

Primjena tehnika dubokog učenja u fizici teških iona



Marko Jerčić, Nikola Poljak, Filip Erhardt, Mirko Planinić, Antonija Utrobičić
Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet, Fizički odsjek

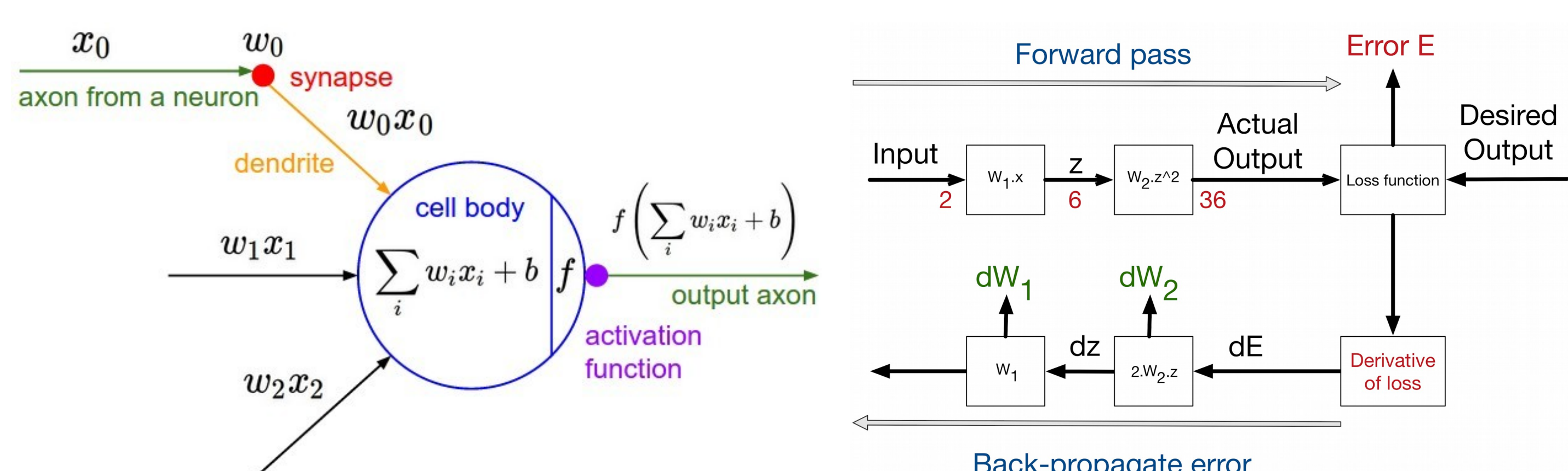


Uvod

Tijekom posljednjih godina metode strojnog učenja uspješno su primijenjene na raznovrsne sustave u fizici teških iona. U takvim sustavima, s velikim brojem stupnjeva slobode, metode strojnog učenja mogu puno bolje djelovati u usporedbi s klasičnim metodama temeljenim na restrikciji nekog broja parametara događaja. Strojno učenje, a osobito tehnike dubokog učenja, omogućuju značajna poboljšanja u klasifikacijskim problemima uzimajući u obzir korelacije u podacima učenjem na događajima iz Monte Carlo simulacija. U ovom radu prikazan je princip rada dubokih neuralnih mreža, algoritama temeljenih na radu ljudskog mozga, koje u posljednje vrijeme doživljavaju ogroman razvoj. Uz to, prikazani su primjeri uspješne primjene tehnika strojnog učenja u fizici teških iona kao što su klasifikacija mlazova [1] i određivanje jednadžbe stanja kromodinamičke (QCD) tvari [2]. Konačno, predstavljene su nove ideje u kojima bi se koristile duboke neuralne mreže za što efikasniju analizu potpisa kvarkovsko-gluonske plazme u sudarima teških iona i sudarima protona.

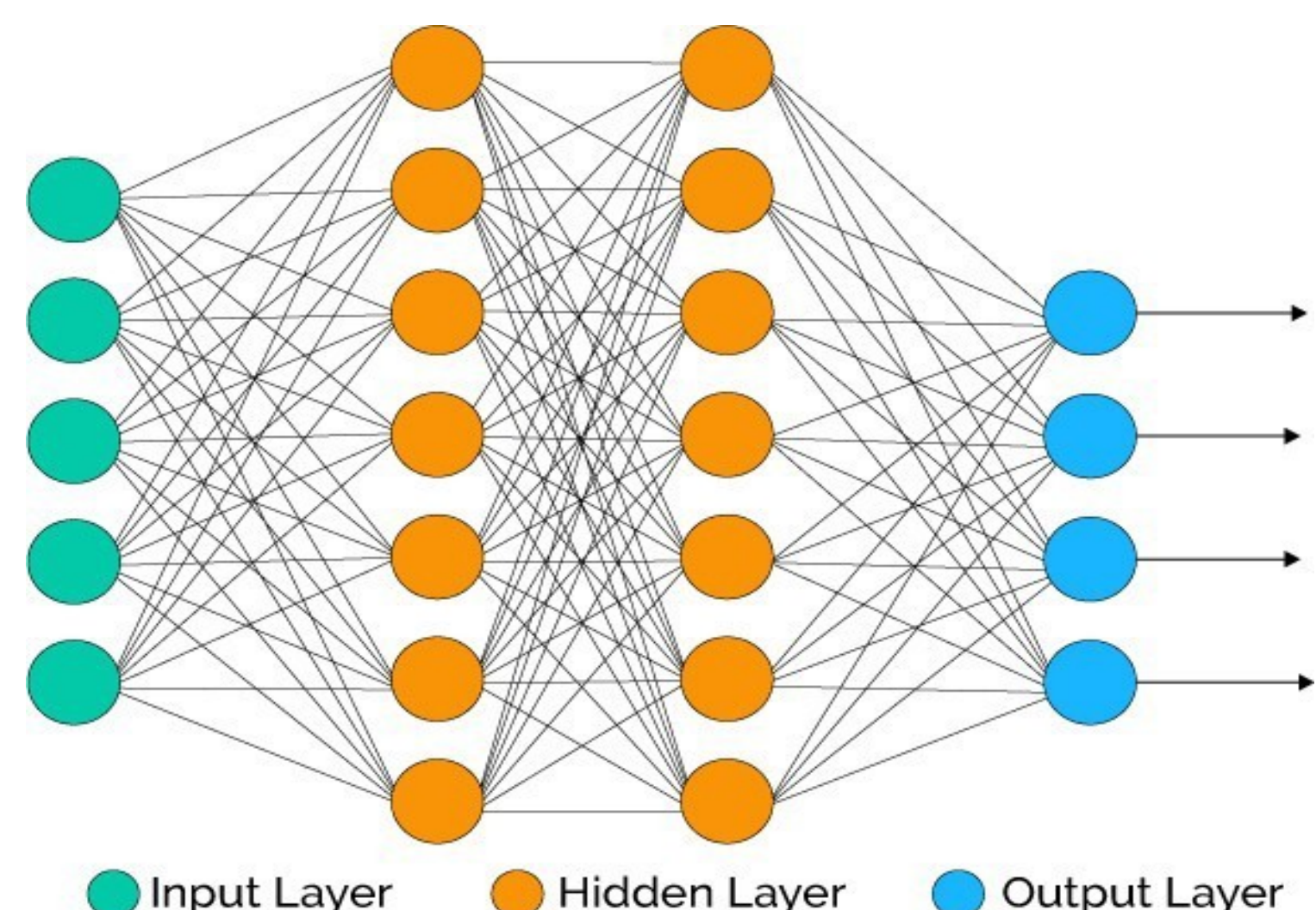
Neuralne mreže

Neuralne mreže su algoritmi temeljeni na radu ljudskog mozga. Takvi sustavi obavljaju zadatke uzimajući u obzir primjere bez poznavanja specifičnih pravila zadatka. Tipične neuralne mreže sastoje se od jednog ili više međusobno povezanih "neurona". Svaki neuron prima signale od ostalih neurona, množi ih s težinom te proslijeđuje nelinearnu funkciju njihove sume.



Princip rada neuralne mreže

Na taj način možemo dobiti apstraktno nelinearno preslikavanje između višedimenzionalnih tenzora. Na primjer, moguće je istrenirati mrežu koja na temelju informacija iz piksela može "prepoznati" životinju te sliku svrstati u neku kategoriju.

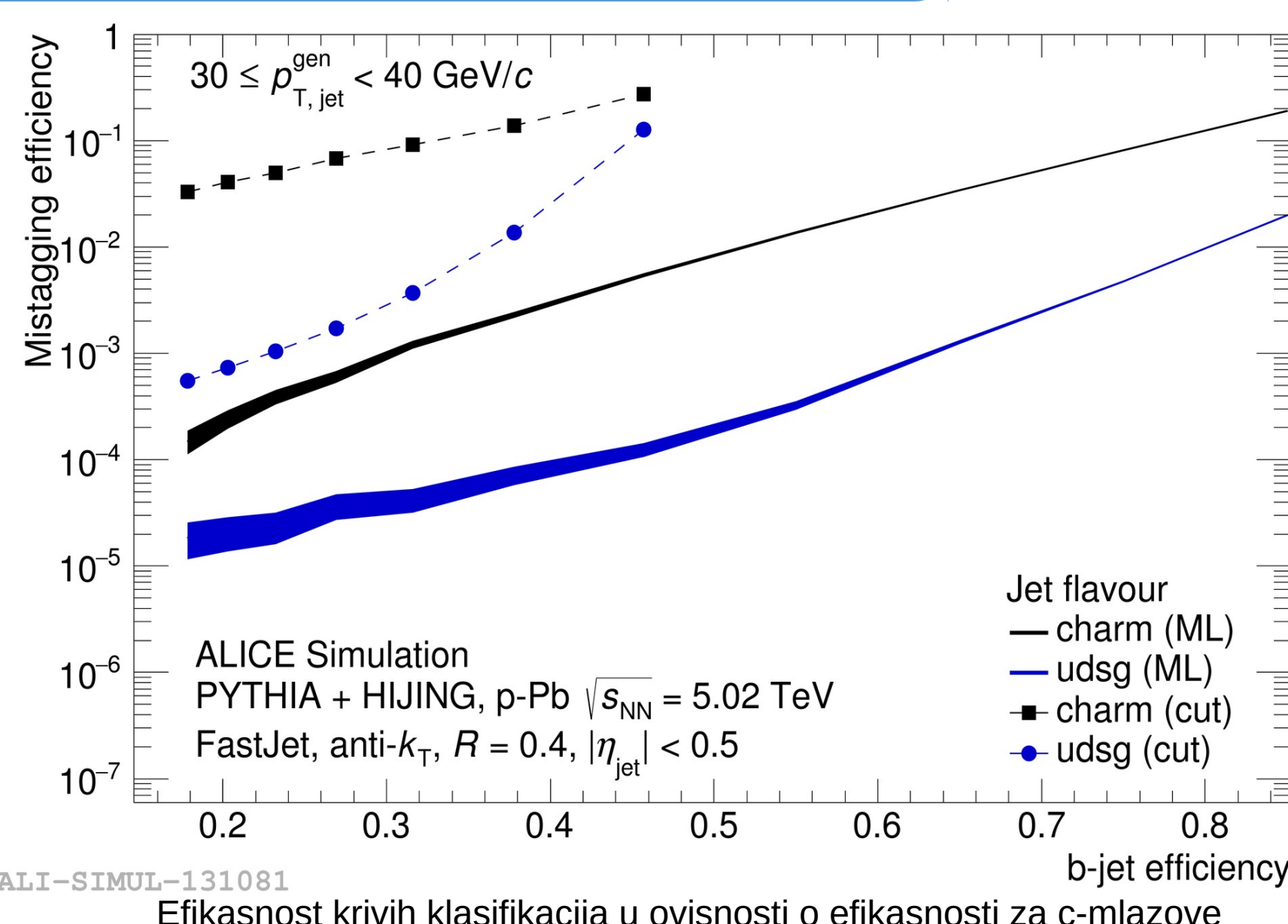


Primjer arhitekture neuralne mreže

Takva preslikavanja se određuju prilagođavanjem težina pojedinih neurona prema primjerima iz baze podataka. Neuralne mreže pokazuju puno bolje rezultate učenjem na velikoj količini podataka. Kako se na velikim ubrzivačima čestica kao što je LHC svakog dana proizvede ogromna količina podataka, eksperimentalna fizika visokih energija može biti obećavajuće područje za primjenu tehnika strojnog učenja.

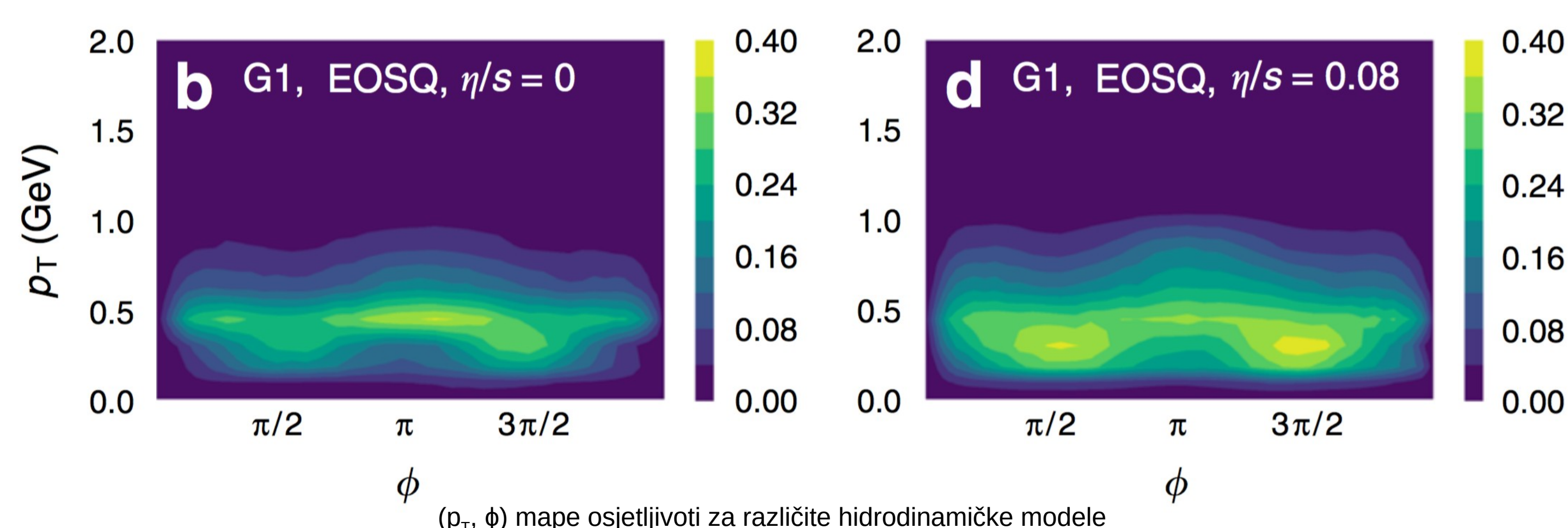
Klasifikacija mlazova

Kvarkovsko-gluonska plazma (QGP) vrući je i gusti medij koji se stvara u sudarima teških iona. Udio pojedinih kvarkova u takvom mediju može dati uvid u svojstva QGP. Analiza okusa mlazova u takvim sudarima može biti moćan alat za određivanje početnog stanja sudara pa tako i glavnih karakteristika QGP.



Na slici je prikazana usporedba algoritma za klasifikaciju c-mlazova temeljenog na dubokim neuralnim mrežama te postojećeg anti- k_T algoritma iz koje vidimo kako neuralne mreže za red veličine manje pogrešno identificiraju mlazove [1].

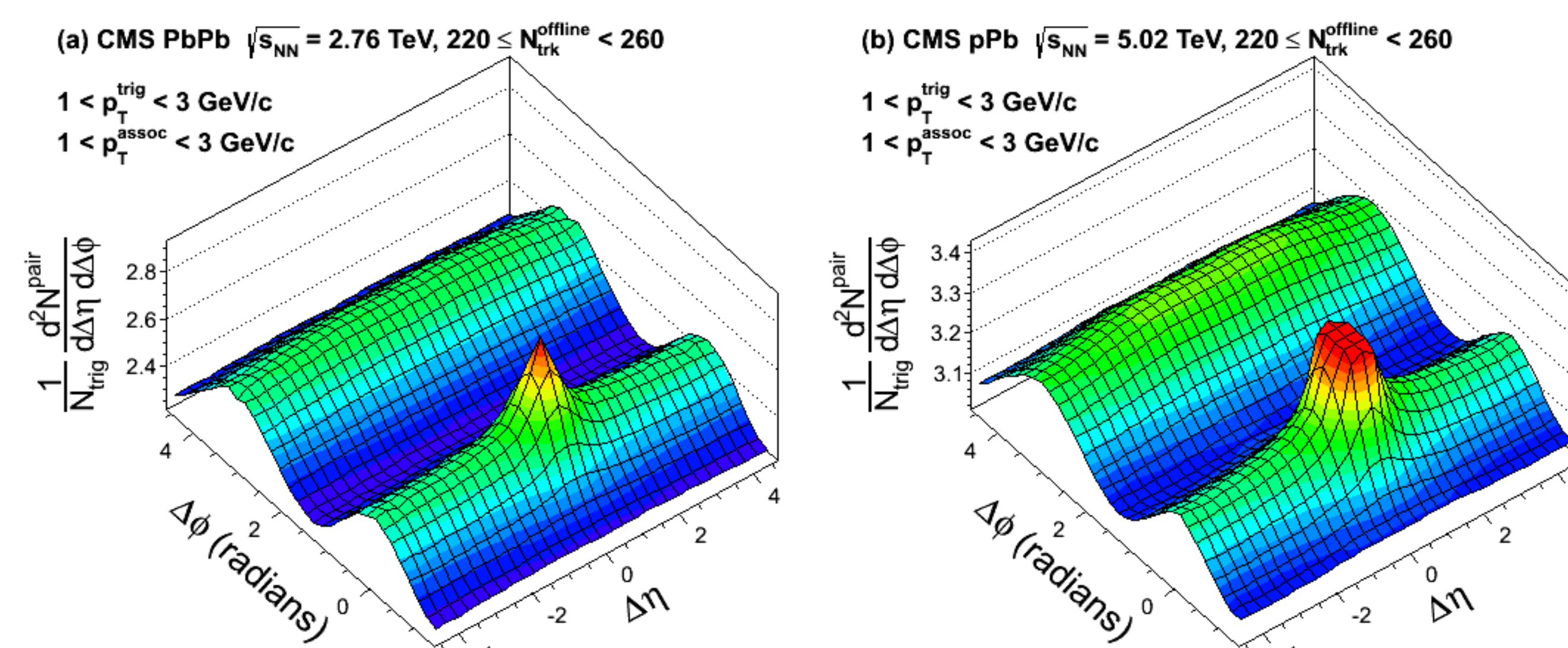
Jednadžba stanja QCD tvari



Pomoću konvolucijske neuralne mreže [2] pokušala se prepoznati jednadžba stanja kromodinamičke tvari iz azimutalnih kutova i transverzalnih impulsa nastalih čestica u sudarima. Učenje je izvršeno na simuliranim podacima temeljenim na dva hidrodinamička modela koji opisuju ponašanje kromodinamičke tvari. Na slikama su prikazane mape osjetljivosti koje opisuju koliko je predviđena jednadžba stanja osjetljiva na određeni dio (p_T, ϕ) faznog prostora.

Plan istraživanja

Jedan od ekperimentalnih potpisa QGP je pojava grebenaste strukture u dvočestičnoj korelacijskoj funkciji.



Primjer dvočestične korelacijske funkcije za PbPb i pPb sudare

Navedeni primjeri u kojima je pokazano da tehnike dubokog učenja omogućuju za red veličine bolju efikasnost u klasifikaciji mlazova te omogućuju ekstrakciju fizikalnih parametara koristeći samo distribucije impulsa i azimutalnog kuta nastalih čestica daju motivaciju za analizu u kojoj bi se pomoću dubokih neuralnih mreža iz low-level parametara (transverzalni impuls, azimutalni kut, pseudorapiditet i okus nastalih čestica) klasificirali mlazovi unutar pojedinih događaja. Potom bi se mogli konstruirati rezovi na događaje koji sadrže određenu klasu mlazova te promatrati utjecaj takvih rezova na dvočestičnu korelacijsku funkciju.

Koristeći grafičku karticu NVIDIA Quadro p6000 dobivenu iz NVIDIA potpore akademskim istraživanjima i radnu stanicu nabavljenu sredstvima iz HRZZ projekta IP-2018-01-4108 imamo sve potrebne hardverske alate za provedbu takve analize.



Grafička kartica NVIDIA Quadro p6000