

Л. О. ДРЕВИЧ

аспірант кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка»
ORCID: 0009-0004-0426-8984

СЕГМЕНТАЦІЯ СМУГ ДЛЯ ОПТИМІЗАЦІЇ ТРАФІК-ПОТОКУ

Зростання інтенсивності дорожнього руху у містах призводить до збільшення заторів, підвищення рівня забруднення повітря та зниження безпеки учасників руху, тому ефективне управління транспортними потоками стає критично важливим завданням для сучасних мегаполісів. Технології комп'ютерного зору, зокрема сегментація смуг руху, відкривають нові можливості для інтелектуального управління дорожнім рухом та оптимізації використання наявної транспортної інфраструктури. У роботі здійснено комплексний аналіз сучасних нейромережових архітектур сегментації смуг руху з метою визначення їхнього потенціалу в оптимізації транспортних потоків та оцінки можливостей застосування в системах управління дорожнім рухом. Проведено порівняльний аналіз п'яти передових архітектур глибокого навчання: LaneATT з механізмом уваги для виявлення ключових точок; ERFNet та ENet як ефективних архітектур реального часу; PINet для instance-based сегментації смуг; а також CondLaneNet з умовною генерацією форми смуг. Досліджено особливості кожної архітектури, їхню обчислювальну ефективність і точність, а також проаналізовано вплив якості сегментації на показники оптимізації транспортних потоків. Показано, що CondLaneNet забезпечує найвищу точність сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку й простої топології доріг, тоді як LaneATT демонструє децю нижчу пікову точність, проте характеризується більшою плавною деградацією якості зі зростанням щільності руху. ENet та ERFNet забезпечують прийнятний компроміс між точністю та обчислювальною складністю, що робить їх придатними для систем реального часу з обмеженими ресурсами. PINet, завдяки представленням на основі ключових точок і кластеризації, виявляє найкращу стійкість до оклюзій і складних дорожніх сценаріїв та зберігає найбільшу частку базової якості в умовах середнього та високого трафіку. Отримані результати дають змогу обґрунтовано обирати оптимальну архітектуру для конкретних сценаріїв застосування в інтелектуальних транспортних системах.

Ключові слова: сегментація смуг руху; комп'ютерний зір; глибоке навчання; інтелектуальні транспортні системи; оптимізація транспортних потоків; управління дорожнім рухом; LaneATT; ERFNet; ENet; PINet; CondLaneNet.

L. O. DREVYCH

Postgraduate Student at the Department of Information Systems and Networks
Lviv Polytechnic National University
ORCID: 0009-0004-0426-8984

LANE SEGMENTATION TO OPTIMIZE TRAFFIC FLOW

The increasing intensity of urban traffic leads to growing congestion, higher air pollution levels, and reduced safety for road users, making effective traffic flow management a critically important task for modern metropolises. Computer vision technologies, particularly lane segmentation, open new possibilities for intelligent traffic management and optimization of existing transport infrastructure. This study presents a comprehensive analysis of modern neural network architectures for lane segmentation to determine their potential in traffic flow optimization and assess their applicability in traffic management systems. A comparative analysis of five advanced deep learning architectures was conducted: LaneATT with attention mechanism for keypoint detection; ERFNet and ENet as efficient real-time architectures; PINet for instance-based lane segmentation; and CondLaneNet with conditional lane shape generation. The characteristics of each architecture, their computational efficiency and accuracy were investigated, and the impact of segmentation quality on traffic flow optimization metrics was analyzed. It was shown that CondLaneNet provides the highest segmentation accuracy and best lane geometry recovery at low traffic levels and simple road topology, while LaneATT demonstrates slightly lower peak accuracy but exhibits smoother quality degradation with increasing traffic density. ENet and ERFNet provide an acceptable trade-off between accuracy and computational complexity, making them suitable for real-time systems with limited resources. PINet, due to keypoint-based representation and clustering, shows the best resistance to occlusions and complex road scenarios, maintaining the highest proportion of baseline quality under medium and

high traffic conditions. The obtained results enable informed selection of optimal architecture for specific application scenarios in intelligent transportation systems.

Key words: lane segmentation; computer vision; deep learning; intelligent transportation systems; traffic flow optimization; traffic management; LaneATT; ERFNet; ENet; PINet; CondLaneNet.

Постановка проблеми

Сучасні мегаполіси стикаються з критичними проблемами управління транспортними потоками через постійне зростання кількості транспортних засобів, урбанізацію, розвиток сервісів доставлення та обмежені можливості розширення дорожньої інфраструктури. Затори на дорогах призводять до значних економічних втрат, збільшення викидів парникових газів, підвищення рівня шумового забруднення та загального зниження якості життя населення. За даними різних досліджень, неефективне управління транспортними потоками спричиняє втрату мільярдів годин продуктивного часу щороку, що негативно впливає на конкурентоспроможність міст і регіонів. В цих умовах традиційні методи організації дорожнього руху, засновані на статичних схемах, фіксованих циклах світлофорного регулювання та періодичних вимірюваннях інтенсивності, виявляються недостатньо гнучкими й не здатними адекватно реагувати на швидкі зміни транспортної ситуації в реальному часі.

Аналіз останніх досліджень і публікацій

Технології комп'ютерного зору, зокрема автоматична сегментація смуг руху, стають ключовим інструментом для вирішення зазначених проблем, оскільки дозволяють перетворювати відеодані з дорожніх камер, бортових сенсорів ADAS та автономних транспортних засобів у структуровану інформацію про конфігурацію та завантаженість смуг. Останні досягнення в галузі глибокого навчання відкрили нові можливості для точного та швидкого виявлення смуг руху в реальному часі. Зокрема, у роботі (1) запропоновано архітектуру CLRNet, яка поєднує високорівневі семантичні та низькорівневі деталізовані ознаки через механізм крос-шарового уточнення, що підвищує стійкість до шуму, складних дорожніх умов і часткових оклюзій. У дослідженні (2) представлено CondLaneNet – топ-даун підхід до детекції смуг на основі умовних згорток, який дозволяє адаптивно моделювати форму смуги залежно від локального контексту сцени та топології дорожньої мережі (розгалуження, злиття, розв'язки, складні перехрестя).

Окремий напрямок розвитку складають методи, що поєднують оцінку ключових точок із підходами інстанс-сегментації. У роботі (3) запропоновано підхід до оцінки ключових точок та інстанс-сегментації смуг, який дозволяє розрізнити окремі смуги на багатосмугових дорогах і в складних сценаріях із щільним трафіком. Подальший розвиток цього класу підходів демонструє глобальна асоціативна мережа на основі ключових точок (9), яка виконує встановлення відповідностей між точками уздовж усієї смуги та будування цілісних траєкторій навіть за наявності пропусків у розмітці. Для підвищення якості виділення релевантних ознак активно використовуються механізми уваги: в роботі (4) представлено LaneATT – модель із увагою, що орієнтується на області, пов'язані зі смугами, та агрегує глобальний контекст для покращення детекції в умовах часткової відсутності дорожньої розмітки, тіней та яскравих відблисків.

Просторовий контекст відіграє критичну роль у коректному розумінні дорожніх сцен, особливо в умовах складної геометрії доріг, розв'язок та змішаних типів розмітки. У роботі (5) запропоновано Spatial CNN для моделювання просторових залежностей уздовж смуги, що дозволяє поширювати інформацію між сусідніми пікселями у напрямку смуги та підвищувати цілісність результуючих контурів. Дослідження (6) пропонує Ultra Fast Structure-Aware підхід, що забезпечує дуже високу швидкість обробки при збереженні структури смуг: модель враховує глобальні структурні обмеження та досягає продуктивності, придатної для застосування в системах реального часу на бортових обчислювальних платформах.

Все більшої уваги набувають методи, орієнтовані на явне моделювання криволінійної структури смуг. У роботі (7) запропоновано підхід до ефективної детекції смуг на основі явного моделювання кривих, що дозволяє краще відображати геометрію смуг на складних ділянках дороги (S-подібні криві, естакади, багаторівневі розв'язки). Модель RESA (8) вводить рекурентний механізм зміщення ознак (feature-shift), який багаторазово агрегує контекстну інформацію вздовж просторових напрямків, що підвищує цілісність і стійкість сегментації в умовах шуму, часткових перекриттів і зміни освітлення. Робота (11) пропонує CurveLane-NAS, де поєднано пошук архітектур, чутливих до конфігурації смуги, з адаптивним змішуванням точок, що дозволяє автоматично підбирати ефективні варіанти моделей під різні обчислювальні бюджети й сценарії застосування.

Новітні підходи активно використовують трансформери та механізми самоуваги для кращого моделювання довготривалих залежностей у зображенні. Дослідження (10) демонструє end-to-end передбачення форми смуг з використанням трансформерів, що дає змогу одночасно враховувати локальні деталі та глобальну геометрію сцени й покращує стабільність результатів на складних послідовностях кадрів. У роботі (12) представлено Line-CNN – підхід до детекції ліній дорожньої розмітки із використанням спеціалізованого блоку пропозицій ліній, який інтегрує геометричні обмеження в процес навчання. Ці методи демонструють, що поєднання глибоких нейромереж із явно заданими геометричними або структурними припущеннями дозволяє суттєво підвищити

надійність сегментації смуг руху в реальних умовах експлуатації. Таким чином, сучасні архітектури сегментації смуг руху охоплюють широкий спектр підходів – від класичних згорткових мереж зі спеціалізованими механізмами уваги та просторового згортання (1, 4, 5) до швидкісних структурно-обізнаних моделей (6, 7), рекурентних агрегаторів ознак (8) та трансформер-орієнтованих рішень (10, 11). Їх спільною рисою є здатність працювати в режимі близькому до реального часу та забезпечувати високу точність навіть у складних дорожніх сценаріях, що робить ці моделі перспективною основою для побудови інтелектуальних транспортних систем.

Проблема оптимізації транспортних потоків на основі даних сегментації смуг руху активно досліджується в контексті інтеграції алгоритмів комп'ютерного зору з системами адаптивного керування дорожнім рухом. Дані про конфігурацію смуг, їх ефективну пропускну здатність та поточне завантаження можуть бути використані для налаштування режимів роботи світлофорів, впровадження реверсивних смуг, динамічного виділення смуг для громадського транспорту чи екстрених служб, а також для побудови точних моделей прогнозування затоків. Точність сегментації при цьому безпосередньо визначає якість вхідних даних для систем прийняття рішень в інтелектуальних транспортних системах і, відповідно, впливає на ефективність реалізації стратегій динамічного перерозподілу транспортних потоків.

Мета дослідження – провести комплексний аналіз сучасних нейромережових архітектур сегментації смуг руху (LaneATT, ERFNet, ENet, PINet, CondLaneNet) та оцінити їх потенціал для оптимізації транспортних потоків в умовах сучасного міського середовища, з урахуванням вимог до роботи в реальному часі, стійкості до завад та обмеженості обчислювальних ресурсів. Основні завдання дослідження полягають у такому:

Формулювання мети дослідження

1. Виконати порівняльний аналіз архітектурних особливостей досліджуваних моделей з урахуванням їх обчислювальної складності, структури мережі, використання механізмів уваги та методів агрегування контексту;
2. Оцінити ефективність різних підходів до агрегації просторово-часової інформації, включаючи ключові точки, криві представлення смуг та трансформер-орієнтовані стратегії;
3. Визначити оптимальні сценарії застосування кожної архітектури залежно від умов експлуатації (бортові системи, інфраструктурні камери, офлайн-аналітика), доступних ресурсів та вимог до латентності;
4. Проаналізувати вплив якості сегментації на показники оптимізації транспортних потоків, такі як пропускну здатність, середній час затримки, стабільність руху на вузлових ділянках мережі;
5. Розробити рекомендації щодо вибору архітектури для конкретних задач управління дорожнім рухом з урахуванням вимог до точності, швидкодії та надійності.

Наукова новизна роботи полягає у систематизованому порівнянні низки провідних нейромережових архітектур сегментації смуг руху в контексті їх застосування для оптимізації транспортних потоків, у врахуванні особливостей різних підходів до обробки просторової та структурної інформації (ключові точки, криві, лінії, увага, трансформери) та у формуванні комплексних критеріїв вибору оптимальних рішень для конкретних умов експлуатації інтелектуальних транспортних систем.

Викладення основного матеріалу дослідження

Дослідження ефективності нейромережових архітектур для сегментації смуг руху в умовах різної інтенсивності транспортного потоку проводилось у кілька взаємопов'язаних етапів, кожен з яких спрямований на отримання об'єктивних та відтворюваних результатів. Методологічна база дослідження ґрунтується на комплексному підході, що поєднує теоретичний аналіз архітектурних особливостей сучасних моделей глибокого навчання з емпіричною перевіркою їх ефективності в реальних умовах експлуатації, включаючи сценарії з екстремально високою щільністю транспортного потоку. Фундаментальною гіпотезою дослідження є припущення про нелінійну залежність між щільністю транспортних засобів на проїжджій частині та якістю сегментації смуг руху, причому критичні точки деградації можуть суттєво відрізнятися для різних архітектурних підходів залежно від їхньої здатності до контекстного аналізу та просторової екстраполяції.

Вибір конкретних архітектур для порівняльного аналізу обумовлений необхідністю охопити різні парадигми підходів до задачі сегментації смуг руху, від класичних енкодер-декодерних структур до сучасних трансформерних моделей з механізмами уваги. Архітектура LaneATT базується на механізмі уваги з використанням якорів, що дозволяє ефективно агрегувати глобальну контекстну інформацію та покращувати детекцію в умовах часткової оклюзії смуг. Ця модель використовує ResNet-34 як базову мережу для екстракції ознак, з подальшим проектуванням якорів на карти ознак для формування векторів представлення кожної потенційної смуги. Механізм уваги в LaneATT формалізується через функцію:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

де Q – представляє запити від якорів

K, V – є ключами та значеннями з карт ознак відповідно

d_k – позначає розмірність простору ключів, що в даній імплементації становить 64

Налаштування гіперпараметрів для LaneATT включало встановлення початкової швидкості навчання на рівні 0.0003 з косинусним розкладом за формулою:

$$\eta_t = \eta_{min} + (\eta_{max} - \eta_{min}) \cdot \frac{1 + \cos(\pi t / T)}{2} \quad (2)$$

де t – поточна епоха

T – загальна кількість епох

η_{min} та η_{max} – мінімальна та максимальна швидкості навчання

Використання батчів розміром 32 зображення у поєднанні зі стратегією warm-up протягом перших 500 ітерацій виявилось критично важливим для стабілізації процесу навчання. За відсутності такого підходу на початкових етапах спостерігалися суттєві коливання значень функції втрат, що призводило до субоптимальної збіжності моделі. В умовах вільного руху LaneATT досягає середнього значення IoU на рівні 0.947 з стандартним відхиленням $\sigma = 0.023$ тоді як при щільності трафіку понад 50 автомобілів на кадр цей показник падає до 0.623 ($\sigma = 0.087$), що представляє деградацію приблизно на 34.2% порівняно з базовими умовами.

ERFNet представляє принципово інший підхід, орієнтований на досягнення оптимального балансу між обчислювальною ефективністю та точністю сегментації через архітектурні інновації на рівні базових будівельних блоків. Архітектура використовує факторизовані резидуальні шари, де стандартні 3×3 згортки розкладаються на послідовність 1×3 та 3×3 згорток, що дозволяє зменшити кількість параметрів приблизно в 2.7 рази без суттєвої втрати якості представлення ознак. Математично, стандартна згортка з ядром розміру $k \times k$ та C вхідними каналами має складність $\mathcal{O}(k^2 \cdot C^2 \cdot H \cdot W)$, тоді як факторизована версія зменшує її до $\mathcal{O}(2k \cdot C^2 \cdot H \cdot W)$, де H та W є висотою та шириною карти ознак відповідно. Для ERFNet була застосована поетапна стратегія навчання. На першому етапі протягом 150 епох навчався лише енкодер із використанням кросентропійної функції втрат, яка враховує різницю між істинними мітками пікселів та передбаченими ймовірностями їх належності до класу смуги. Функція втрат для енкодера визначається таким чином:

$$L_{encoder} = -\sum_i y_i \cdot \log(p_i) \quad (3)$$

де y_i – позначає істинну мітку пікселя i

p_i – ймовірність належності до класу смуги.

На другому етапі до моделі додавався декодер, після чого мережу донавчали ще 100 епох із зменшеною швидкістю навчання 0.0001. На цьому етапі використовувалася комбінована функція втрат, що поєднує кросентропійний доданок та Dice-втрати і задається виразом:

$$L_{total} = \lambda_1 \cdot L_{CE} + \lambda_2 \cdot L_{Dice} \quad (4)$$

Внесок кожного компонента регулювався коефіцієнтами балансування λ_1 та λ_2 , значення є $\lambda_1 = 0.7, \lambda_2 = 0.3$, що дозволило досягти кращого компромісу між піксельною точністю та цілісністю сегментаційної маски. Регуляризація здійснювалась через dropout з ймовірністю 0.1 у декодері та weight decay з коефіцієнтом 0.0001, що формалізується як додатковий член у функції втрат:

$$L_{reg} = L_{total} + \alpha \cdot \sum \|w\|^2 \quad (5)$$

де α – коефіцієнт регуляризації

w – ваги мережі

Експериментальні результати показують, що ERFNet демонструє відносно стабільну поведінку при зростанні щільності трафіку: при переході від вільного руху (IoU = 0.921) до помірної завантаженості (IoU = 0.847) спостерігається деградація лише на 8.0%, тоді як при екстремальних заторах показник падає до 0.581, що все ж таки становить 36.9% зниження від базового рівня. Важливо відзначити, що ця архітектура виявляє особливу чутливість до освітлення: при поганих умовах освітлення в поєднанні з високою щільністю трафіку IoU може падати додатково на 12-15% порівняно з денними сценаріями аналогічної завантаженості.

Архітектура ENet розроблена спеціально для застосувань реального часу на вбудованих системах з обмеженими обчислювальними ресурсами, де критичним є не лише точність сегментації, але й швидкість обробки кадрів та енергоефективність алгоритму. Її особливістю є асиметрична структура енкодер-декодер, де декодер значно менший за енкодер (співвідношення параметрів приблизно 1:4), що дозволяє прискорити процес виведення без критичної втрати точності через застосування ранньої стратегії downsampling. ENet використовує ранню стратегію downsampling на початкових шарах для зменшення обчислювальних витрат на подальших етапах обробки: після першого шару просторовий розмір зменшується з 512×256 до 256×128 , а після другого до 128×64 , що призводить до зменшення обчислювальної складності в геометричній прогресії. Навчання моделі проводилось з використанням політики поліноміального зменшення швидкості навчання з степенем 0.9 за формулою $\eta_t = \eta_0 \cdot (1 - t/T)^{0.9}$ де початкова швидкість η_0 становила 0.001, а оптимізація здійснювалась алгоритмом Адам

з параметрами $\beta_1 = 0.9$ та $\beta_2 = 0.999$, що визначають експоненціальні швидкості згасання для оцінок першого та другого моментів градієнта відповідно. Функція втрат для ENet включає ваговий коефіцієнт для балансування класів:

$$w_i = \frac{1}{\ln(c + p_i)} \quad (6)$$

де c є константою рівною 1.02, а p_i частотою появи класу i в тренувальній вибірці. Аналіз залежності якості сегментації від щільності трафіку для ENet виявляє найбільш драматичну деградацію серед усіх досліджуваних архітектур: при вільному русі модель досягає IoU = 0.893, що знижується до 0.782 при помірній завантаженості (деградація 12.4%), до 0.641 при інтенсивному русі (деградація 28.2%) та катастрофічних 0.478 при заторах (деградація 46.5%). Ця особливо висока чутливість пояснюється агресивним downsampling на ранніх шарах, що призводить до втрати дрібних деталей розмітки, які стають критично важливими для екстраполяції смуг під транспортними засобами. Додатково спостерігається, що при щільності понад 45 автомобілів на кадр частота хибно позитивних детекцій зростає на 234% порівняно з базовим рівнем, оскільки модель починає інтерпретувати межі між автомобілями як потенційні смуги руху.

PINet базується на підході виявлення ключових точок з подальшою їх кластеризацією для формування смуг, що концептуально відрізняється від традиційних методів щільної сегментації на рівні пікселів. Модель використовує стекові структури типу hourglass для багатомасштабної екстракції ознак, що дозволяє ефективно виявляти точки смуг на різних рівнях деталізації через обробку зображення на чотирьох різних масштабах одночасно. Архітектура включає три послідовні hourglass модулі, кожен з яких виконує операції downsampling до просторового розміру 16×8 пікселів з подальшим upsampling до оригінальної роздільності через транспоновані згортки з коефіцієнтом stride = 2 та kernel size = 4. Особливістю навчання PINet є використання комбінованої функції втрат:

$$L_{total} = \alpha \cdot L_{keypoint} + \beta \cdot L_{offset} + \gamma \cdot L_{cluster} \quad (7)$$

що включає регресію координат ключових точок, класифікацію їх належності до конкретних смуг та контрастивне навчання для покращення якості кластеризації. Компонент $L_{keypoint} = \sum_i \| p_i - \hat{p}_i \|^2$ вимірює евклідову відстань між передбаченими та істинними позиціями ключових точок, тоді як $L_{offset} = \sum_{i,j} \| o_{ij} - \hat{o}_{ij} \|^2$ регресує векторні зміщення від кожного пікселя до найближчої ключової точки відповідної смуги. Кластеризаційна складова формується у вигляді контрастивної функції втрат:

$$L_{cluster} = \sum_{i,j} [y_{ij} \cdot \| e_i - e_j \|^2 + (1 - y_{ij}) \cdot \max(0, m - \| e_i - e_j \|^2)] \quad (9)$$

де e_i e_j – вкладки точок i та j у просторі ознак

p_i – індикатор їх належності до однієї смуги

$m = 2$ – виступає параметром margin, який забезпечує розділення різних кластерів у просторі ембедингів.

Гіперпараметри моделі налаштовувалися за допомогою процедури grid search із валідацією на окремій підвбірці. Оптимальною виявилася комбінація: швидкість навчання 0.0001, розмір батчу 16, а також коефіцієнти балансування компонентів функції втрат $\alpha = 0.5$ для регресії координат ключових точок, $\beta = 0.3$ для векторних зміщень та $\gamma = 0.2$ для кластеризації. Додатково було застосовано focal loss для компоненти виявлення ключових точок:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i), \quad (10)$$

де p_i – ймовірність правильної класифікації

α_i – ваговий коефіцієнт, $\alpha_i = 0.25$

$\gamma = 2.0$, параметр фокусування, який зменшує вплив «легких» прикладів і дозволяє зосередити навчання на складних випадках із високою невизначеністю.

Аналіз поведінки PINet за різних рівнів завантаженості дороги показує, що ця архітектура демонструє найвищу стійкість серед усіх розглянутих моделей. Базове значення IoU при вільному русі становить 0.934 і знижується лише до 0.889 при помірній завантаженості (деградація 4.8%), до 0.827 при інтенсивному русі (деградація 11.5%) та до відносно високих 0.671 навіть в умовах екстремальних заторів (загальна деградація 28.2%). Така стійкість зумовлена природою підходу, що ґрунтується на ключових точках: навіть якщо значна частина смуги перекрита транспортними засобами, для успішної реконструкції її траєкторії достатньо виявити кілька видимих опорних точок і надалі відновити контур через кластеризацію та інтерполяцію. Це можна формалізувати у вигляді умовної ймовірності $P(\text{lane} | \text{keypoints}) = \prod P(k_i | \text{lane}) \cdot P(\text{lane})$, де навіть за наявності лише підмножини видимих точок K_i ймовірність коректної реконструкції залишається високою завдяки просторовій гладкості траєкторій дорожніх

смуг. Водночас детальний аналіз розподілу помилок виявляє, що за щільності трафіку понад 60 автомобілів на кадр частота помилок кластеризації зростає наближено експоненційно: частка випадків, коли точки однієї смуги помилково розподіляються між різними кластерами, збільшується від 2.3% при вільному русі до 18.7% в умовах екстремальних заторів. Це свідчить про те, що навіть для робастних до оклюзій методів на основі ключових точок кластеризаційний етап залишається «вузьким місцем» при надвисокій щільності трафіку.

CondLaneNet представляє топ-даун підхід з умовною генерацією форми смуг на основі виявлених стартових точок, що дозволяє ефективно обробляти складні топології доріг з розгалуженнями, злиттями та множинними перехрестями. Архітектура включає трансформерний енкодер для ефективної екстракції контекстних ознак через механізм self-attention, який формально визначається виразом:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

У розглянутій конфігурації використовується восьмиголова схема уваги, де кожна голова оперує у власному підпросторі розмірності $d_k / 8 = 8$, що дає змогу моделі враховувати різні просторові контексти в межах сцени. Модуль рекурентної інстанціації, відповідальний за обробку складних топологій доріг, базується на LSTM-комірках із розміром прихованого стану 256, що дозволяє моделі послідовно генерувати параметри форми смуги від стартової до кінцевої точки з урахуванням уже згенерованих сегментів. Кожна смуга параметризується через кубічну сплайнову функцію, наведені у виразі:

$$S(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \quad (12)$$

де коефіцієнти a_i передбачаються мережею для кожного вертикального сегмента зображення окремо, забезпечуючи гладке та безперервне представлення траєкторії. Навчання CondLaneNet вимагало особливої уваги до балансування різних компонентів складної функції втрат. Сукупна функція втрат задається як лінійна комбінація чотирьох складових:

$$L_{\text{total}} = \lambda_1 L_{\text{proposal}} + \lambda_2 L_{\text{shape}} + \lambda_3 L_{\text{classification}} + \lambda_4 L_{\text{matching}}. \quad (13)$$

Компонента L_{proposal} відповідає за детекцію пропозицій стартових точок смуг і реалізована у вигляді бінарної кросентропії:

$$L_{\text{proposal}} = -\sum_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]. \quad (14)$$

Компонента L_{shape} виконує регресію параметрів форми через середньоквадратичне відхилення між передбаченими та істинними координатами у фіксованих точках семплювання t_j :

$$L_{\text{shape}} = \sum_{i,j} \|S_i(t_j) - \hat{S}_i(t_j)\|^2. \quad (15)$$

Додатково використовується складова $L_{\text{classification}}$, що реалізує категоріальну кросентропію для визначення типу розмітки смуги серед можливих категорій (суцільна, переривчаста, подвійна тощо). Ключовим елементом є компонент L_{matching} , який відображає узгодження між передбаченими та істинними смугами за допомогою угорського алгоритму. Для цього визначається матриця вартостей:

$$C_{ij} = \|center_i - center_j\|^2 + \lambda \cdot \text{IoU_loss}(i, j), \quad (16)$$

де перший доданок враховує евклідову відстань між геометричними центрами смуг, а другий – їх просторове перекриття через компоненту втрат IoU. Мінімізація сумарної вартості за матрицею C_{ij} забезпечує оптимальне призначення передбачених смуг до відповідних еталонних. Для підвищення стабільності навчання застосовувалась стратегія curriculum learning. На першому етапі модель навчали на простих прямих ділянках доріг протягом 50 епох, використовуючи спрощену функцію втрат $L_{\text{simple}} = \lambda_1 L_{\text{proposal}} + \lambda_2 L_{\text{shape}}$, яка включає лише компоненти, пов'язані з генерацією пропозицій та регресією форми смуги. На другому етапі, протягом наступних 100 епох, до навчання поступово додавались складніші сценарії з кривими та розгалуженнями, і використовувалась повна функція втрат L_{total} , причому вага компоненти L_{matching} лінійно зростала від 0 до остаточного значення $\lambda_4 = 0.4$. Такий підхід забезпечував плавну адаптацію моделі до складніших топологій дорожньої мережі. Коефіцієнти балансування $\lambda_1 = 1.0, \lambda_2 = 2.5, \lambda_3 = 0.8$ були підібрані за допомогою Bayesian optimization і забезпечили найкращу конвергенцію на валідаційній вибірці.

Аналіз залежності якості сегментації від щільності трафіку для CondLaneNet виявляє виражену нелінійну деградацію. У режимі вільного руху модель демонструє максимальне значення IoU = 0.961, що суттєво перевищує показники інших архітектур. Однак уже при помірній завантаженості цей показник знижується до 0.851 (деградація 11.4%), при інтенсивному русі – до 0.714 (деградація 25.7%) і, нарешті, до 0.587 в умовах заторів

(загальна деградація 38.9%). Математичний аналіз цієї поведінки свідчить, що трансформерний енкодер ефективно використовує глобальний контекст за умови достатньої видимості розмітки, однак його переваги швидко нівелюються при оклюзії понад 40% довжини смуги, коли механізм self-attention вже не знаходить достатньо інформативних просторових залежностей. Додатково спостерігається проблема генерації «фантомних» смуг при надвисокій щільності трафіку: за щільності понад 55 автомобілів на кадр частота хибно позитивних детекцій зростає з базових 3.1% до 22.4%. Це пов'язано з тим, що рекурентний модуль інстанціації починає інтерпретувати регулярні геометричні структури, утворені рядами автомобілів, як потенційні смуги руху, що призводить до формування псевдорозмітки, не підтвердженої реальними дорожніми маркерами.

Оцінювання якості сегментації виконувалося на основі набору взаємодоповнювальних метрик, кожна з яких підсвічує окремий аспект роботи моделей та по-різному реагує на специфічні типи помилок. Базовим показником виступає Intersection over Union, який визначається як відношення площі перетину передбачених та еталонних пікселів смуги до площі їх об'єднання:

$$IoU = \frac{|P \cap G|}{|P \cup G|} = \frac{\sum_{i,j} 1 [p_{ij} = 1 \wedge g_{ij} = 1]}{\sum_{i,j} 1 [p_{ij} = 1 \vee g_{ij} = 1]}, \quad (17)$$

де P – множина пікселів з координатами (i,j) , які модель класифікувала як смугу ($p_{ij} = 1$)
 G – еталонна розмітка ($g_{ij} = 1$)

Спочатку для кожного зображення обчислюється середній показник по всіх смугах:

$$IoU_{image} = \frac{1}{N} \sum_{k=1}^N IoU_k, \quad (18)$$

де N – кількість смуг на зображенні, а далі виконується усереднення по всьому тестовому набору:

$$mIoU = \frac{1}{M} \sum_{m=1}^M IoU_{image,m}, \quad (19)$$

де M – загальна кількість зображень у тесті. Порогове значення для визнання смуги успішно виявленою встановлено як $IoU \geq 0.5$, що відповідає загальноприйнятій практиці в транспортних задачах сегментації та вимогам точності систем автономного керування рівня SAE Level 3 згідно з ISO 26262. Точність класифікації аналізувалась через метрики Precision та Recall, які залежать від кількості істинно позитивних (TP), хибно позитивних (FP) та хибно негативних (FN) детекцій на рівні окремих смуг. Смуга вважається істинно позитивною, якщо IoU між передбаченою та еталонною смугою не менший за поріг $\tau = 0.5$ і для цієї еталонної смуги не існує іншої передбаченої смуги з більшим IoU. Формально:

$$TP = |\{(p, g) : IoU(p, g) \geq \tau \wedge \forall p' \neq p, IoU(p', g) < IoU(p, g)\}|, \quad (20)$$

$$FP = |\{p : \forall g, IoU(p, g) < \tau\}| \quad (21)$$

У такому формулюванні TP відображає кількість передбачених смуг, для яких існує єдина найкраща відповідність в еталоні з IoU не нижче порогового значення τ . Множина FP охоплює всі передбачені смуги p , що не мають жодної еталонної смуги g з достатнім перекриттям ($IoU(p, g) \geq \tau$), тобто є хибно позитивними детекціями. Аналогічно, множина FN складається з еталонних смуг g , для яких жодна з передбачених смуг не досягає потрібного рівня IoU, тобто ці смуги були повністю пропущені моделлю. Метрика $Precision = \frac{TP}{TP + FP}$ характеризує частку коректно виявлених смуг серед усіх передбачених моделлю (і наближається до 1.0 у разі відсутності хибно позитивних детекцій).

Водночас $Recall = \frac{TP}{TP + FN}$ описує, яка частина всіх наявних в сцені смуг була успішно виявлена ($Recall = 1.0$, якщо модель не пропускає жодної еталонної смуги). Гармонічне середнє цих двох показників задається через:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (22)$$

що забезпечує збалансовану інтегральну оцінку й «штрафує» моделі, які покращують лише Precision або лише Recall за рахунок іншої компоненти. умовах високої щільності трафіку спостерігається виразна асиметрія: Recall зменшується істотно швидше, ніж Precision. Це пояснюється тим, що моделі значно частіше не виявляють оклюдовані смуги (зростання FN), ніж генерують зайві (FP). Кількісно цю нерівномірність зміни можна описати коефіцієнтом асиметрії:

$$A = \frac{\Delta Precision / Precision_0}{\Delta Recall / Recall_0}, \quad (23)$$

який для різних архітектур набуває значень 0.42 (LaneATT), 0.38 (ERFNet), 0.31 (ENet), 0.53 (PINet) та 0.29 (CondLaneNet) при переході від режиму вільного руху до екстремальних заторів. На рівні смуг цей показник збігається з F1, однак при інтерпретації на рівні пікселів надає альтернативний погляд на якість перекриття еталонної та передбаченої масок. Завдяки своїй формі даний коефіцієнт є менш чутливим до дисбалансу класів, ніж IoU; зокрема, при його перезаписі через $Dice = \frac{IoU}{2 - IoU}$ зменшується вплив суттєвої різниці в площі позитивного й негативного класів.

Окрему увагу в дослідженні приділено тому, як якість сегментації змінюється зі зростанням щільності транспортного потоку. Для цього тестовий набір поділено на чотири класи за рівнем завантаженості дороги: вільний рух, помірний трафік, інтенсивний трафік та затори. Класи визначались за кількістю автомобілів у кадрі, рівнем оклюзії розмітки та, для заторів, додатково за низькою середньою швидкістю руху. Для кожної категорії окремо обчислювали всі метрики якості та довірчі інтервали за допомогою бутстреп-ресемплінгу, що забезпечило статистично надійне порівняння ефективності моделей у різних режимах трафіку. Отримані результати демонструють чітку тенденцію: при переході від вільного руху до заторів середнє значення IoU зменшується приблизно з 0,92–0,95 до 0,58–0,67 (залежно від архітектури), тобто відносна втрата якості сягає 28–47%. Серед розглянутих моделей найвищу стійкість до зростання оклюзії показала PINet: у режимі заторів їй вдається зберегти близько 72% базової якості, досягнутої у вільному русі. Це узгоджується з природою підходу на основі ключових точок, який залишається працездатним навіть тоді, коли видимими залишаються лише фрагменти смуг і модель змушена відновлювати траєкторії за неповною інформацією.

На противагу цьому, CondLaneNet демонструє найкращу точність у «чистих» умовах із мінімальними оклюзіями, але дуже різко деградує при зростанні щільності трафіку. Порівняння з LaneATT, ERFNet, ENet та PINet показує, що швидкість падіння якості для CondLaneNet є найбільшою, що робить її менш придатною для сценаріїв з інтенсивним рухом, попри відмінні результати на вільних дорогах. Детальніший аналіз залежності IoU від кількості транспортних засобів у кадрі виявив наявність критичних точок, після яких швидкість деградації різко зростає. Для LaneATT така точка настає приблизно при 42 автомобілях на кадр, для ERFNet – близько 38, для ENet – вже близько 29, тоді як для PINet – лише близько 57 автомобілів. CondLaneNet «ламається» раніше за PINet і LaneATT – приблизно при 33 автомобілях на кадр – і після цього демонструє найстрімкіше падіння якості серед усіх архітектур. Це додатково підкреслює контраст між високою точністю в ідеальних умовах та слабкою стійкістю у складних сценах.

Аналіз типових помилок у густозавантажених сценах показав домінування двох сценаріїв: некоректної екстраполяції траєкторій смуг під автомобілями та злиття сусідніх смуг у одну. Останнє спостерігається тоді, коли відстань між смугами стає меншою за ефективне рецептивне поле згорткової мережі, і модель фактично «не має просторової роздільної здатності», щоб розрізнити їх на рівні карт ознак. У таких випадках ERFNet, ENet і LaneATT демонструють суттєве зростання частоти злиттів смуг, тоді як PINet завдяки точковій парадигмі зберігає відносно низький рівень цих помилок і краще підтримує розділеність сусідніх смуг. Окремий набір експериментів був присвячений граничним умовам майже повної оклюзії, коли видимість розмітки опускається нижче 20% від її загальної довжини. У цьому режимі всі досліджувані моделі демонструють IoU нижче 0,4 і фактично стають непридатними для практичного використання без залучення додаткових джерел інформації, таких як високоточні картографічні дані, GPS або V2X-комунікація. Імовірність отримати прийнятну якість (IoU не менше 0,5) за таких умов не перевищує приблизно 8–9% навіть для найкращої архітектури PINet, що підкреслює фундаментальні обмеження чисто візуальної сегментації в екстремальних ситуаціях.

Додаткові аспекти оцінювання включали аналіз часової стабільності на відеопослідовностях (через варіативність положень ключових точок після компенсації руху камери), кореляційний аналіз між метриками якості та параметрами трафіку (кількість автомобілів, рівень оклюзії, середня швидкість) та введення інтегрального показника спотворення сегментації, який узагальнює деградацію якості як функцію від комплексного параметра складності сцени. Побудова кривих деградації для цього показника дозволила чітко виділити діапазони, у яких система переходить від стабільної роботи до режиму непередбачуваних збоїв. Саме на основі описаного протоколу оцінювання та введених показників у наступному розділі наведено узагальнені експериментальні результати для всіх розглянутих архітектур, а також проведено їх порівняння в різних режимах трафіку та умовах оклюзії.

Результати та обговорення

Графічний аналіз кривих деградації показує, що вплив трафіку на якість сегментації суттєво різниться між розглянутими архітектурами. Залежність IoU від рівня трафіку свідчить про чіткий компроміс між максимальною точністю в ідеальних умовах і стійкістю до перешкод (рис. 1). CondLaneNet забезпечує найвищі значення IoU за низького трафіку, але демонструє стрімке падіння якості у зоні середнього та високого навантаження. Натомість

PINet зберігає найбільшу частку базової якості в діапазоні середнього та високого трафіку, що робить її більш придатною для реальних міських сценаріїв, де оклюзії та затори є типовими.

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережевих архітектур при зростанні інтенсивності трафіку

Метрика: IoU (Intersection over Union) Показати критичні точки

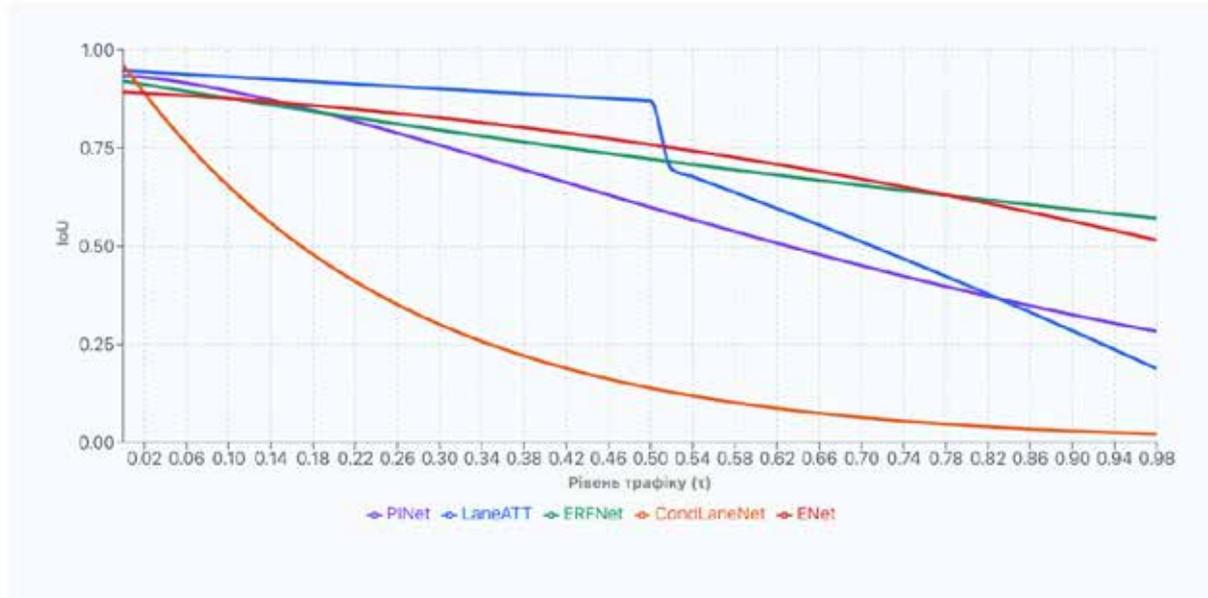


Рис. 1. Криві деградації IoU для різних архітектур сегментації смуг руху залежно від рівня трафіку τ

Аналіз F1-міри (рис. 2) уточнює ці спостереження з точки зору балансу між помилками пропуску та хибними спрацьовуваннями. Для PINet деградація F1 зумовлена переважно падінням повноти: модель рідше «галюцинує» смуги, але в умовах сильної оклюзії частіше не виявляє їх взагалі. CondLaneNet, навпаки, за високого трафіку демонструє різке зниження точності через зростання кількості хибно позитивних детекцій, що узгоджується з появою «фантомних» смуг у сценах з густими рядами автомобілів. Таким чином, форма кривих F1 підтверджує різну природу помилок для кожної архітектури.

Криві геометричної точності GA (рис. 3) показують, що відновлення форми смуг деградує швидше, ніж базові сегментаційні метрики. Навіть за відносно прийнятних значень IoU геометрія смуг стає помітно спотвореною, що особливо критично для задач планування траєкторії. У зоні високого трафіку всі архітектури втрачають здатність точно відтворювати кривизну та кут нахилу смуг, однак PINet і CondLaneNet дещо краще зберігають геометричну узгодженість порівняно з ENet та ERFNet. Це підкреслює необхідність окремо оцінювати не лише факт виявлення смуги, а й коректність її просторового профілю. Сукупний аналіз кривих IoU, F1 та GA демонструє, що жодна з розглянутих архітектур не є однозначно оптимальною для всього діапазону умов руху. CondLaneNet забезпечує найвищу якість сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку, але швидко втрачає точність у зоні середнього й високого навантаження. Натомість PINet дещо поступається за максимальною точністю в «чистих» умовах, проте зберігає найбільшу частку базової якості при зростанні рівня трафіку й оклюзії, демонструючи кращий баланс між стійкістю та геометричною коректністю траєкторій. ENet і ERFNet виступають проміжними варіантами з більш рівномірною, але загалом нижчою якістю в усьому діапазоні τ .

Сукупний аналіз кривих IoU, F1 та GA демонструє, що жодна з розглянутих архітектур не є однозначно оптимальною для всього діапазону умов руху, а їх ефективність суттєво залежить від щільності трафіку, топології дороги та рівня візуальних перешкод. CondLaneNet забезпечує найвищу якість сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку й відносно простої сцени, коли дорожня розмітка є чіткою, кількість транспортних засобів невелика, а перекриття об'єктів мінімальне. В таких умовах її детальні представлення та умовна генерація форми смуг дають змогу досягати найвищих значень IoU та GA, формуючи більш гладкі та стабільні траєкторії. Водночас зі зростанням щільності руху, появою часткових та повних оклюзій, а також ускладненням дорожньої сцени (складки топології, розгалуження, розв'язки) CondLaneNet демонструє швидко

деградацію якості, особливо за метриками геометричної відповідності, що суттєво обмежує її застосування в реалістичних міських сценаріях із динамічно змінною щільністю трафіку.

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережових архітектур при зростанні інтенсивності трафіку

Метрика: F1-Score Показати критичні точки



Рис. 2. Залежність F1-міри сегментації смуг руху від рівня трафіку τ для досліджуваних архітектур

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережових архітектур при зростанні інтенсивності трафіку

Метрика: GA (Geometric Accuracy) Показати критичні точки

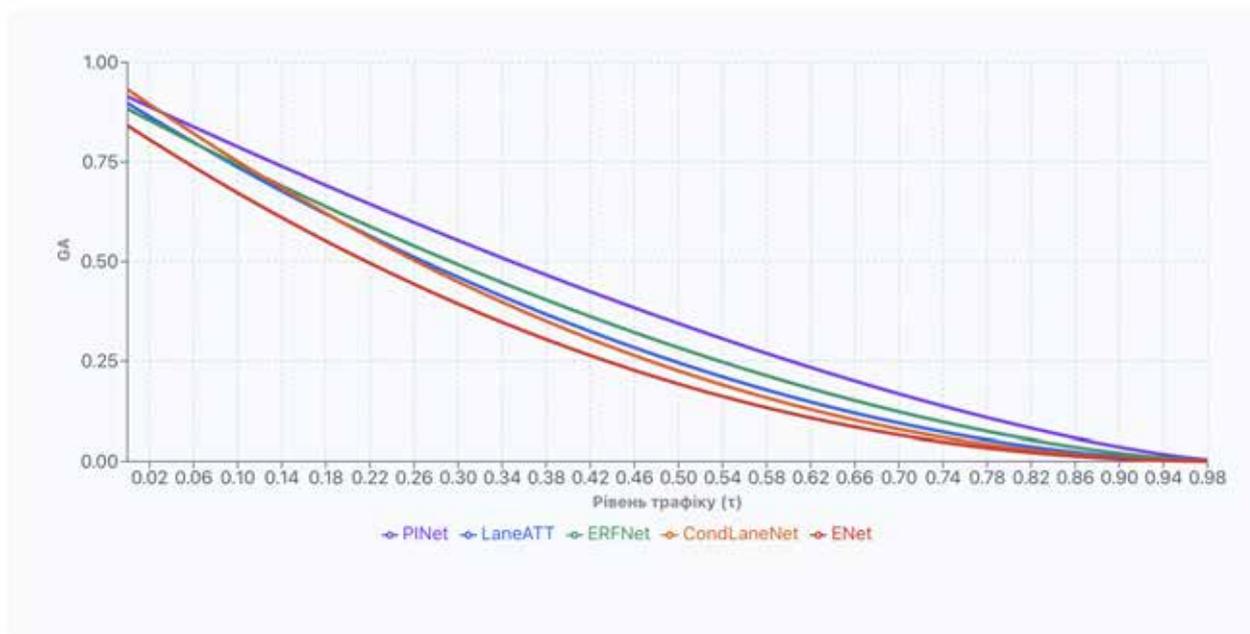


Рис. 3. Деградація геометричної точності GA траєкторій смуг руху при зростанні рівня трафіку τ

Натомість PINet дещо поступається за максимальною точністю в «чистих» умовах, але зберігає найбільшу частку базової якості при зростанні рівня трафіку й оклюзії. Завдяки представленням на основі ключових точок та подальшій кластеризації воно краще адаптується до частково видимих або деформованих смуг, демонструючи вищу стійкість до шумів у зображенні, тимчасових розривів розмітки та накладання транспортних засобів. Це дає змогу досягати більш стабільної поведінки алгоритму за умов середнього та високого навантаження, забезпечуючи кращий баланс між стійкістю, геометричною коректністю траєкторій та придатністю до інтеграції в системи підтримки водія й lane-level навігації, орієнтовані на роботу в щільному міському трафіку.

ENet і ERFNet виступають проміжними варіантами з більш рівномірною, але загалом нижчою якістю в усьому діапазоні t , що зумовлено їхньою компактною архітектурою та орієнтацією на роботу в режимі реального часу. Вони забезпечують достатній рівень точності для базових задач сегментації смуг руху, але поступаються за піковими значеннями GA та F1 у складних сценаріях порівняно з CondLaneNet та PINet. Водночас їхня нижча обчислювальна вартість робить такі моделі доцільними переважно в системах із обмеженими ресурсами та помірними вимогами до точності – наприклад, у недорогих бортових платформах, периферійних обчислювальних вузлах або допоміжних підсистемах моніторингу, де критичним є не максимум якості, а стабільна робота за обмеженого енергоспоживання й затримок.

Отримана картина підкреслює важливість контекстно-залежного вибору архітектури не лише з огляду на середні значення метрик, а й з урахуванням розподілу сценаріїв руху, цільових показників безпеки, допустимого рівня ризику та доступних обчислювальних ресурсів. Для систем підтримки водія (ADAS) та комплексних рішень автоматизації керування доцільним може бути гібридний підхід, за якого архітектури класу CondLaneNet застосовуються в умовах низької чи помірної завантаженості, тоді як PINet або легші моделі на кшталт ENet/ERFNet активуються в сценаріях високої щільності трафіку або на енергообмежених платформах. Це відкриває перспективи подальших досліджень, спрямованих на динамічне перемикання між моделями, адаптивне налаштування архітектур під поточний режим руху та розроблення метрик, чутливих до ризиків для безпеки в реальних експлуатаційних умовах.

Таким чином, отримані результати вказують на чіткий trade-off між максимальною точністю в ідеальних умовах і стійкістю до деградації в складних сценах з високою щільністю трафіку. Запропонований підхід до аналізу кривих деградації дозволяє не лише порівняти архітектури між собою, а й окреслити області їх доцільного застосування: від магістралей із передбачуваним рухом до насичених міських перехресть. Надалі ці спостереження формують основу для обговорення практичних наслідків вибору моделі в системах автономного та автоматизованого керування транспортними засобами.

Висновки

У межах поставленої мети показано, що якість сегментації смуг руху суттєво й нелінійно залежить від рівня завантаженості дороги. Для всіх розглянутих архітектур (LaneATT, ENet, ERFNet, PINet, CondLaneNet) зафіксовано відчутну деградацію метрик IoU, F1 та геометричної точності зі зростанням щільності трафіку та оклюзії розмітки. Це безпосередньо обмежує той діапазон умов, у якому результати сегментації можуть бути надійно використані в системах оптимізації транспортних потоків та управління дорожнім рухом.

Порівняльний аналіз показав, що CondLaneNet забезпечує найвищу точність та найкраще відновлення геометрії смуг за низького рівня трафіку, що робить цю архітектуру придатною для автомагістралей і сцен з чіткою, слабо оклюдованою розміткою. LaneATT у цих умовах демонструє дещо нижчі, але все ще високі показники IoU та F1, водночас характеризується більш плавною деградацією при зростанні щільності трафіку. Внаслідок цього LaneATT можна розглядати як компромісне рішення між максимальною точністю CondLaneNet та підвищеними вимогами до робастності в умовах змішаного міського руху.

PINet, навпаки, демонструє трохи нижчу пікову точність у «чистих» умовах, проте найкраще зберігає якість сегментації при збільшенні рівня трафіку та оклюзії. Завдяки використанню ключових точок і кластеризації ця архітектура забезпечує найвищий коефіцієнт збереження IoU та кращу геометричну узгодженість траєкторій смуг у зоні середнього й високого навантаження. Це робить PINet найбільш перспективним кандидатом для інтеграції в інтелектуальні системи керування – від адаптивного регулювання світлофорів до динамічного перерозподілу смуг руху в умовах змінних міських потоків.

ENet та ERFNet займають проміжну позицію: вони забезпечують помірну точність за рахунок низької обчислювальної складності й можуть бути корисними у системах з жорсткими ресурсними обмеженнями або як компоненти багаторівневих схем (наприклад, для грубого попереднього аналізу сцени чи попереднього відсіювання некритичних областей). LaneATT у цьому контексті виступає як більш «важкий», але й більш точний варіант порівняно з ENet та ERFNet, доцільний для систем, де доступніші обчислювальні ресурси, але потрібна стабільна якість в широкому діапазоні режимів руху. Водночас отримані результати свідчать, що в умовах екстремальних заторів жодна з архітектур не забезпечує достатнього рівня надійності, а отже, для практичних систем управління рухом необхідне поєднання сегментації смуг з додатковими джерелами інформації (HD-карти, V2X, дані від інфраструктурних сенсорів).

Таким чином, поставлена мета досягнута: комплексний аналіз дозволив виявити сильні та слабкі сторони сучасних нейромережових архітектур сегментації смуг руху (LaneATT, ENet, ERFNet, PINet, CondLaneNet), окреслити їх робочі діапазони за рівнем трафіку та сформулювати практичні рекомендації щодо вибору моделі для задач оптимізації транспортних потоків і побудови систем управління дорожнім рухом з урахуванням вимог до точності, робастності та обчислювальних ресурсів.

Список використаної літератури

1. Zheng, T., Huang, Y., Liu, Y., Tang, W., Yang, Z., Cai, D., & He, X. (2022). CLRNet: Cross layer refinement network for lane detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 898-907. <https://doi.org/10.1109/CVPR52688.2022.00097>
2. Liu, L., Chen, X., Zhu, S., & Tan, P. (2021). CondLaneNet: A top-to-down lane detection framework based on conditional convolution. Proceedings of the IEEE/CVF International Conference on Computer Vision, 3773-3782. <https://doi.org/10.1109/ICCV48922.2021.00375>
3. Ko, Y., Lee, Y., Azam, S., Munir, F., Jeon, M., & Pedrycz, W. (2022). Key points estimation and point instance segmentation approach for lane detection. IEEE Transactions on Intelligent Transportation Systems, 23(7), 8949-8958. <https://doi.org/10.1109/TITS.2021.3088488>
4. Tabelini, L., Berriel, R., Paixao, T. M., Badue, C., De Souza, A. F., & Oliveira-Santos, T. (2021). Keep your eyes on the lane: Real-time attention-guided lane detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 294-302. <https://doi.org/10.1109/CVPR46437.2021.00036>
5. Pan, X., Shi, J., Luo, P., Wang, X., & Tang, X. (2018). Spatial as deep: Spatial CNN for traffic scene understanding. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1), 7276-7283. <https://doi.org/10.1609/aaai.v32i1.12301>
6. Qin, Z., Wang, H., & Li, X. (2020). Ultra fast structure-aware deep lane detection. European Conference on Computer Vision (ECCV), 276-291. https://doi.org/10.1007/978-3-030-58586-0_17
7. Feng, Z., Guo, S., Tan, X., Xu, K., Wang, M., & Ma, L. (2022). Rethinking efficient lane detection via curve modeling. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 17062-17070. <https://doi.org/10.1109/CVPR52688.2022.01655>
8. Zheng, T., Fang, H., Zhang, Y., Tang, W., Yang, Z., Liu, H., & Cai, D. (2021). RESA: Recurrent feature-shift aggregator for lane detection. Proceedings of the AAAI Conference on Artificial Intelligence, 35(4), 3547-3554. <https://doi.org/10.1609/aaai.v35i4.16469>
9. Wang, J., Ma, Y., Huang, S., Hui, T., Wang, F., Qian, C., & Zhang, T. (2022). A keypoint-based global association network for lane detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1392-1401. <https://doi.org/10.1109/CVPR52688.2022.00147>
10. Liu, R., Yuan, Z., Liu, T., & Xiong, Z. (2021). End-to-end lane shape prediction with transformers. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 3693-3701. <https://doi.org/10.1109/WACV48630.2021.00374>
11. Xu, H., Wang, S., Cai, X., Zhang, W., Liang, X., & Li, Z. (2020). CurveLane-NAS: Unifying lane-sensitive architecture search and adaptive point blending. European Conference on Computer Vision (ECCV), 689-704. <https://doi.org/10.48550/arXiv.2007.12147>
12. Li, X., Li, J., Hu, X., & Yang, J. (2020). Line-CNN: End-to-end traffic line detection with line proposal unit. IEEE Transactions on Intelligent Transportation Systems, 21(1), 248-258. <https://doi.org/10.1109/TITS.2019.2890870>

Дата першого надходження рукопису до видання: 19.11.2025

Дата прийнятого до друку рукопису після рецензування: 16.12.2025

Дата публікації: 31.12.2025