



**Fakulta elektrotechniky
a informatiky**



Fakulta Elektrotechniky a Informatiky

Katedra kybernetiky a umelej inteligencie

Predmet : **Základy hlbokého učenia**
kurz 2018 / 2019

Zadanie číslo # 1 (esej) :

CNN pri riadení autonómnych vozidiel

klúčové slova : neural network, self-driving car, machine learning,
convolution neural networks

Spracovali :
Filip Hudzík

Obsah

1	Úvod	2
2	Špecifikácie oblasti aplikácie a popis problému	2
2.1	ALVINN	2
2.2	NVIDIA End-to-end systém	3
3	Výber typu neurónových sietí a jej topológia	3
3.1	Topológia konvolučnej neurónovej siete	3
3.2	Hyperparametre a parametre siete	3
3.3	Výber L funkcie a jej matematický popis	5
4	Vyhodnotenie popísaného experimentu	5
5	Zhodnotenie a záver	6

1 Úvod

Inteligentné technológie prenikajú do najrôznejších oblastí ľudskej spoločnosti. Jednou z domén, v ktorej prínos algoritmov umelej inteligencie prináša markantné výsledky sú autonómne vozidlá.

Autonómne vozidlá zmenia spôsob, akým žijeme a pracujeme. Za najväčšie plus je pri autonómnych vozidlách označované zníženie nákladov na spotrebu paliva, zníženie potreby parkovacích miest a transformácia bežného modelu transportu na model zdieľania. Na realizáciu týchto revolučných riešení však budeme v budúcnosti potrebovať obrovské množstvo výpočtového výkonu. Analýza dát a následná optimalizácia môže viesť k zvýšeniu efektívnosti jazdy vozidla taktiež napomáha obmedzeniu alebo úplnému eliminovaniu ľudských chýb a tak k zvýšeniu bezpečnosti dopravy.

2 Špecifikácie oblasti aplikácie a popis problému

Kvolúčnu neurónovú sieť, ktorú budem opisovať pre svoj základ používa poznatky získane pred viac ako 10 rokmi v projekt DARPA Autonomous Vehicle (DAVE). DAVE bol trénovaný na dátach z riadenia, pri jazde prevádzanej človekom. Tréninové dáta obsahovali video z dvoch kamier a riadiace príkazy poslané ľudským operátorom.

DAVE bol v mnohých smeroch inšpirovaný priekopníckou prácou, ktorá v roku 1989 vybuodovala projekt Autonómne pozemné vozidlo ALVINN. ALVINN je predchodcom DAVE a poskytol prvotný dôkaz o koncepcii, že neurónová sieť by bola jedného dňa schopná riadiť vozidlo na verejných komunikáciách.

2.1 ALVINN

Ako už bolo spomínané prvým pokusom o riešenie tohto problému bol projekt ALVINN (autonómne pozemné vozidlo v neurónovej sieti). Riešenie pracuje zo získaným obrazom z kamery a laserového skenera a ako výstup produkuje smer vozidla. Učenie neurónovej siete sa uskutočnilo pomocou simulovaných obrázkov cesty. Architektúra neurónovej siete so spätnou väzbou a jednou skrytou vrstvou, ktorú využíva model ALVINN pozostáva zo 3 vstupov. Prvý snímok je vstupom z kamery, ktorá sníma cestu v rozlíšení 30x32 pozostávajúci z 3002 vstupných neurónov.

Aktivačná funkcia každej jednotky je v pásme modrej farby zodpovedajúcej ceste. Využívanie tohto modrého pásma je veľmi výhodné pretože poskytuje najvyšší kontrast medzi cestou a okolím. Druhým vstupom je taktiež snímok v rozlíšení 8x32, ktorý prichádza z LIDARu. Výstupná vrstva pozostáva z 46 neurónov rozdelených do dvoch skupín reprezentujúcich zakrivenie obvodu pozdĺž ktorého sa má vozidlo pohybovať.

2.2 NVIDIA End-to-end systém

Inžinieri spoločnosti NVIDIA v snahe o napredovanie inovácií vybudovali vlastné riešenie postavené na NVIDIA DRIVE teda škálovateľnej platforme pre autonómne vozidlo. NVIDIA využitím tohto riešenia vo svojej aplikácii používa konvolučné neurónové siete pre predikovanie riadiacích príkazov. Systém je koncipovaný tak, aby sa automaticky naučil potrebné kroky spracovania, ako je napríklad zisťovanie vlastností cesty, pričom ako tréningový signál slúži len uhol riadenia.

Experti teda explicitne nevyškoli, systém aby zistil napríklad obrys cesty. Na rozdiel od iných metód využívajúcich rôzne rozklady problému, ako je napríklad detekcia jazdného pruhu, plánovanie cesty a riadenie, NVIDIA end-to-end systém optimalizuje všetky kroky spracovania súčasne. NVIDIA zvolila tento prístup, keďže je možné pri pomerne malých tréningových dátach od ľudí spoľahlivo riadiť vozidlo a to aj v rôznych prostrediach. Takto navrhnuté riešenie zvláda situácie aj s nejasným vizuálnym obsahom. Takéto situácie môžu nastať napríklad pri jazde na ceste bez označenia pruhov, alebo pri príchode na parkovacie miesta.

3 Výber typu neurónových sietí a jej topológia

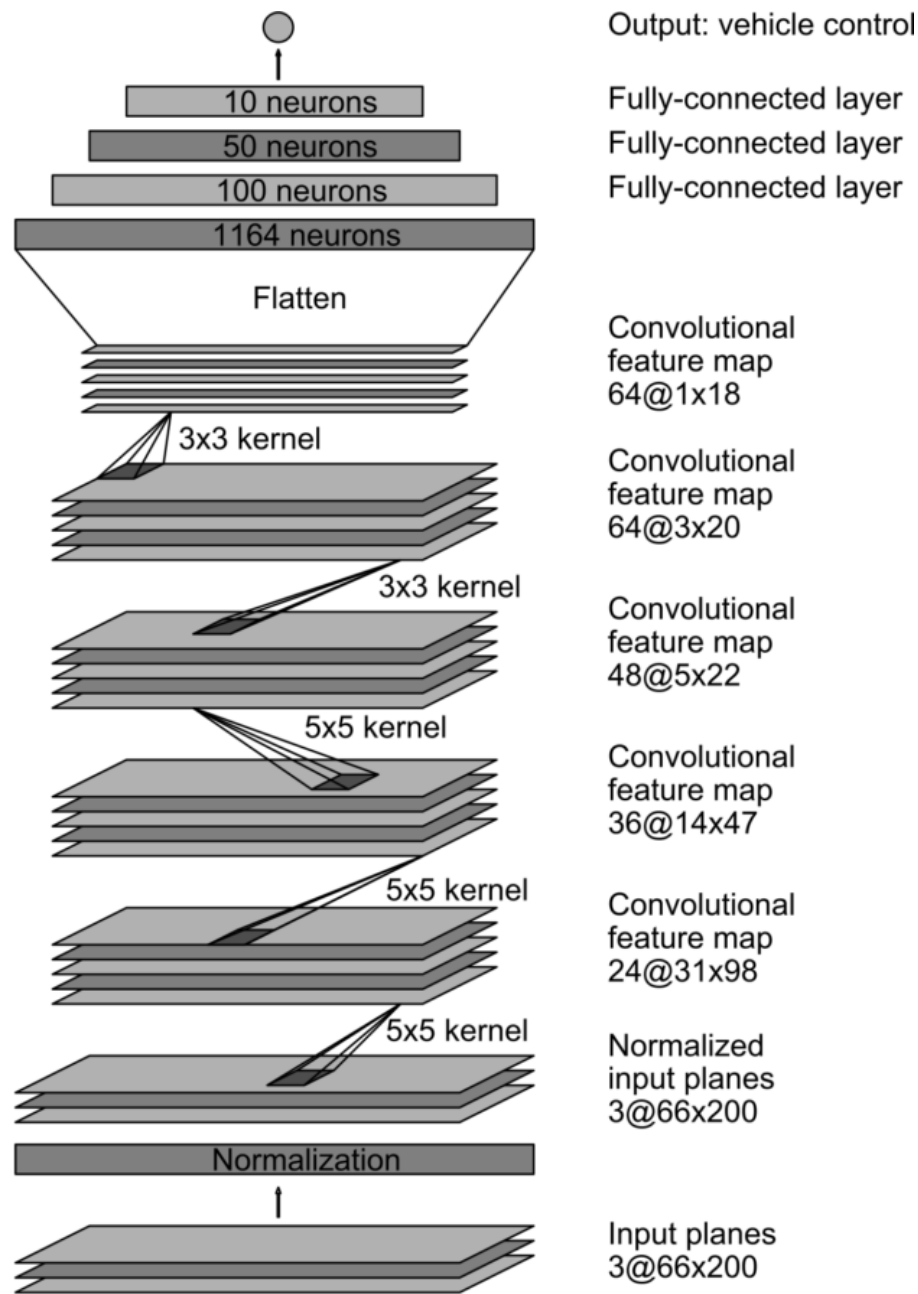
Konvolučné neurónové siete sú špeciálnym druhom neurónových sietí, ktoré sa používajú pre rozpoznávanie vizuálnych vzorov priamo z pixelov obrázka. Cieľom tejto siete je identifikácia objektov na základe podobností vzorov bez ohľadu na deformáciu, zmenu veľkosti alebo posunutia pozície.

3.1 Topológia konvolučnej neurónovej siete

Typická konvolučná sieť sa skladá z niekoľkých vrstiev. NVIDIA vo svojej aplikácii používa konvolučnú neurónovú sieť s architektúrou, ktorá sa skladá z 9 vrstiev, vrátane normalizačnej vrstvy, 5 konvolučných vrstiev a troch plne prepojených vrstiev. Princíp siete spočíva na mapovaní pixelov z prednej kamery umiestnenej na streche vozidla a na základe toho pomerne presne predikuje riadiace príkazy.

3.2 Hyperparametre a parametre siete

Sieť má asi 27 miliónov prepripení a 253 803 parametrov. Vstupný obraz do siete je definovaný ako tri YUV roviny o veľkosti 66x200 pixelov. Prvá vrstva siete vykonáva normalizáciu obrazu. Normalizácia je pevne zadefinovaná a teda nie je upravovaná parametrami v procese učenia. Vykonávanie normalizácie v konvolučnej neurónovej sieti však umožňuje to, aby sa normalizačné schéma spojila so sieťovou architektúrou a sa tak následne výpočet mohol realizovať prostredníctvom spracovania v GPU. Konvolučné vrstvy sú navrhnuté tak, aby vykonávali



Obr. 1: Topológia a popis konvolučnej neurónovej siete

extrakciu príznakov a sú vybrané empiricky pomocou série experimentov. V prvých troch konvolučných vrstvách sa použijú konvolúcie so stride 2×2 a jadrom 5×5 a na zvyšných dvoch non-stride konvolúcia s veľkosťou jadra 3×3 . Hodnota stride udáva posunutie konvolučnej masky. Následne po konvolúcií sieť pozostáva z troch plne spojených (dense) vrstiev, čo vedie ku konečnej hodnote, ktorá je teda inverzným polomerom otáčania vozidla. Plne prepojené vrstvy sú navrhnuté tak, aby fungovali ako riadiaca jednotka pre riadenie.

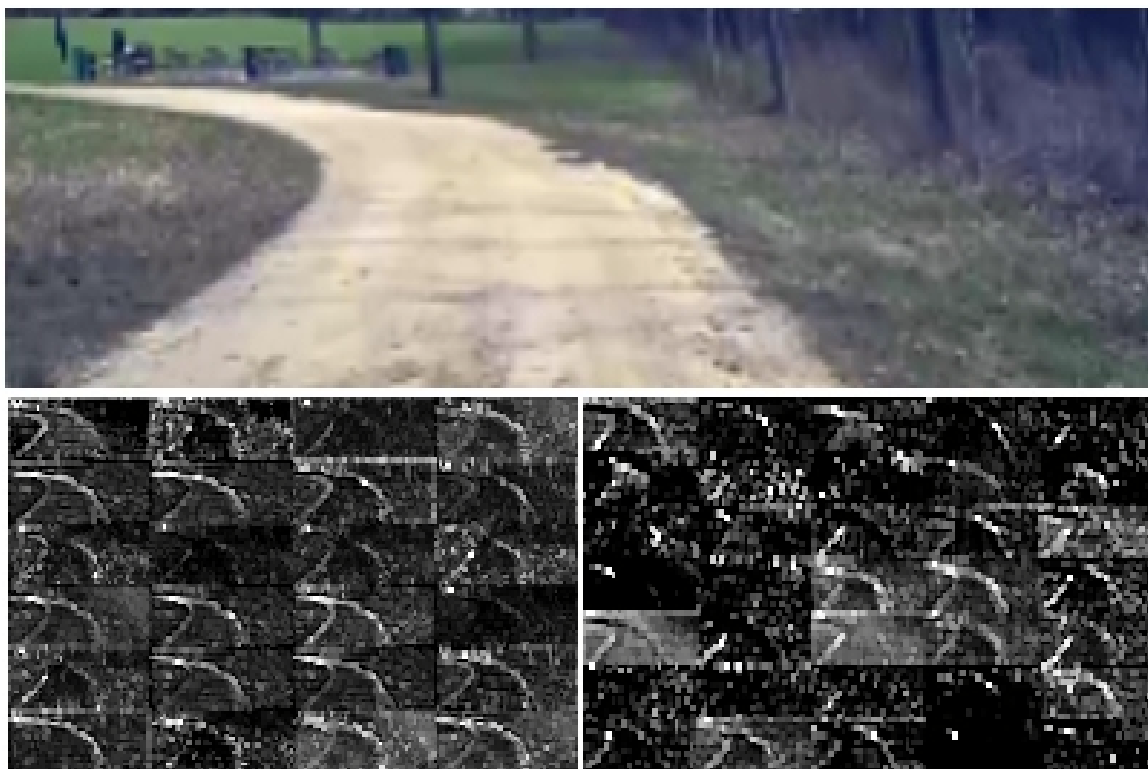
3.3 Výber L funkcie a je matematický popis

Systém je koncipovaný tak, aby sa automaticky naučil roznávať užitočné vlastností z obrazu, pričom ako tréningový signál slúži len uhol ľudského riadenia. Keďže výstupná vrstva obsahuje 10 neurónov môžeme zadať chybovú funkciu L kde $\hat{y}(i)$ je očakávanou hodnotou i -teho neurónu na výstupnej vrstve, teda uhol ľudského riadenia. Keďže uhol riadenia je jednorozmerná hodnota, bude potrebné definovať enkóder, ktorý túto hodnotu zakóduje na rovnaký tvar ako má výstupná vrstva siete a teda 10-rozmerný vektor. Na záver definujeme vypočítanú hodnotou na i -tom neuróne výstupnej vrstvy ako $y(i)$. Chybovú funkciu L potom môžeme definovať takto:

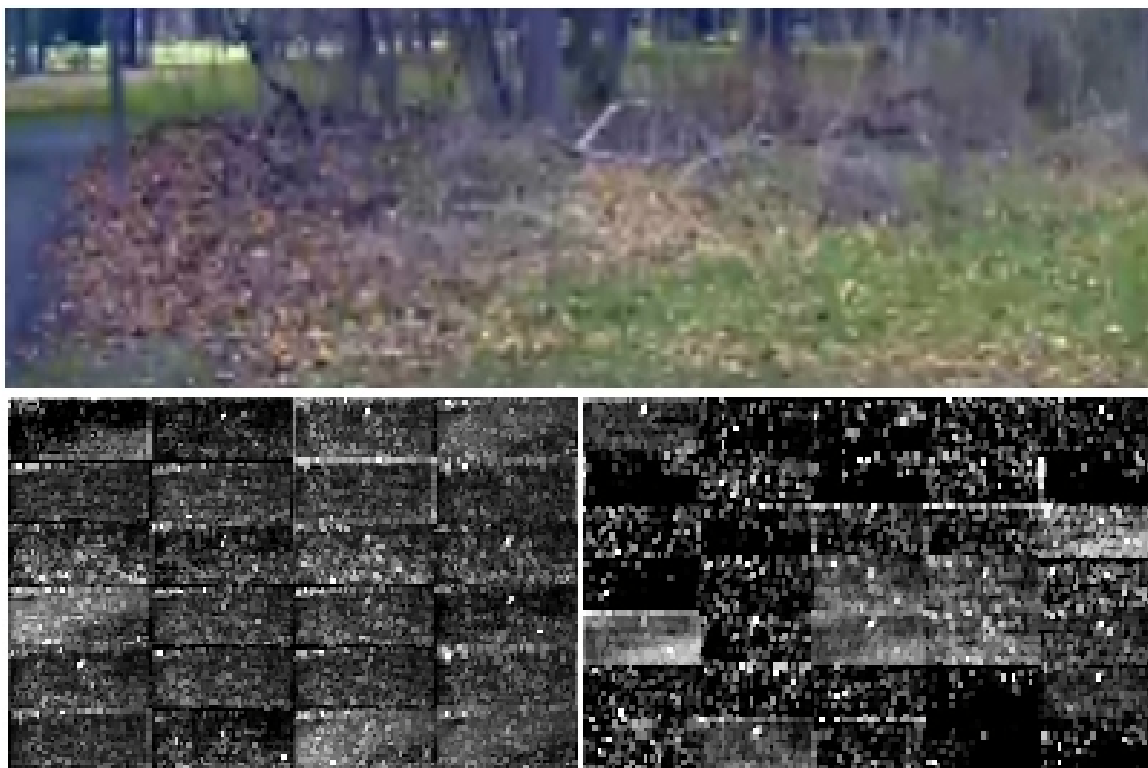
$$L = \frac{1}{2} \sum_{i=1}^{10} (\hat{y}(i) - y(i))^2 \quad (1)$$

4 Vyhodnotenie popísaného experimentu

Výstupy riešenia demonštrujú, že CNN sa naučila extrahovať užitočné vzory cesty sama a to len so vstupnými dátami reprezentujúcimi ľudským uhlom riadenia ako tréningový signál.



Obr. 2: Zobrazte aktivácie prvých dvoch vrstiev



Obr. 3: Zobrazte aktivácie prvých dvoch vrstiev

Obrázky 2 a 3 znázorňujú aktiváciu prvých dvoch vrstiev, ktoré extrahujú užitočné vzory pre dva rôzne príklady vstupov, nespevnenú cestu a les. V prípade nespevnenej cesty, aktivácia prvkov jasne ukazuje obrys cesty, zatiaľ čo v prípade lesa obsahujú prevažne šum. S toho vyplýva, že CNN nenachádza v tomto obrázku žiadne užitočné informácie. Týmto je možné demonštrovať, že CNN sa naučila detekovať užitočné sama a nebola explicitne trébovaná na detekciu obrysov ciest.

5 Zhodnotenie a záver

Táto práca bola venovaná problematike využitia neurónových sietí pre riadenie a navigáciu autonómnych vozidiel. NVIDIA empiricky demonštrovala, že CNN sú schopné pokryť celú úlohu riadenia bez manuálneho rozkladu na detekciu jazdných pruhov alebo ďalších podúloh. Taktiež svojím riešením demonštrovali, že malé množstvo trébových dát z menej ako sto hodín jazdy bolo postačujúce na to, aby bolo vozidlo schopné jazdiť v rôznych podmienkach teda na diaľniciach, miestnych komunikáciach a to za slnečných, oblačných či daždivých podmienok. Systém sa taktiež naučil napríklad zisťovať obrys cesty bez potreby explicitných trébovaných vzoriek počas učenia.

Literatúra

- [1] ALVINN: AN AUTONOMOUS LAND VEHICLE IN A NEURAL NETWORK. Author: Dean A. Pomerleau <https://papers.nips.cc/paper/95-alvinn-an-autonomous-land-vehicle-in-a-neural-network.pdf>
- [2] END TO END LEARNING FOR SELF-DRIVING CARS, Authors: Bojarski, Mariusz and Del Testa, Davide and Dworakowski, Daniel and Firner, Bernhard and Flepp, Beat and Goyal, Praseem and Jackel, Lawrence D and Monfort, Mathew and Muller, Urs and Zhang, Jiakai and others <https://images.nvidia.com/content/tegra/automotive/images/2016/solutions/pdf/end-to-end-dl-using-px.pdf>
- [3] EXPLAINING HOW A DEEP NEURAL NETWORK TRAINED WITH END-TO-END LEARNING STEERS A CAR, Authors: Bojarski, Mariusz and Yeres, Philip and Choromanska, Anna and Choromanski, Krzysztof and Firner, Bernhard and Jackel, Lawrence and Muller, Urs <https://arxiv.org/pdf/1704.07911.pdf>