر التحكم الدقيقة للمركبة (التسارع، الفرملة، التوجيه). في هذه المرحلة، يتم تدريب النموذج بأكمله (بما في ذلك تحديثات طفيفة للنموذج اللغوي باستخدام تقنية LoRA) لكي يتمكن من التفكير في الموقف، والإجابة على الأسئلة بشكل منطقي، وتوليد الأوامر الصحيحة لقيادة السيارة.

الشكل (3-6): تدريب النموذج

### 6.5.3- معايير التقييم

تم تقييم دقة توصيف المشهد بالإضافة إلى الفعل المتخذ من خلال

* حساب الخطأ في توقع عدد السيارات وعدد المشاة
* حساب دقة وجود إشارة المرور والخطأ في بعدها
* حساب الخطأ في الفعل المتخذ

بالإضافة إلى تقييم دقة الأسئلة والأجوبة التي تم توليدها، وتم ذلك بالاعتماد على التقييم البشري وعلى ChatGPT-3.5 Turpo لتقييم هذه الأسئلة بعد استخدام prompt مناسبة تجعله يقيم بشكل صحيح

### 7.5.3- النتائج

نتائج تقييم توصيف المشهد مع الفعل المتخذ

| E(lat) | E(lon) | D(TL) | ACC(TL) | E(pedestrian) | E(car) |
| --- | --- | --- | --- | --- | --- |
| 0.014 | 0.066 | 6.624 | 0.718 | 0.313 | 0.066 |

الجدول (3-3): نتائج التقييم

نتائج تقييم الأسئلة والأجوبة

| Human Grading | GPT Grading |
| --- | --- |
| 7.71 | 8.39 |

### 8.5.3- المشاكل

* التعامل المحدود مع مشاهد القيادة وذلك لأن التجارب محدودة بالجغرافيا التي تم الحصول على البيانات منها والكثافة المرورية بالإضافة إلى عدم وجود مشاهد نادرة مثل وجود حوادث
* الحاجة إلى موارد حسابية مرتفعة

## 6.3- الخاتمة

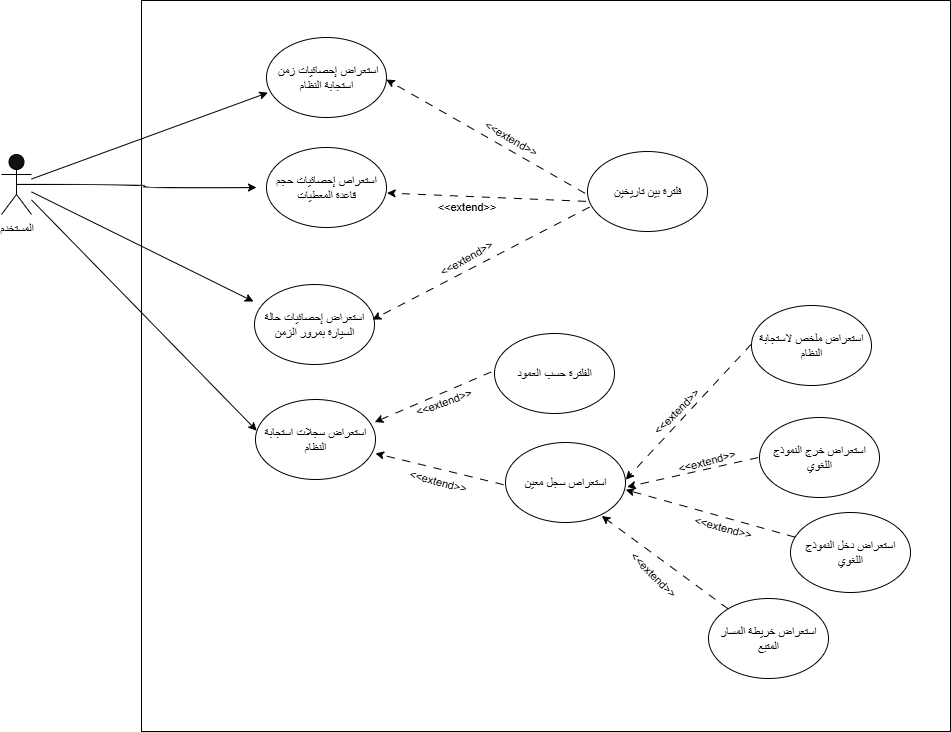
| النموذج | المزايا | العيوب |
| --- | --- | --- |
| OmniDrive | أداء قوي، بنية VLM، بيانات استدلال متقدمة | معقد جدًا، دمج بصري ثقيل |
| تعزيز وكلاء القيادة الذاتية المعتمدة على النماذج اللغوية الكبيرة من خلال تجارب مجسدة واجتماعية DriVLMe | دمج اجتماعي، تفاعل واقعي، مهام طويلة | بطء الاستدلال، تعقيد بالحوار |
| القيادة باستخدام النماذج اللغوية الكبيرة: دمج متجه مستوى الكائن في القيادة الذاتية القابلة للتفسير | تفسير ممتاز، موارد قابلة للتحكم، قابل للتجريب بسهولة | يغطي مشاهد محدودة جغرافيًا فقط |
| نموذج تتبع سيارة ذاتية القيادة شبيه بالإنسان باستخدام التعلم المعزز العميق | أداء فعلي ممتاز، ومختبر على بيانات واقعية | ضعف التعميم وتعلم سوكيات خاطئة |

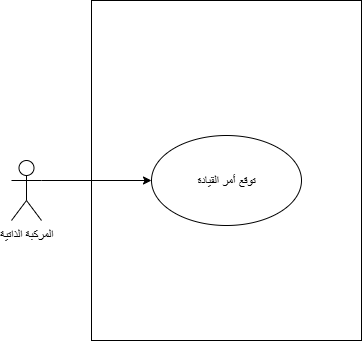
# الفصل الرابع

# الدراسة التحليلية

*يوضح هذا الفصل عملية تحليل النظام ودراسة متطلباته*

## 1.4- مخطط حالات الاستخدام

الشكل(4-1): مخطط حالات الاستخدام



## 2.4- السرد النصي لحالات الاستخدام

### 1.2.4-استعراض إحصائيات زمن استجابة النظام

| اسم الحالة: استعراض إحصائيات زمن استجابة النظام | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض مخططات تظهر توزع أزمنة الاستجابة في مخططات مختلفة ضمن مجالات زمنية مختلفة | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | تحديد المجال الزمني المراد لعرض البيانات | | | |
| الشروط اللاحقة Postcondition | عرض إحصائيات استجابة النظام ضمن المجال الزمني المحدد | | | |

**سير الأحداث**

**السيناريو الأساسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يحدد المجال الزمني المراد عرض الإحصائيات خلاله |
| 2- يعرض النظام المخططات الإحصائية ضمن هذا المجال |  |

**المسارات البديلة**

لا يحدد المجال الزمني فيقوم النظام بناء الإحصائيات بناءً على بيانات اليوم الحالي

**المسارات الخطأ**

E1: في المرحلة رقم 1 في حال اختار المستخدم تاريخ نهاية المجال الزمني قبل تاريخ بداية المجال الزمني فيعيد النظام رسالة خطأ يطلب فيها تحديد تاريخ نهاية المجال بعد تاريخ نهاية المجال

### 2.2.4- استعراض إحصائيات حالة السيارة بمرور الزمن

| اسم الحالة: استعراض إحصائيات حالة السيارة بمرور الزمن | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض مخططات تظهر فيها إحصائيات قيم التسارعات والمكابح والتفاف التي قامت بها المركبة، بالإضافة إلى استعراض إحصائيات لمشاهد القيادة كعدد المشاة والمركبات التي تم تحديدها خلال كل مشهد | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | تحديد المجال الزمني المراد لعرض البيانات | | | |
| الشروط اللاحقة Postcondition | عرض إحصائيات قيم التسارعات والمكابح والتفاف وعدد الأشخاص والمركبات في كل مشهد قيادة | | | |

**سير الأحداث**

**السيناريو الأساسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يحدد المجال الزمني المراد عرض الإحصائيات خلاله |
| 2- يعرض النظام المخططات الإحصائية ضمن هذا المجال الزمني  للقيم التي اتخذتها المركبة خلال القيادة من تسارع وكبح والتفاف بالإضافة إلى بيانات مشاهد القيادة كعدد السيارات والمشاة |  |

**المسارات البديلة**

لا يحدد المجال الزمني فيقوم النظام بناء الإحصائيات بناءً على بيانات اليوم الحالي

**المسارات الخطأ**

E1: في المرحلة رقم 1 في حال اختار المستخدم تاريخ نهاية المجال الزمني قبل تاريخ بداية المجال الزمني فيعيد النظام رسالة خطأ يطلب فيها تحديد تاريخ نهاية المجال بعد تاريخ نهاية المجال

### 3.2.4- استعراض إحصائيات حجم قاعدة البيانات

| اسم الحالة: استعراض إحصائيات حجم قاعدة البيانات | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض حجم البيانات التي تم تخزينها خلال كل يوم | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | تحديد المجال الزمني المراد | | | |
| الشروط اللاحقة Postcondition | عرض إحصائيات حجم البيانات المخزن خلال هذا المجال الزمني بالإضافة إلى إحصائيات للقيم الأكثر تكراراً لقيم التسارعات والمكابح والالتفاف | | | |

**سير الأحداث**

**السيناريو الرئيسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يحدد المجال الزمني المراد عرض الإحصائيات خلاله |
| 2- يعرض النظام مخطط إحصائيات عدد البيانات المخزنة خلال المدة الزمنية المختارة بالإضافة إلى عرض أوامر القيادة الأكثر تكراراً |  |

**المسارات البديلة**

لا يحدد المجال الزمني فيقوم النظام بناء الإحصائيات بناءً على بيانات اليوم الحالي

**المسارات الخطأ**

E1: في المرحلة رقم 1 في حال اختار المستخدم تاريخ نهاية المجال الزمني قبل تاريخ بداية المجال الزمني فيعيد النظام رسالة خطأ يطلب فيها تحديد تاريخ نهاية المجال بعد تاريخ نهاية المجال

### 4.2.4- استعراض سجلات استجابة النظام

| اسم الحالة: استعراض سجلات استجابة النظام | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض سجلات مشاهد القيادة | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | لا يوجد | | | |
| الشروط اللاحقة Postcondition | عرض سجلات مشاهد القيادة | | | |

**سير الأحداث**

**السيناريو الأساسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يطلب المستخدم عرض سجلات مشاهد القيادة |
| 2- يعرض النظام هذه السجلات بالترتيب حسب زمن السجل من الأحدث للأقدم |  |

**المسارات البديلة**

يختار المستخدم ترتيب السجلات بناءً على عمود آخر تصاعدياً أو تنازلياً فيعيد النظام ترتيب السجلات بناءً على العمود المختار وبالترتيب المطلوب تصاعدياً أو تنازلياً

**المسارات الخطأ**

لا يوجد

### 5.2.4- استعراض سجل معين

| اسم الحالة: استعراض سجل معين | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض سجل معين من بين سجلات القيادة | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | استعراض سجلات النظام | | | |
| الشروط اللاحقة Postcondition | استعراض واجهة لاختيار الجزء المراد عرضه من السجل | | | |

**سير الأحداث**

**السيناريو الرئيسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يحدد السجل المراد عرضه |
| 2- يعرض النظام واجهة خاصة بالسجل والتي تمثل المشهد |  |

**المسارات البديلة**

لا يوجد

**المسارات الخطأ**

لايوجد

### 6.2.4- استعراض ملخص سجل

| اسم الحالة: استعراض ملخص سجل | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض ملخص سجل | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | استعراض سجل معين | | | |
| الشروط اللاحقة Postcondition | استعراض ملخص سجل يحتوي على ما تم مشاهدته ضمن مشهد القيادة من مركبات ومشاة وبعد كل منهم بالإضافة إلى الطريق الواجب اتخاذه | | | |

**سير الأحداث**

**السيناريو الأساسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يختار المستخدم عرض الملخص ضمن سجل معين |
| 2- يعرض النظام المشهد الذي تنبأ به النظام من عدد سيارات وعدد مشاة وبعد كل منها بالإضافة إلى الأمر المتوقع |  |

**المسارات البديلة**

لا يوجد

**المسارات الخطأ**

لايوجد

### 7.2.4- استعراض خرج النموذج اللغوي

| اسم الحالة: استعراض سجل معين | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض خرج النموذج بدون أي معالجة أو تقسيم إلى فعل أو توصيف | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | استعراض سجل معين | | | |
| الشروط اللاحقة Postcondition | استعراض نص غير منسق ومتسلسل بشكل كامل | | | |

**سير الأحداث**

**السيناريو الرئيسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يختار المستخدم عرض خرج النموذج اللغوي |
| 2- يعرض النظام التنبؤ الذي قام به النموذج اللغوي |  |

**المسارات البديلة**

لا يوجد

**المسارات الخطأ**

لايوجد

### 8.2.4- استعراض خريطة المسار المتبع

| اسم الحالة: استعراض خريطة المسار المتبع | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم المستخدم باستعراض خريطة المسار المتبع | | | |
| الفاعلين Actors | المستخدم | | | |
| الشروط السابقة Precondition | استعراض سجل معين | | | |
| الشروط اللاحقة Postcondition | استعراض خريطة تحتوي على المسار المسلوك الذي تنبأ به النظام بالإضافة إلى مواقع المشاة والمركبات على الخريطة | | | |

**سير الأحداث**

**السيناريو الرئيسي الناجح**

| النظام | المستخدم |
| --- | --- |
|  | 1- يختار المستخدم عرض الخريطة |
| 2- يعرض النظام خريطة تحتوي على المسار المسلوك الذي اقترحه النظام |  |

**المسارات البديلة**

لا يوجد

**المسارات الخطأ**

لايوجد

### 9.2.4- توقع أمر القيادة

| اسم الحالة: توقع أمر القيادة | | | | |
| --- | --- | --- | --- | --- |
| الوصف Description | يقوم النظام بتوقع فعل الحركة وتقوم المركبة | | | |
| الفاعلين Actors | المركبة | | | |
| الشروط السابقة Precondition | أن يرسل النظام بيانات المشهد الرقمي إلى النموذج | | | |
| الشروط اللاحقة Postcondition | تنفيذ أمر الحركة الذي قام النظام بتوقعه وإرسال بيانات المشهد الرقمي مرة أخرى | | | |

**سير الأحداث**

**السيناريو الرئيسي الناجح**

| النظام | المركبة |
| --- | --- |
|  | 1- ترسل بيانات المشهد الرقمية إلى النظام |
| 2- يقوم النظام باستقبال البيانات |  |
| 3- يتنبأ النظام بالمشهد والفعل الواجب اتخاذه |  |
| 4- استخراج الفعل الواجب اتخاذه من النص المتُنبَأ به وإرساله إلى المركبة |  |
|  | 5- تنفيذ الفعل المتوقع |

**المسارات البديلة**

لا يوجد

**المسارات الخطأ**

E1:في حال حدوث خلل في الشبكة أو خطأ في النظام أثناء عملية التوقع ، فتبقى المركبة دون تنفيذ أي فعل وترسل بيانات المشهد مرة أخرى

# الفصل الخامس

# المنهجية المقترحة

*يعرض هذا الفصل المنهجية المقترحة ضمن مشروع القيادة الذاتية*

## 1.5- مقدمة

بعد النظر إلى الأبحاث في مجال الأنظمة ذاتية القيادة، تم اعتماد آخر منهجية تم ذكرها وهي دمج نمط جديد من البيانات مع النماذج اللغوية الكبيرة لتطوير نظام قيادة قابل للتفسير،ولكن لن نقوم باستخدام الأسئلة والأجوبة وذلك بسبب أن زمن الاستجابة هو X ثانية بحيث أن المستخدم لن يقدر على كتابة السؤال وإرساله إلا ويكون الفعل التالي قد تم اتخاذه، وبدلاً من ذلك للحفاظ على ثقة المستخدم بالنموذج سيتم إرسال الفعل مع التوصيف النصي للمشهد مما يوفر على المستخدم كتابة السؤال، وبالإضافة إلى ذلك نكون قد حافظنا على ثقة المستخدم بالنظام.

## 2.5- النموذج المقترح

بداية نعرف دخل النموذج، حيث يتكون الدخل من الأشعة التي تمثل مشهد القيادة، وخرج النموذج والذي سيكون عبارة عن توصيف نصي بالإضافة إلى الفعل المقترح ويكون الخرج على الشكل التالي

I'm observing 0 cars and 3 pedestrians

A pedestrian; Angle in degrees: -35.50; Distance: 4.63m; Direction of travel: same direction as me Crossing: False

A pedestrian; Angle in degrees: 34.86; Distance: 10.49m; Direction of travel: same direction as me; Crossing: False

A pedestrian; Angle in degrees: -10.83; Distance: 14.34m; Direction of travel: same direction as me; Crossing: False

The distance to the closest intersection is 14.32m

It is a roundabout.

There is no traffic lights.

My car 0.00m from the lane center and 0.00 degrees left off center.

My current speed is 0.00 mph\nSteering wheel is 0.00% right.

The next turn is 84 degrees right in 51.26m.

- Accelerator pedal 62%\n- Brake pedal 0%\n- Going to steer 0% to the right.

ونقوم باستخلاص الفعل من النص و إرساله إلى بيئة المحاكاة لتقوم بتنفيذ المطلوب من خلال تابع تم إنشاءه

### 3.5- مكونات النموذج

سنقوم بتفصيل مكونات الأساسية للنموذج بالتفصيل

### 1.3.5- مرمّز الأشعة

دخل هذا المكون هو:

* وصف الطريق وهو مصفوفة بحجم 17\*30 حيث 30 تمثل نقاط المسار و كل نقطة تمثل ب 17 بعد
* وصف المركبات وهو مصفوفة بحجم 33\*30 حيث 30 تمثل عدد السيارات وكل سيارة تمثل ب 33 بعد
* وصف المشاة هو مصفوفة بحجم 9\*20 حيث 20 تمثل عدد المشاة وكل مار يمثل ب 9 أبعاد
* وصف المركبة الذاتية والتي تمثل ب 31 بعد

**1. طبقات الإسقاط MLP**

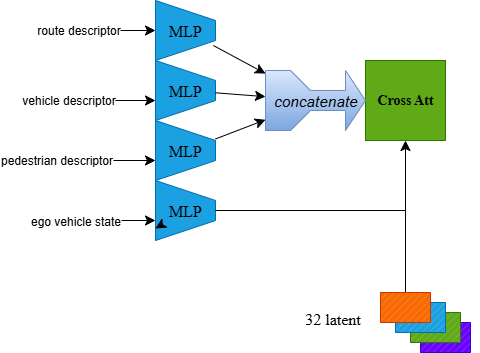
* **المكوّنات**
  1. إسقاط لكل نقطة مسار (route descriptor)
  2. إسقاط لكل مركبة محيطة (vehicle descriptor)
  3. إسقاط لكل مشاة (pedestrian descriptor)
  4. إسقاط لحالة السيارة الذاتية (ego state)
* **الهدف** تحويل كل شعاع خام (بُعده الأصلي يختلف بحسب نوعه) إلى تمثيل داخلي ثابت الطول (256) يساعد على توحيد مقياس المعلومات وتجهيزها للدخول إلى آليات الانتباه.

**2. دمج التمثيلات (Concatenate)**

* بعد مرور كل مدخل عبر الـ MLP الخاص به، نجمع كل المخرجات في سياق واحد.
* نحتفظ بقناع (mask) يُحدد أيُّ مدخلاتٍ فعليّة (مركبة/مشاة/نقطة طريق) وأيُّها مجرد حشو أي متجهات صفرية.
* يصبح لدينا مصفوفة كبيرة تمثّل المشهد كله بشكل متسلسل ويكون حجمها 80\*256.

**3. الانتباه المتقاطع الأولي (Cross-Attention)**

* **آلية العمل**
  1. تبدأ بـ 32 متجهاً كامناً جُمعت من حالة المركبة الذاتية.
  2. لكل استعلام، تبحث في كامل السياق عن الأجزاء الأكثر صلة، وتلخّصها في نتيجة واحدة.
* **النتيجة** مجموعة مصغرة (32 متجهاً) تحوي تلخيصاً مركزياً للمشهد بأكمله، مما يقلل حجم البيانات دون فقدان المعلومات الحرجة.

الشكل (5-1): مرمز الأشعة

### 2.3.5- مشكل الأشعة

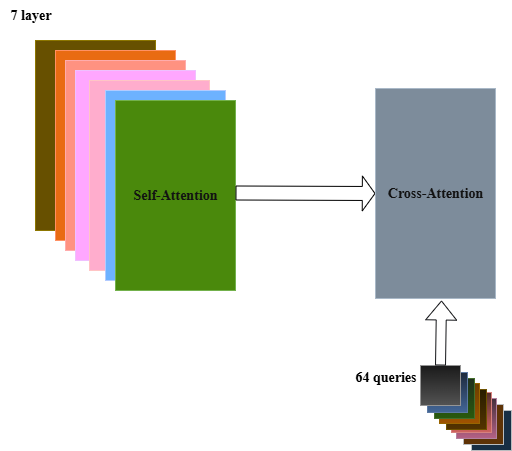
يتكون من الأجزاء التالية:

**1.سلسلة الانتباه الذاتي (Self-Attention Stack)**

* **آلية العمل**سبع طبقات متوالية، في كلٍ منها:
  1. كل متجه كامن يستخرج المعلومات من باقي المتجهات في المجموعة (بما فيها نفسه).
  2. يدمج ما رآه مع حالته الحالية عبر طبقة إسقاط داخلية.
* **النتيجة** تمثيلات كامنة متزنة، قادرة على احتواء العلاقات المتبادلة بين كل مقطع من المشهد (مركبة–مشاة–طريق) في حزمة واحدة.

**2. الانتباه المتقاطع النهائي (Output Cross-Attention)**

* **آلية العمل** يستخدم استعلامات جديدة (output queries) ليتواصل مع المتجهات الكامنة المحسّنة ويستدعي منها النقاط الأكثر إفادة لتفسير المشهد واتخاذ القرار.
* **النتيجة** متجه وحيد مُركّز، جاهز للإدراج في الـ LLM لتوليد النصّ والفعل المقترح مباشرة.



الشكل(5-2): مشكل الأشعة

### 3.3.5-النموذج اللغوي مع التكيف المنخفض الرتبة

هو نموذج لغوي محسن بآلية LoRA حيث يتم تجميد النموذج اللغوي وهو deepseek-1.3-base

* **آلية العمل**

1. توليد prompt النصي

### Instruction:

(تعليمات المستخدم)

### Input:

(جزء نصّي فارغ)

### Response:

1. يتم ترميز هذا prompt إلى سلسلة من input\_ids و attention\_mask و labels
2. يتم تضمين خرج مشكّل المتجه ودمجه مع تضمين prompt في مكان محدد وهذا المكان هو بالضبط بعد تضمين كلمة input بحيث نحصل على prompt كامل يمثل المشهد
3. النموذج مجمد بالكامل ويتم تطبيق LoRA مستهدفاً وحدات q\_proj و v\_proj في كل طبقة يكون dropout=0.05 و a=8 و rank=4

## الشكل (5-3): نموذج LORA

## 4.5- معايير التقييم المتبعة

سيتم الأخذ بعين الاعتبار المعايير التي تم اتباعها في البحث الأخير المذكور في الدراسة المرجعية والتي هي

1. TL accuracy
2. Vehicle count
3. Pedestrian count
4. Lon control
5. Lat control

# الفصل السادس

# تصميم النظام

*يعرض هذا الفصل القرارات التصميمية التي بني من خلالها النظام*

## 1.6- مقدمة

يهدف هذا التصميم إلى تحقيق تكامل مرن وفعال بين بيئة المحاكاة و النموذج اللغوي الكبير، وذلك من خلال نظام يسمح بجمع المشاهد الشعاعية من البيئة كل مدة محددة، ثم تفسير النتائج واسترجاع الأفعال اللازمة للتحكم بالمركبة الذاتية في الزمن الحقيقي بما يضمن سلامة التنقل ووضوح التفسير.

## 2.6- بيئة المحاكاة

تم اختيار بيئة محاكاة CARLA والتي هي عبارة عن منصة مفتوحة المصدر توفر لنا واجهات برمجة التطبيقات APIs، وتسمح للمستخدمين بتخصيص كل جانب من جوانب المحاكاة بدءاً من المشاهد والخرائط وصولاً إلى سلوكيات القيادة والمشاة،وتقدم مجموعة واسعة من الحساسات المدعومة مثل كاميرات RGB وكاميرات العمق، بالإضافة إلى التكامل مع مكتبات مثل Pytorch و TensorFlow.

## 3.6- النموذج اللغوي

تم استخدام النمط الهجين حيث تم دمج النموذج اللغوي مع ترميز الموجهات بالإضافة إلى fine tune للنموذج اللغوي، وتم اعتماد modular architecture في تقسيم ملفات النظام، مما يوفر مرونة في تغيير النموذج اللغوي المستخدم، بالإضافة إلى استثمار الموارد من خلال استخدام تقنيات 8bit.

## 4.6- الواجهات الخلفية Back-End

تم اتخاذ معمارية vertical slice architecture لبناء هذه المكون من النظام، إذ تقوم هذه المعمارية على تقسيم التطبيق إلى شرائح تحتوي كل شريحة على جميع طبقات الهندسة المعمارية،على

سبيل المثال واجهة المستخدم، الاتصال مع قاعدة المعطيات ومنطق العمل...( لتوفير حالة استخدام كاملة مستقلة عن باقي

حالات الاستخدام ضمن النظام.

تقدم هذه المعمارية مجموعة من الفوائد أهمها:

* تحترم مبدأ single responsibility حيث أن كل شريحة (slice) تحتوي على المنطق الكامل الخاص بميزة معينة (feature) أي حالة استخدام، مما يسمح باختبارها وصيانتها بسهولة بمعزل عن الميزات الأخرى التي يقدمها النظام.
* تحترم مبدأ High Cohesion بين مكونات الشريحة حيث أن كل مكونات الشريحة متماسكة فيما بينها وتخدم هدفاً واحداً.

من الجدير بالذكر أن هذه البنية تعاني من بعض المشاكل أهمها :

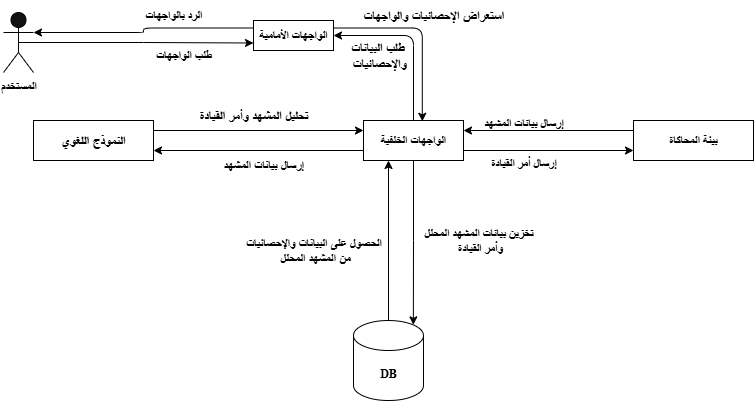
تكون مكونات الشريحة غير قابلة لإعادة الاستخدام ضمن شرائح أخرى، مما قد يؤدي إلى إعادة بناء بعض المكونات في شرائح

أخرى كان من الممكن إعادة استخدامها من الشريحة الأساسية .

## 5.6- الواجهات الأمامية

تم بناء واجهة المستخدم باستخدام بنية قائمة على المكونات مما يوفر العديد من الفوائد أهمها:

* النسيقية (Modularity) : تسمح بتقسيم موقع الويب إلى وحدات أصغر (مكونات) مستقلة. يغلف كل مكون وظيفة محددة، مما يجعل إدارته وفهمه أسهل .
* قابلية إعادة الاستخدام: بمجرد إنشاء أحد المكونات، يمكن إعادة استخدامه عبر أجزاء مختلفة من موقع الويب أو حتى في مشاريع مختلفة. يقلل هذا من التكرار ويسرع التطوير ويضمن الاتساق في واجهة المستخدم والسلوك .
* سهولة تصحيح الأخطاء: تعمل المكونات المعزولة على تبسيط عملية تصحيح الأخطاء، حيث يمكن تتبع المشكلات إلى مكونات معينة بدلاً من البحث ضمن قاعدة الرماز الكبيرة (huge codebase)



الشكل(6-1):مثال عن الواجهة الأمامية

# الفصل السابع

# الأدوات المستخدمة

*يعرض هذا الفصل الأدوات المستخدمة لتنجيز النظام*

## 1.7-Docker

هي منصة مفتوحة مصممة لأتمتة عملية نشر التطبيقات وتوسيع نطاقها (scaling) وإدارتها في حاويًت خفيفة الوزن (containers)

الحاوية (container) هي بيئة معزولة تجمع بين التطبيق وجميع تبعياته، مما يسمح بتشغيله بشكل متسق عبر

بيئات حوسبة مختلفة، مثل أنظمة تشغيل مختلفة أو مزودي الخدمات السحابية .

## 2.7-MongoDB

MongoDB هو برنامج قاعدة بيانات NoSQL مفتوح المصدر شائع الاستخدام. يندرج ضمن فئة قواعد البيانات الموجهة

للمستندات (document-oriented databases) ، مما يعني أنه يخزن البيانات بتنسيق يشبه JSON يعرف باسم BSON وهو (JSON الثنائي).

تم تصميم MongoDB ليكون مرناً وقابلًا للتوسع وسهل التعامل معه لكل من المطورين والمسؤولين (administrators)

## 3.7-GitHub

هي منصة تعمل ضمن الويب توفر أدوات للتحكم في الإصدارات والتعاون بين المطورين لتطوير البرامج، حيث أنها مبنية على

نظام التحكم في الإصدارات Git . تسمح هذه المنصة بتخزين الكود البرمجي ومشاركته بين مجموعة من المطورين كما تقدم

مجموعة من الأدوات للتعامل مع بيئات التطوير المستندة إلى السحابة (cloud) وإستضافة مواقع الويب الثابتة .

## 4.7-Hugging Face

هي منصة مفتوحة المصدر وشركة تقنية متخصصة في مجال الذكاء الاصطناعي ومعالجة اللغة الطبيعية. توفر المنصة مكتبة ضخمة من النماذج الجاهزة مثل نماذج GPT وBERT وغيرها، والتي يمكن استخدامها في تطوير تطبيقات الذكاء الاصطناعي بسهولة وسرعة. كما توفر Hugging Face بيئة متكاملة للمطورين، تتيح لهم استضافة ومشاركة النماذج ومجموعات البيانات والتطبيقات بشكل مفتوح. من أبرز فوائد Hugging Face تسهيل الوصول إلى أحدث تقنيات الذكاء الاصطناعي، وتمكين المطورين والباحثين من التعاون في تطوير النماذج، وتوفير وقت وجهد كبيرين في عملية التدريب واختبار النماذج، بالإضافة إلى دعم المجتمع المفتوح وتبادل المعرفة بشكل فعّال.

## 5.7- Pytorch

هي مكتبة برمجية مفتوحة المصدر، تم تطويرها من قِبل شركة ميتا ، وتُستخدم بشكل واسع في مجالات الذكاء الاصطناعي والتعلم العميق.

تتميز PyTorch بمرونة عالية في التعامل مع البيانات والنماذج، وتوفر بيئة برمجية بسيطة وسهلة الاستخدام، مما يجعلها خياراً مثالياً للباحثين والمطورين. تقدم PyTorch دعماً قوياً لحوسبة المصفوفات والعمليات الحسابية المتقدمة عبر وحدات معالجة الرسوميات (GPU)، كما تتيح إمكانية بناء وتدريب الشبكات العصبية بسهولة، مع توفير نظام حساب تلقائي للتدرجات (AutoGrad). وتُستخدم PyTorch بشكل واسع في الأبحاث وتطوير التطبيقات في مجالات مثل معالجة اللغة الطبيعية، رؤية الحاسوب، والروبوتات، ما يسرّع الابتكار والتطوير في تقنيات الذكاء الاصطناعي.

## 6.7- FastAPI

هو إطار عمل (Framework) حديث ومفتوح المصدر، مخصص لبناء واجهات برمجية (APIs) سريعة وعالية الأداء باستخدام لغة بايثون.

يتميز FastAPI بسهولة الاستخدام، والتوافق الكامل مع معايير OpenAPI وJSON Schema، ما يسهّل توثيق واختبار التطبيقات بشكل تلقائي. يستخدم FastAPI مكتبة Pydantic للتحقق من صحة البيانات وإدارة النماذج، ويدعم بشكل كامل البرمجة غير المتزامنة (Asynchronous Programming)، مما يسمح بتحقيق أداء عالٍ و توسّعية كبيرة في تطوير تطبيقات الويب. بفضل أدائه العالي وسهولة استخدامه، أصبح FastAPI من أكثر الأدوات شهرة وانتشارًا في بناء الخدمات الخلفية (Backend) للأنظمة الحديثة وتطبيقات الذكاء الاصطناعي، وكذلك في تقديم حلول سريعة وآمنة للمطورين.

## 7.7-React

هي مكتبة برمجية مفتوحة المصدر تم تطويرها بواسطة شركة فيسبوك (ميتا حاليًا)، تُستخدم في بناء وتطوير واجهات المستخدم (UI) لتطبيقات الويب الحديثة بطريقة سهلة ومنظمة. تعتمد React على مفهوم المكونات (Components)، التي تتيح تقسيم واجهات المستخدم إلى أجزاء صغيرة قابلة لإعادة الاستخدام والصيانة بسهولة، مما يوفر الوقت ويزيد من كفاءة التطوير. كما تتميز React باستخدام نموذج البرمجة التصريحية (Declarative Programming)، وتوفر آلية الـ Virtual DOM، مما يُسرّع أداء التطبيقات ويضمن تجربة مستخدم سلسة وسريعة. تُستخدم React بشكل واسع في الشركات العالمية لتطوير تطبيقات ويب متقدمة وقابلة للتوسّع، وتتميز بوجود مجتمع كبير ونشط من المطورين الذين يساهمون في دعمها وتطوير أدواتها بشكل مستمر.

## 8.7- محاكي CARLA

هو محاكي قيادة ذاتية مفتوح المصدر. تم تطويره من الصفر ليكون منصة مرنة وقابلة للتخصيص، تهدف إلى دعم مجموعة متنوعة من المهام المتعلقة بمشاكل القيادة الذاتية. أحد الأهداف الأساسية هو المساهمة في ديمقراطية البحث والتطوير في مجال القيادة الذاتية، من خلال تقديم أداة يمكن للجميع الوصول إليها وتخصيصها بسهولة. لتحقيق ذلك، يجب على المحاكي تلبية متطلبات الاستخدامات المختلفة (مثل تعلم سياسات القيادة، تدريب خوارزميات الإدراك، وغير ذلك).

يعتمد على محرك Unreal Engine لإجراء المحاكاة، ويستخدم معيار ASAM OpenDRIVE لتحديد الطرق والبيئات الحضرية. يتم التحكم بالمحاكاة عبر واجهة برمجة تطبيقات (API) بلغة Python و ++C، وهي تتطور باستمرار.

يتكون هيكل CARLA من خادم وعميل. الخادم مسؤول عن تنفيذ المحاكاة، بما في ذلك الفيزياء، وتحديث العالم، وعرض الحساسات. أما العميل فيتحكم في منطق السيناريوهات وظروف البيئة. كما أن هناك عدة وحدات إضافية، مثل:

* **مدير الحركة المرورية**: يتحكم في المركبات غير المستخدمة للتعلم، لمحاكاة حركة مرور واقعية.
* **الحساسات**: مثل الكاميرات والرادارات والليدار، وتوفر معلومات مفصلة عن البيئة.
* **دعم ROS وAutoware**: للربط مع بيئات تعلم أخرى.
* **أصول مفتوحة**: مثل خرائط قابلة للتخصيص ومكتبة كائنات.
* **Scenario Runner**: لتوفير سيناريوهات تعلم قابلة للتكرار، واستخدامها في تحديات CARLA.

## 9.7- DeepSeek

هو مشروع مفتوح المصدر يقدّم نماذج ذكاء اصطناعي متقدمة، تركز بشكل أساسي على النماذج اللغوية الكبيرة (LLMs) التي تُستخدم في مجموعة واسعة من التطبيقات مثل معالجة اللغة الطبيعية، و الترميز، والذكاء الاصطناعي التوليدي. تتميز نماذج DeepSeek بكفاءتها العالية في فهم السياقات اللغوية، وتوليد نصوص دقيقة ومتماسكة، وتوفير إمكانات تدريب متقدمة للمستخدمين. يحظى المشروع بدعم واسع من مجتمع المطورين والباحثين بفضل توفير نماذج فعّالة مفتوحة المصدر يمكن تطويرها وتخصيصها بسهولة لمهام متنوعة.

## 10.7- Kaggle

هي منصة مميزة للمطورين ومبرمجي علوم البيانات، إذ توفر بيئة متكاملة ومجانية لكتابة الأكواد وتطوير النماذج عبر بيئة الـ Notebooks التفاعلية التي تدعم لغات برمجية متعددة مثل Python و R.

تقدّم المنصة مصادر برمجية جاهزة، ومكتبات مثبتة مسبقًا مثل TensorFlow، وPyTorch، وscikit-learn، ما يسمح للمستخدمين البدء مباشرةً في تحليل البيانات وتدريب النماذج واختبارها بسهولة. كما تحتوي Kaggle على مكتبة ضخمة من مجموعات البيانات المتنوعة والمتاحة مجانًا للاستخدام في المشاريع والتجارب الشخصية أو العملية.

بالإضافة لذلك، توفر المنصة موارد حسابية ممتازة حيث تقدم 15G GPU P-100.

# الفصل الثامن

# تنجيز النظام

*يعرض هذا الفصل كيفية تنجيز النظام مع تفصيل كل جزء من أجزاءه*

## 1.8- مقدمة

خلال عملية تطوير وبناء النظام يتم رفع الرماز المصدري الخاص بكل مكون باستخدام Git إلى مستودعات (Repositores) حيث يتم تجميع الرماز البرمجي ضمن فرع (branch) يسمى (dev) يرمز إلى مرحلة التطوير(development)، بحيث أنه عند بناء كل مكون يتم إنشاء فرع جديد من الفرع (dev) ويتم دمجها مع الفرع (dev) عند إنجاز هذا المكون كاملاً.

يحقق هذا الأسلوب مبدأ (Continues Integration) حيث يتم تجميع الرماز البرمجي لكل خدمة ضمن مستودع مركزي

يمكن جميع المطورين من الوصول والعمل على نفس الرماز.

## 2.8- الواجهات الأمامية

### 1.2.8- آلية التنجيز

بناء واجهة المستخدم بإستخدام إطار العمل Vite مع مكتبة React، حيث نستطيع الاستفادة من جميع خواص React مع السرعة والأداء العالي الذي يقدمه Vite في عملية التطوير والبناء. إذ لا بد من استخدام إطار عمل حديث يساعد على بناء واجهات تفاعلية لعرض الإحصائيات الزمنية وخرائط المسار في الوقت الفعلي. إذ تم اختيار مكتبة Recharts لبناء الرسوم البيانية التفاعلية ومكتبة Radix UI لتوفير مكونات واجهة المستخدم عالية الجودة.

تم بناء واجهة عرض الإحصائيات والتحليلات من البداية مع الاستفادة من مكتبة ShadCN/UI لعدة أسباب أهمها:

من أجل تحقيق التحكم الكامل بعرض البيانات والإحصائيات الخاصة بنموذج القيادة الذاتية وإدارة التكامل معها.

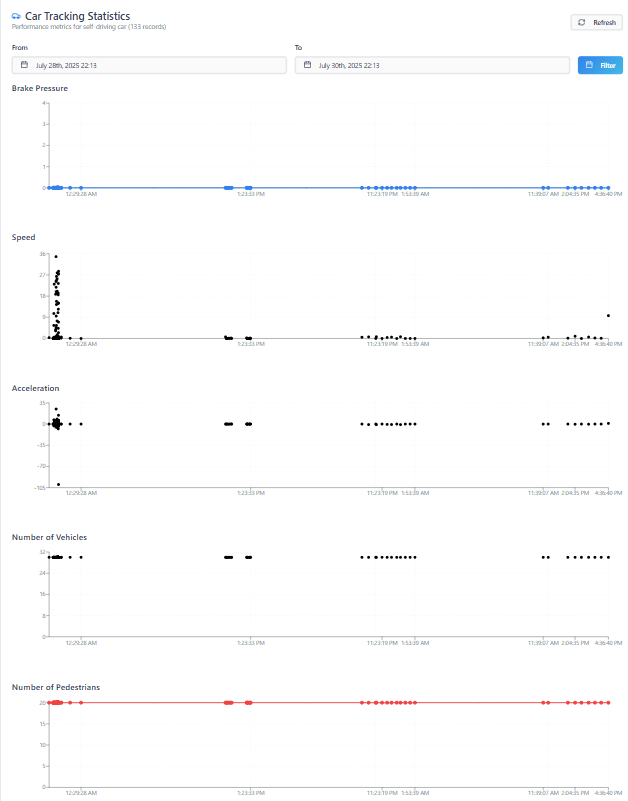
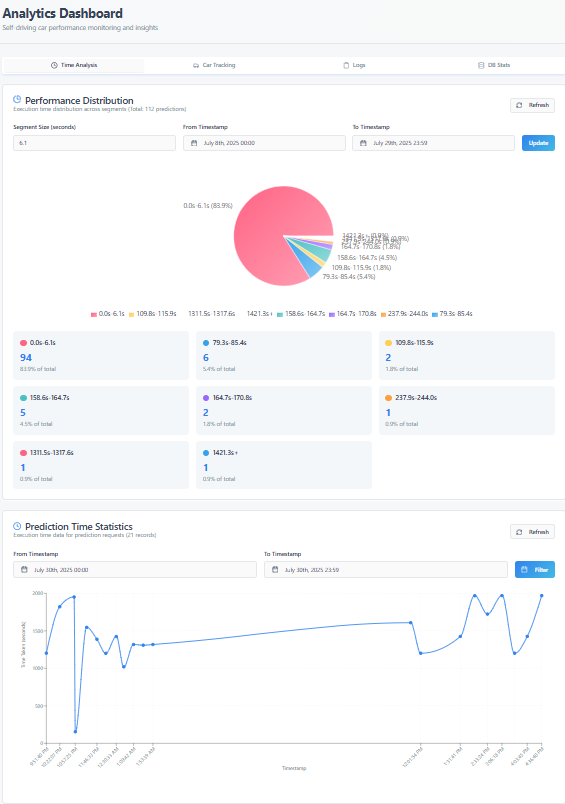
التحكم بتنسيق الخرائط والرسوم البيانية، حيث أن معظم مكتبات التصور الجاهزة لا تدعم متطلبات القيادة الذاتية المتخصصة.

إمكانية تصدير البيانات والتقارير بصيغ متعددة، عن طريق تحويل البيانات المعروضة إلى تنسيقات CSV وJSON للتحليل الخارجي.

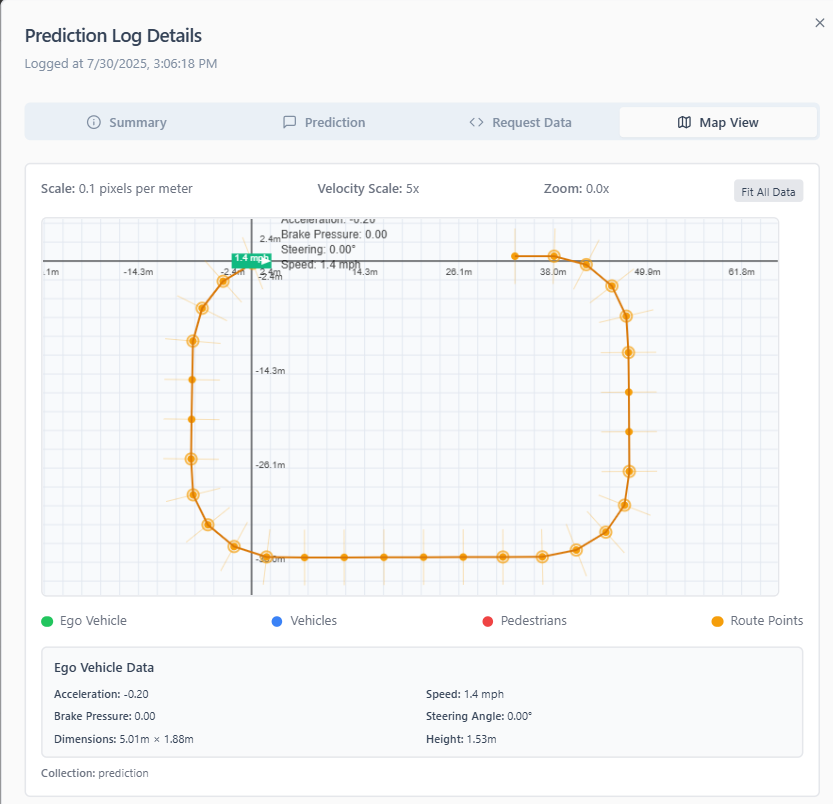
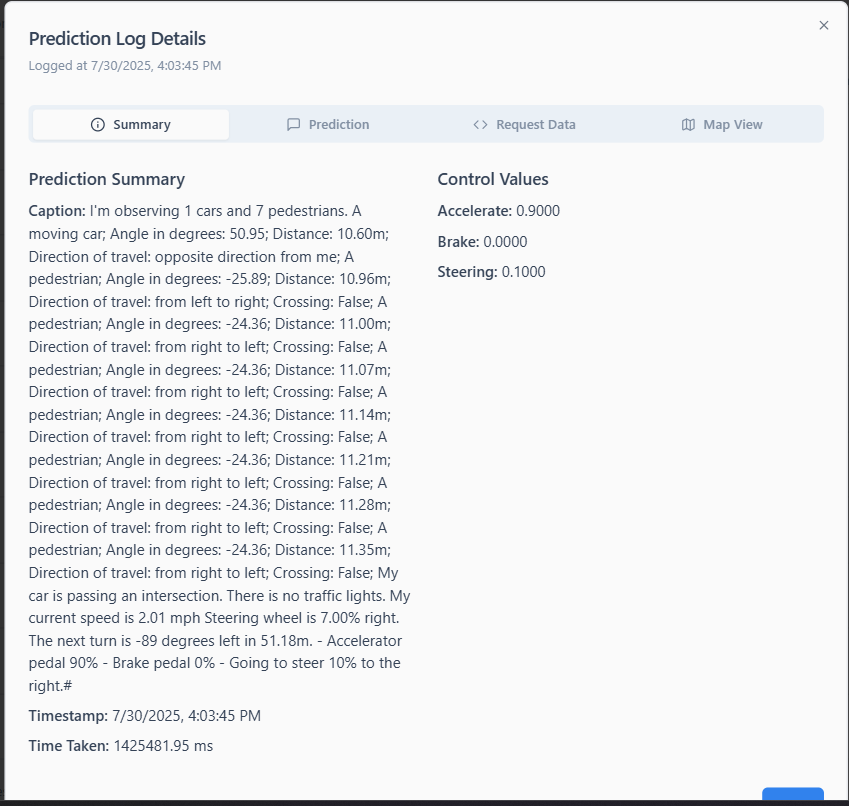
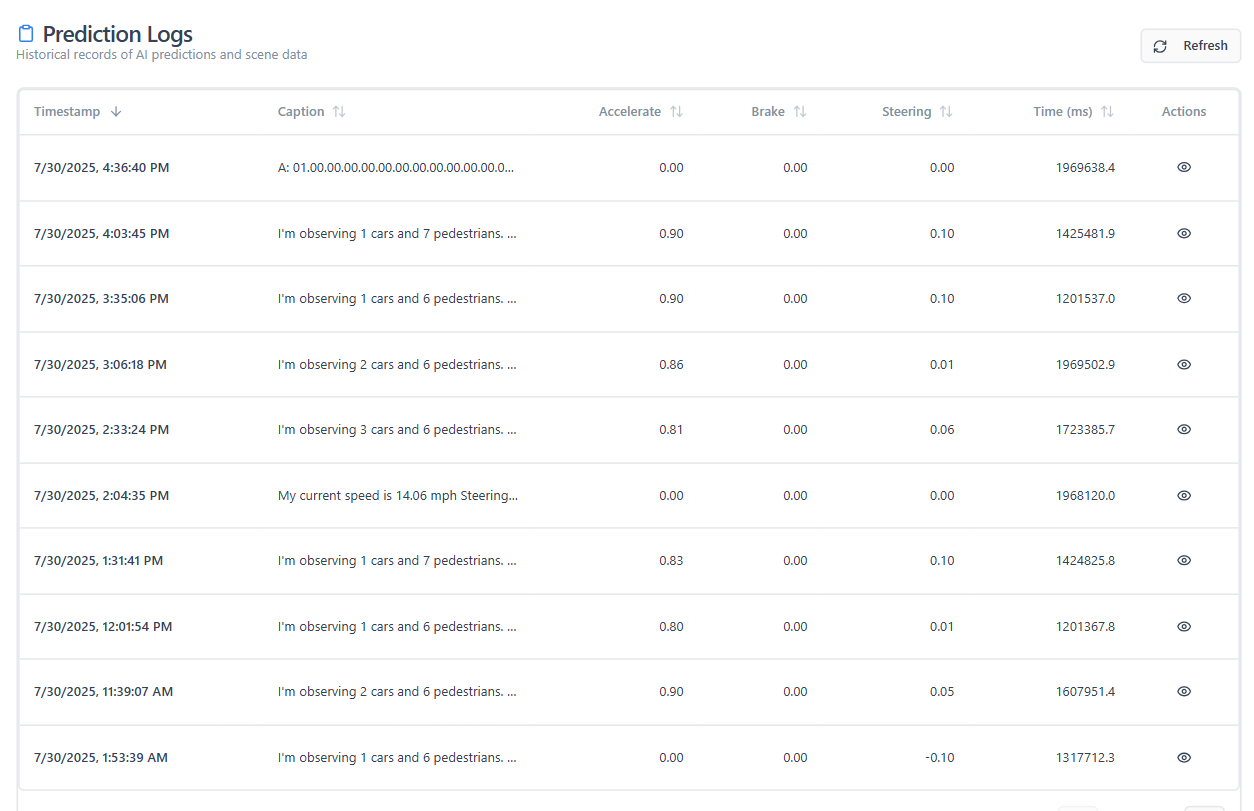
تم دعم خدمة واجهة المستخدم بنظام تصميم موحد باستخدام Tailwind CSS مع دعم الوضع المظلم عبر مكتبة next-themes، وتم تطبيق مبادئ التصميم المتجاوب لضمان عمل الواجهة على جميع الأجهزة والشاشات.

### 2.2.8- بعض واجهات لوحة التحكم

تقدم لوحة التحكم مجموعة من الخدمات من عرض الإحصائيات في الوقت الفعلي، فلترة البيانات حسب النطاق الزمني، تصدير التقارير، عرض خرائط المسار التفاعلية، مراقبة أداء النموذج، وتتبع قيم السرعة والمكابح والالتفاف المقترحة من النموذج.



الشكل (8-1): واجهات لوحات التحكم



## 3.8- الواجهات الخلفية

### 1.3.8- مقدمة

تم بناء هذا المكون باستخدام إطار العمل fastapi بلغة البرمجة python ، تم اختيار python عن غيرها من لغات البرمجة

بسبب الدعم الكبير لها في مجال التعامل والاتصال مع نماذج اللغات الكبيرة إذ تحتوي على العديد من المكتبات أهمها:

(langchain, crewAi, langFlow, groq, …) التي تساعد على بناء التطبيق بشكل سريع وتحقيق مبدأ don’t repeat yourself .

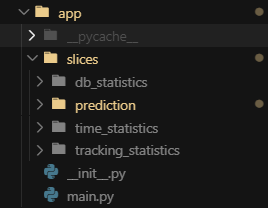
تم الاعتماد على إطار العمل fastapi بدلاً من flask و Django ، لكونه إطار العمل الأسرع كأداء ضمن هذه الإطارات

ولكونه مدعوماً )من حيث المكتبات والمجتمع( أكثر من flask في المقام الأول، وبالرغم من أنه مدعوم بشكل أقل من Django

ولكنه يعد كافياً في مجال بناء RESTAPI ويقدم المطلوب.

### 2.3.8- تفاصيل التنجيز

تم بناء الواجهات الخلفية باستخدام معمارية الشرائح ( Slices Architecture) المبنية على إطار العمل FastAPI، حيث تم تقسيم النظام إلى أربع شرائح متخصصة ومترابطة وظيفياً. تتميز هذه المعمارية بالفصل الكامل بين الطبقات التقنية والوظيفية، مما يضمن قابلية التوسع الأفقي والعمودي للنظام.

الشكل (8-2): تفاصيل التنجيز

**الطبقة الأساسية للتطبيق (Application Layer)**

تم إنشاء نقطة دخول موحدة عبرfastapi مع تكوين شامل لـ CORS middleware لدعم التفاعل مع الواجهة الأمامية. يتم تجميع جميع موجهات الشرائح (Slice Routers) في ملف main.py الرئيسي، مما يوفر نقطة تحكم مركزية لإدارة جميع واجهات التطبيق البرمجية.

**شريحة التنبؤ (Prediction Slice)**

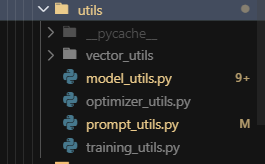
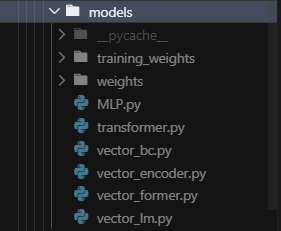
* طبقة نماذج الذكاء الاصطناعي (AI Models Layer)

تم تكامل نموذج deepseek-ai/deepseek-coder-1.3b-base مع تقنية LoRA للضبط الدقيق بمعايير محددة (r=4, alpha=8, dropout=0.05)

استخدام AutoTokenizer مع دعم الرموز الخاصة للتعامل مع البيانات الشعاعية والرموز الخاصة

تحميل النموذج مرة واحدة عند بدء التشغيل مع دعم CUDA للمعالجة المتسارعة عبرmodel\_utils.py

توفير مجموعة من الأدوات المساعدة في app/slices/prediction/ai/utils.

الشكل(8-3): الطبقات الأساسية

* طبقة الخدمات ( Services Layer)

خدمة التنبؤ الرئيسية: تنظم عملية التنبؤ الكاملة من استقبال البيانات إلى إرجاع النتائج

خدمة التسجيل المتقدمة: تحويل البيانات الرقمية إلى نماذج منظمة وحفظها في قاعدة البيانات

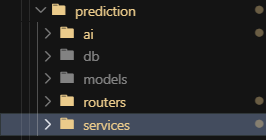
واجهة الخدمات الموحدة تعريف الواجهة الأساسية للخدمات

* طبقة إدارة البيانات (Data Management Layer)

تعريف شامل لعناصر المشهد بما يشمل المركبات، المشاة، الطرق، والمركبة الذاتية

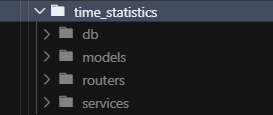
تعريف هياكل البيانات للنتائج والتوقعات

حفظ واسترجاع البيانات التاريخية والإحصائيات

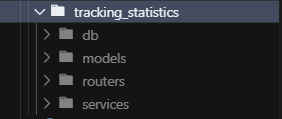


**شريحة الإحصائيات الزمنية**

تتولى قياس وتحليل أزمنة الاستجابة للنموذج، تسجيل أوقات المعالجة، وحساب المتوسطات والانحرافات المعيارية للأداء.

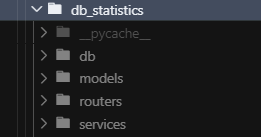
**شريحة إحصائيات التتبع** 

مسؤولة عن معالجة وتحليل مسارات السيارة، حساب المسافات المقطوعة، تحليل أنماط الحركة، وإنتاج البيانات اللازمة لرسم الخرائط التفاعلية.



**شريحة إحصائيات قاعدة البيانات**

تدير عمليات التجميع والتحليل المتقدم للبيانات التاريخية، حساب الإحصائيات التراكمية، وتوفير واجهات للاستعلام والفلترة الزمنية



## 4.8- بيئة المحاكاة

قمنا بتثبيت CARLA على بيئة Docker لتسهيل عملية الإدارة وضمان توافق التبعيات البرمجية مع نظام التشغيل المستخدم. بدايةً، تم تحميل صورة CARLA الرسمية من مستودع Docker Hub وتشغيلها

وللحصول على البيانات المرادة تم بناء تابع للحصول على البيانات المرادة المناسبة كدخل للنموذج اللغوي، ثم قمنا ببناء نيسب لإرسال هذه البيانات إلى النموذج اللغوي بشكل دوري إلى خادم خارجي بواسطة طلب POST، واستقبال الأوامر الحركية وتطبيقها على المركبة الذاتية.

استخدمنا الحساسات الأساسية GNSS و IMU للحصول على بعض بيانات المركبة الذاتية ولكن لم تكن الحساسات الموجودة ضمن البيئة كافية للحصول على البيانات المرادة حيث لا يوجد حساس يقدم إحداثيات x,y,z مباشرة، ولذلك بافتراض أن كل سيارة تمتلك جهاز GPS فيمكننا من خلال مركز بيانات مرتبط بأجهزة الGPS فيمكننا تحديد هذه القيم مباشرة وهو قابل للتحقيق، فلذلك قمنا ببناء هذا الميزة من خلال كتابة رماز برمجي يوفر لنا بيانات x,y,z مباشرة عبر استخدام APIs المقدمة من CARLA

### الملف simulator

يدير المحاكاة الكاملة: تهيئة CARLA، تحميل المشهد، تشغيل الحلقة الرئيسية (loop)، إدارة الحساسات، التعامل مع المستخدم.

### الملف snapshot\_builder

مسؤول عن استخراج التمثيل الهندسي/الشعاعي من عالم CARLA، سواء بيانات السيارة الذاتية أو المركبات والمشاة والمسار.

### الملف traffic\_manager\_utils

يحتوي دوال إضافية مثل: تهيئة المركبات والمشاة، ضبط الإشارات، إعدادات Traffic Manager لجعل بقية السيارات تقود بشكل منطقي.

### الملفapi\_client.py

للتعامل مع واجهة الـ API الخارجية (الإرسال والاستقبال مع النموذج)، وتبسيط الكود في الملف الأساسي.

**الملف controls.py**

لتطبيق الأوامر المستلمة على السيارة الذاتية (throttle, brake, steer)، وضبط كل ما يخص القيادة من أوامر.

# الفصل التاسع

# الاختبارات والنتائج

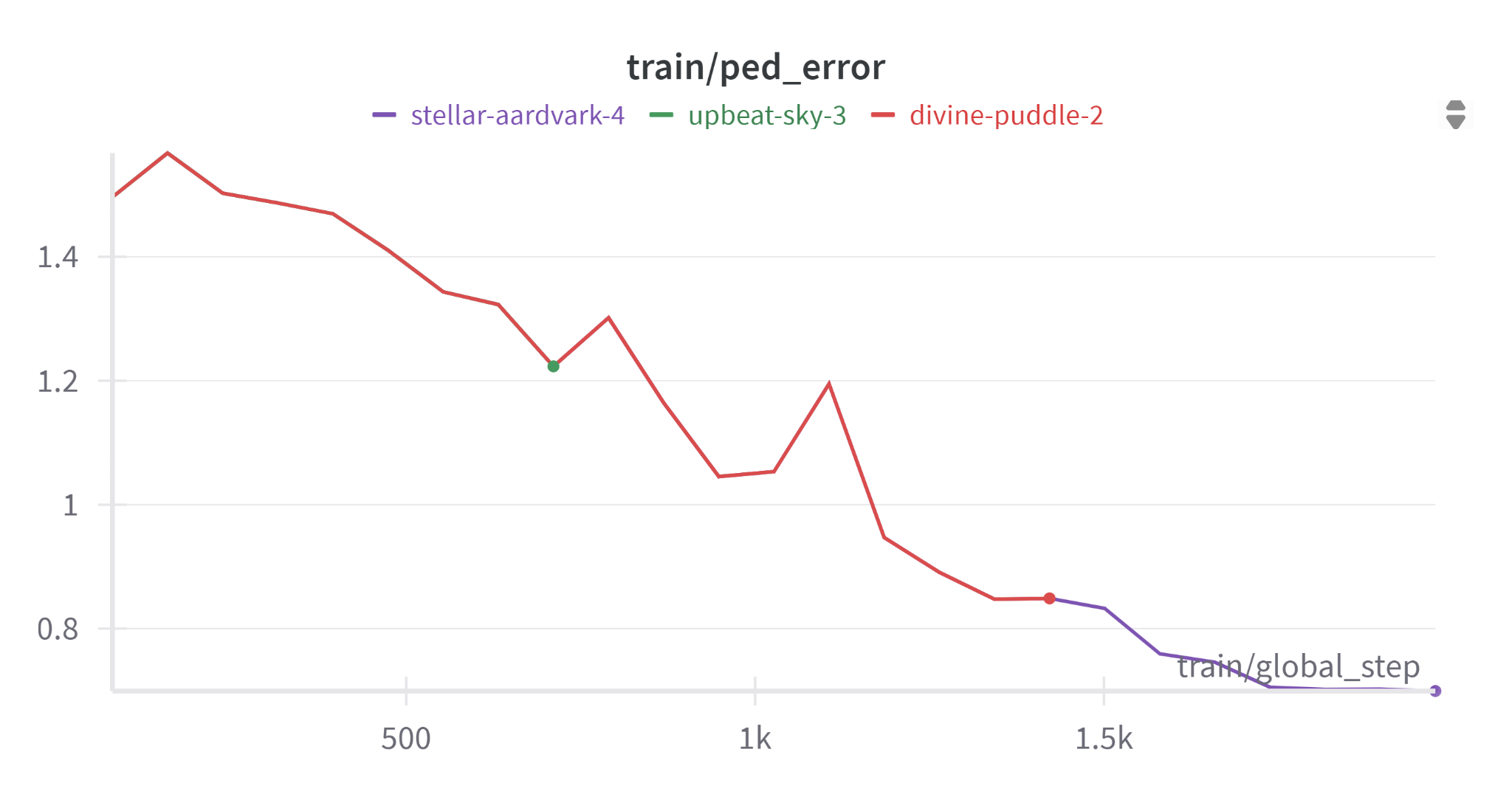
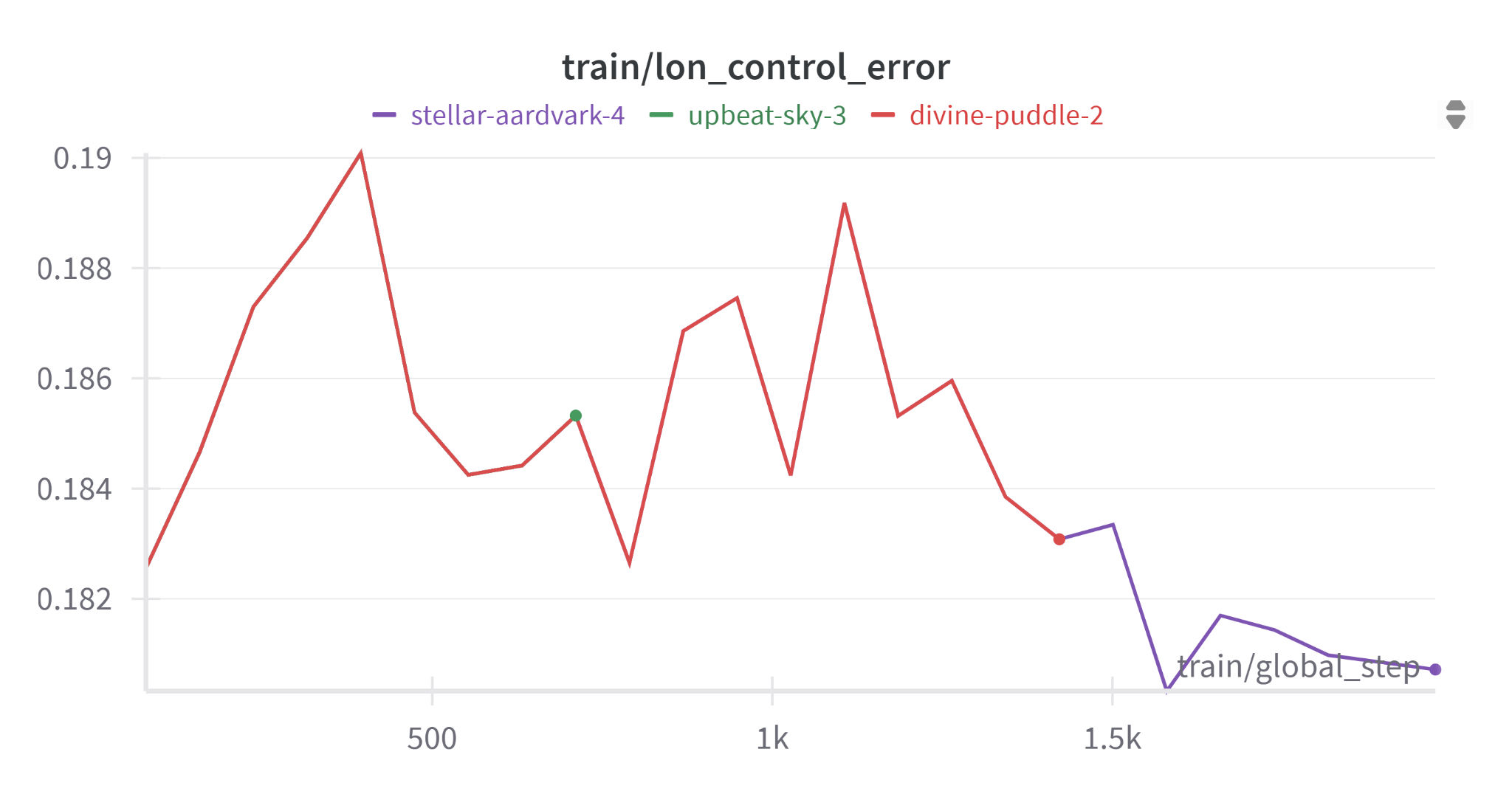
## 1.9- مقدمة

سوف نقوم باستعراض التجارب التي طبقناها والنتائج التي حصلنا عليها

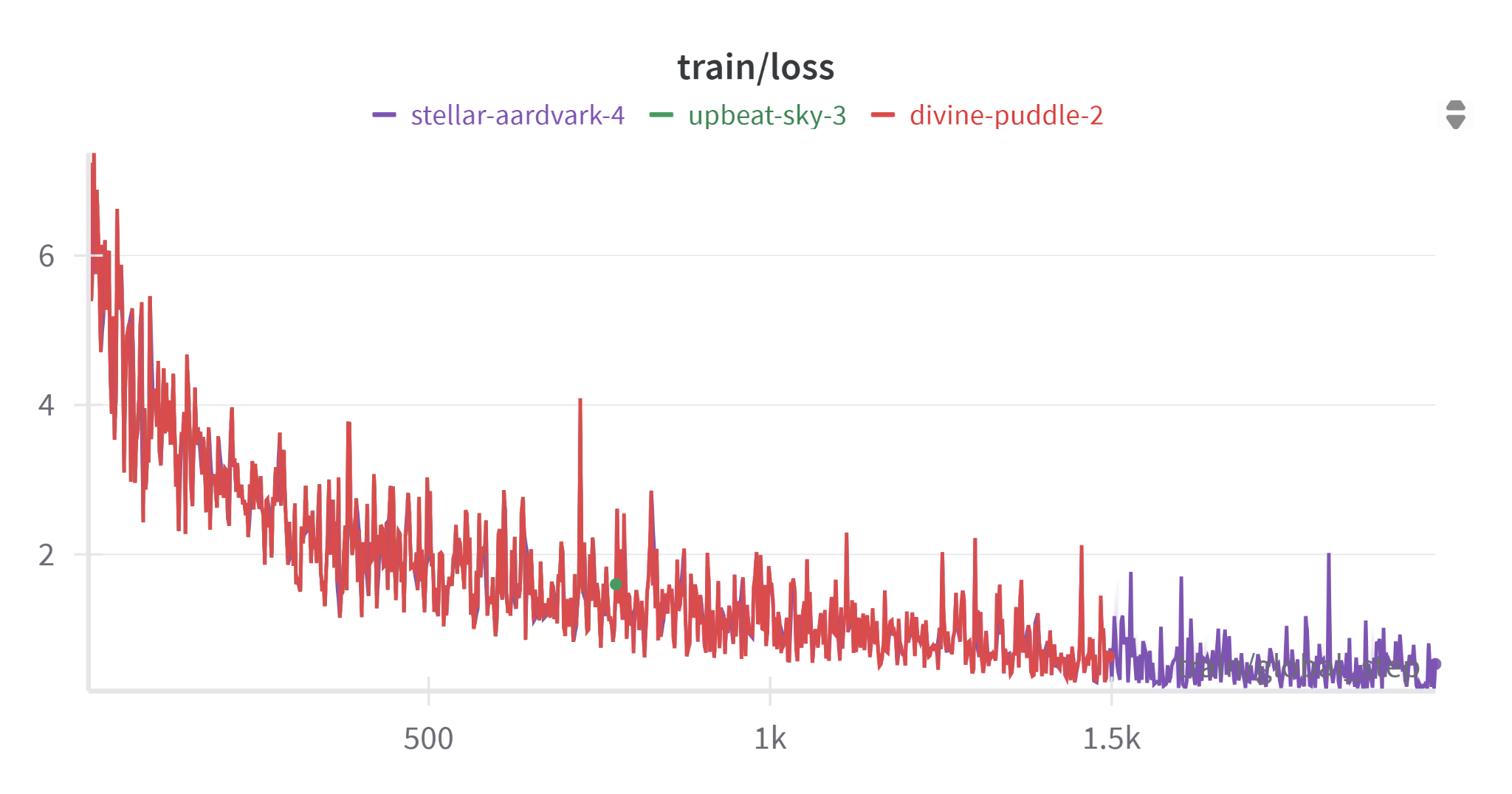
## 2.9- اختبار Vector Encoder و Vector Former

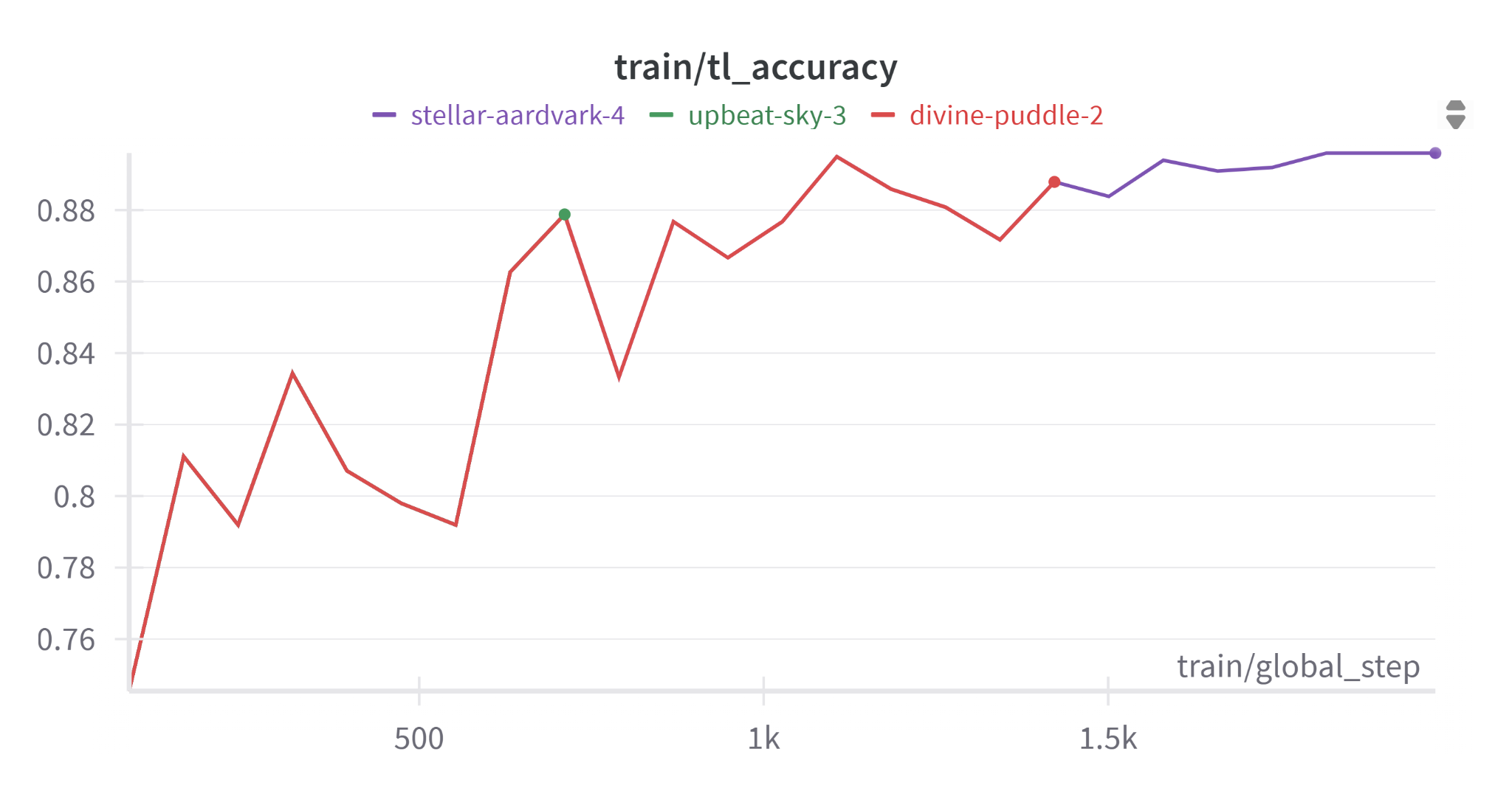
لقد قمت باختبار الجزء المكون من Vector encoder و Vector former وكان الدخل عبارة عن بيانات الدخل نفسها (أشعة المركبات والمشاة والطريق والمركبة الحالية)

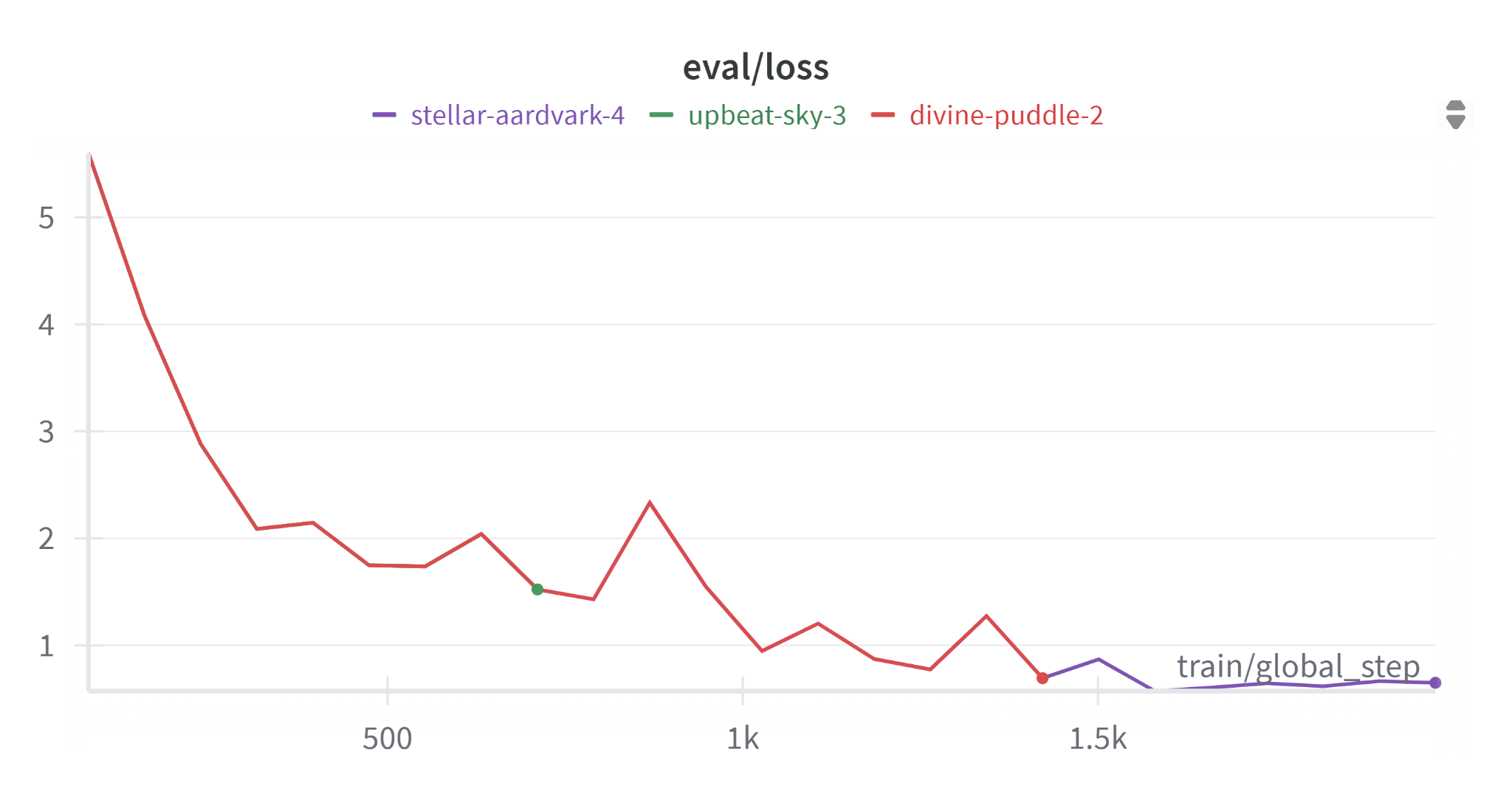
وأما الخرج هو عبارة عن قيم التسارع وقيمة الكبح وقيمة الالتفاف واكتشاف إشارات المرور وتحديد بعدها











## 3.9- مراجعة نتائج Vector Encoder و Vector Former

## 

بناءً على نتائج اختبارات هذا الجزء من الشبكة توضح لنا أنها قادرة على اكتشاف العلاقات بين الأشعة المختلفة وتحديد عدد الأشخاص والمركبات وبعد إشارات المرور والأمر الواجب تنفيذه بدقة ولذلك يمكن الاعتماد عليها وقبولها، ويمكن الاستفادة منها وتوظيفها في مشروعنا.