Particle identification using a convolutional neural network on EMC cluster data

Marvin Peter

II. Physikalisches Institut, Justus-Liebig-Universität Giessen

Outline

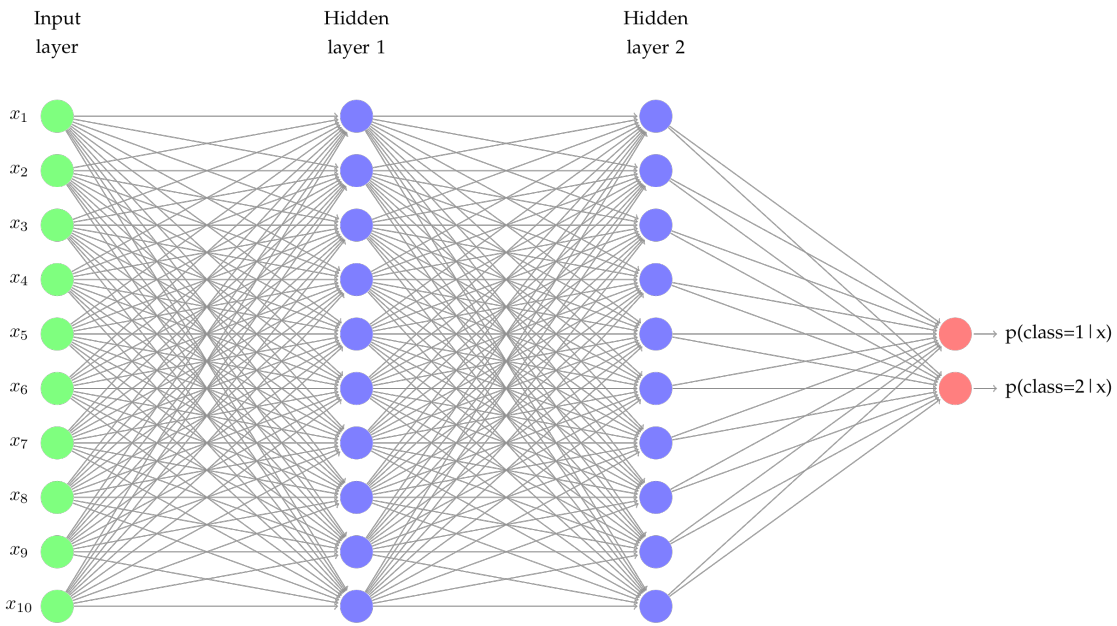
JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Basic Idea

- Use EMC cluster data
- Create images of the clusters
  (each crystal in the EMC is a pixel)
- Train a convolutional neural network on this image data
- The neural network should classify the images

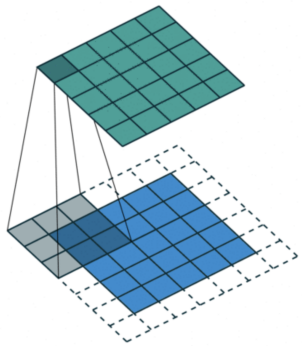
Multilayer Perceptron (MLP)

- Layers of fully connected nodes
- Supervised machine learning



Convolutional Neural Network (CNN)

- Convolution layers extract high level features out of the image
- MLP classifies the data
- Simple CNN
  - Two convolution layers
  - Two fully connected layers



Box Generator

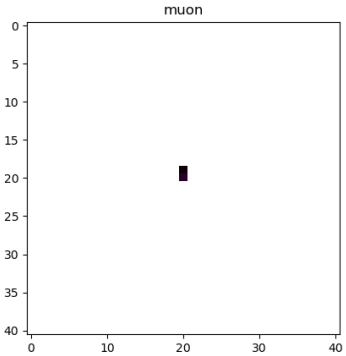
- Simulated data with PANDAroot
- Generation of 250 000 muons and 250 000 pi^+ with the box generator
- In the direction ϕ=[170 °,190 °], θ=[70 °,100 °]
- In momentum ranges
  p_0=[0 GeV,2 GeV], p_1=[2 GeV,4 GeV], p_2=[4 GeV,6 GeV]
- Equal amount of events simulated for muons and pions at each momentum range
- Results in approx. 470 000 usable cluster images

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

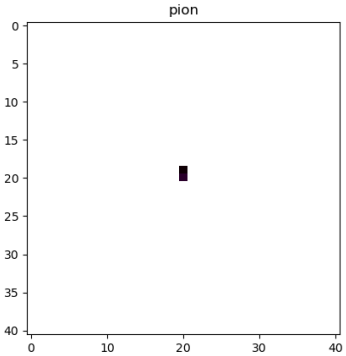
Image Data

- Write cluster data to a .csv file and create images with python
- Clusters with faulty data will be dropped
- Images are cut to 41 x 41 pixels (adjustable) and larger clusters will be dropped
- Each cluster is centered inside an image

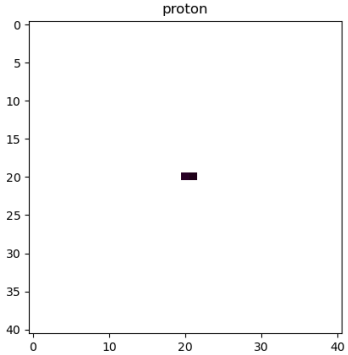
Image Data



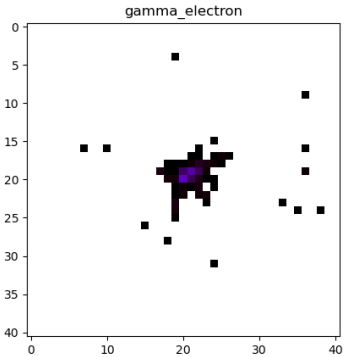
(a) μ^+/μ^-



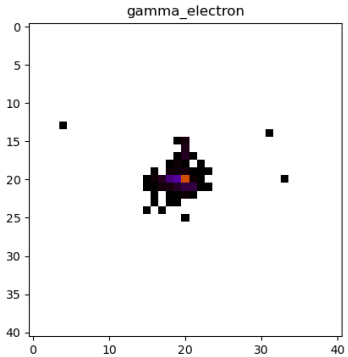
(b) $\mathrm{pi}^+/\mathrm{pi}^-$



(c) $\mathrm{p}/\bar{\mathrm{p}}$



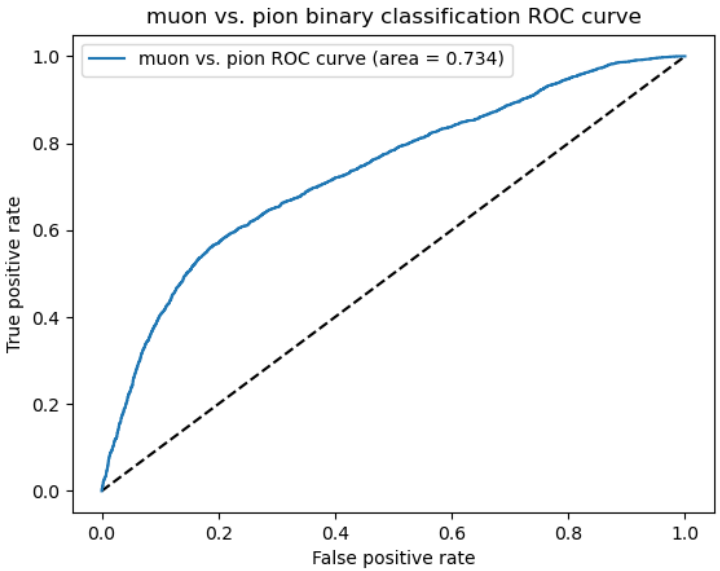
(d) γ or $\mathrm{e}^+/\mathrm{e}^-$

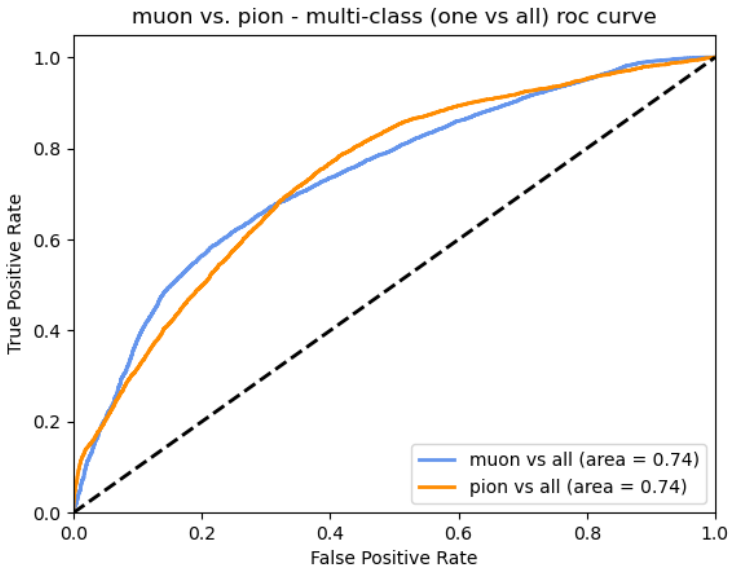


(e) γ or $\mathrm{e}^+/\mathrm{e}^-$

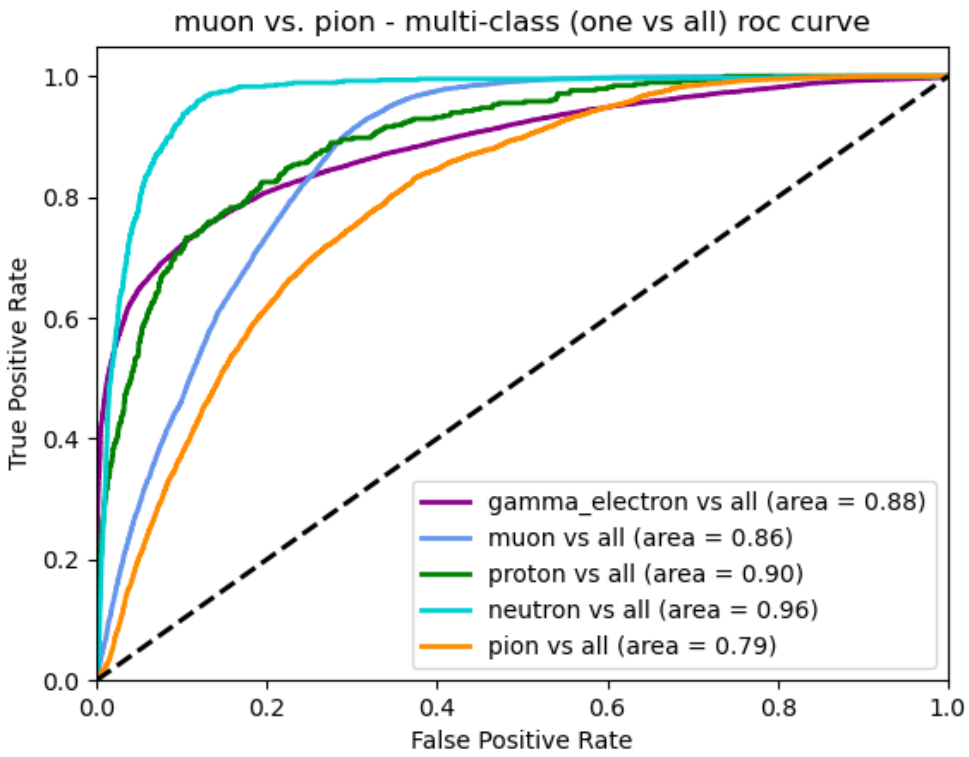
Each image: 41×41 crystal matrix out of the barrel. White space means 0

First Results





First Results

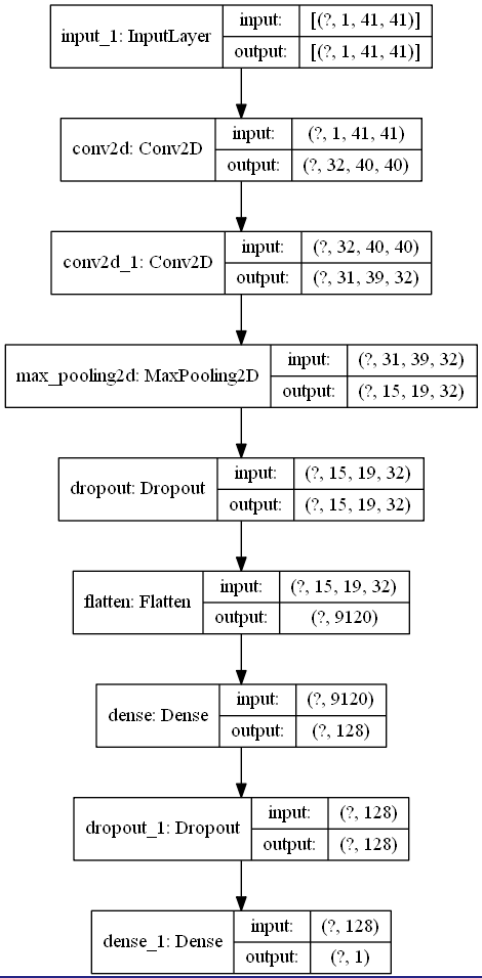


Outlook

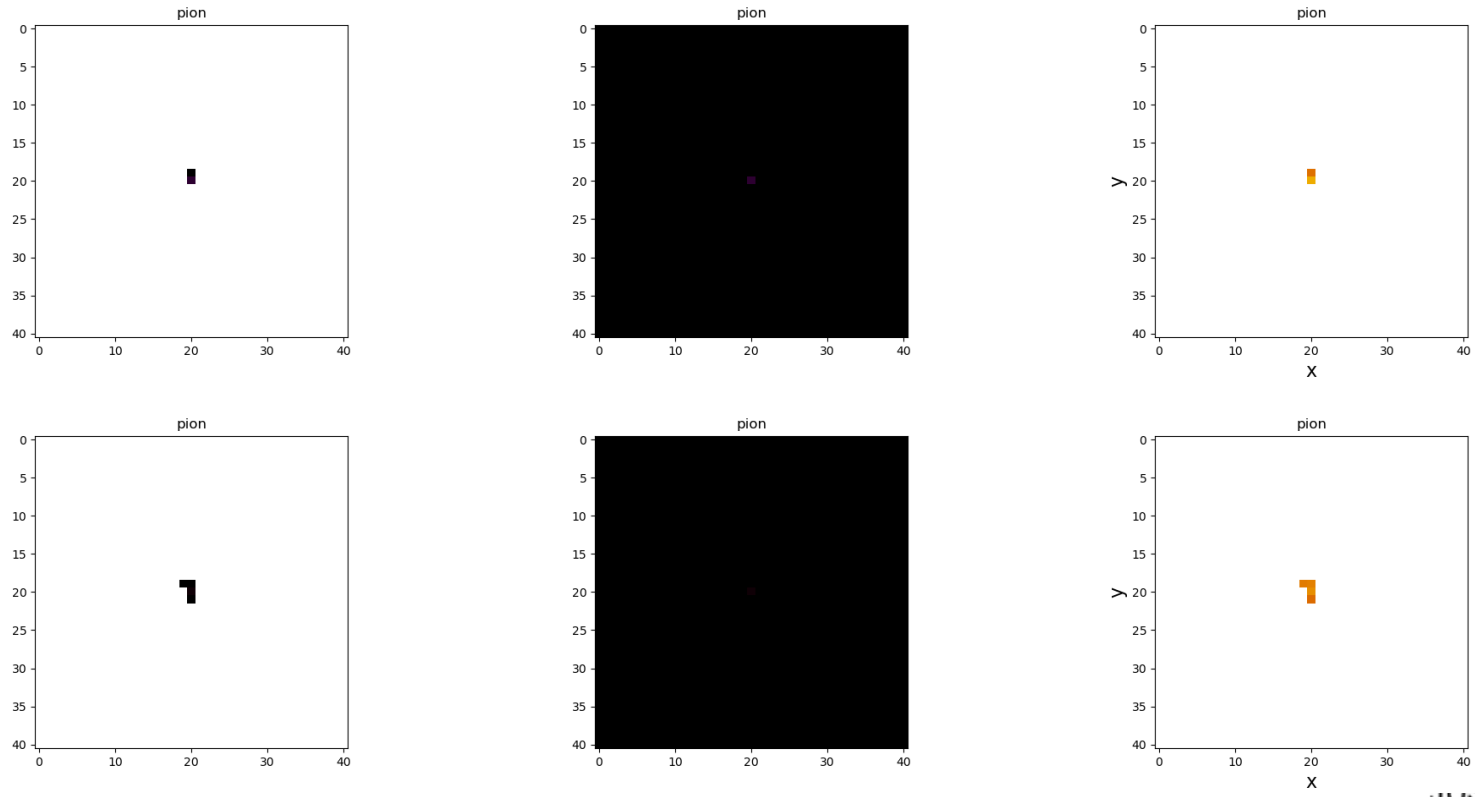
- Look at more data → simulate more types of particles
- Try different neural network architecture
- Change output categories
- Compare to other particle identification algorithms

Neural Network Structure

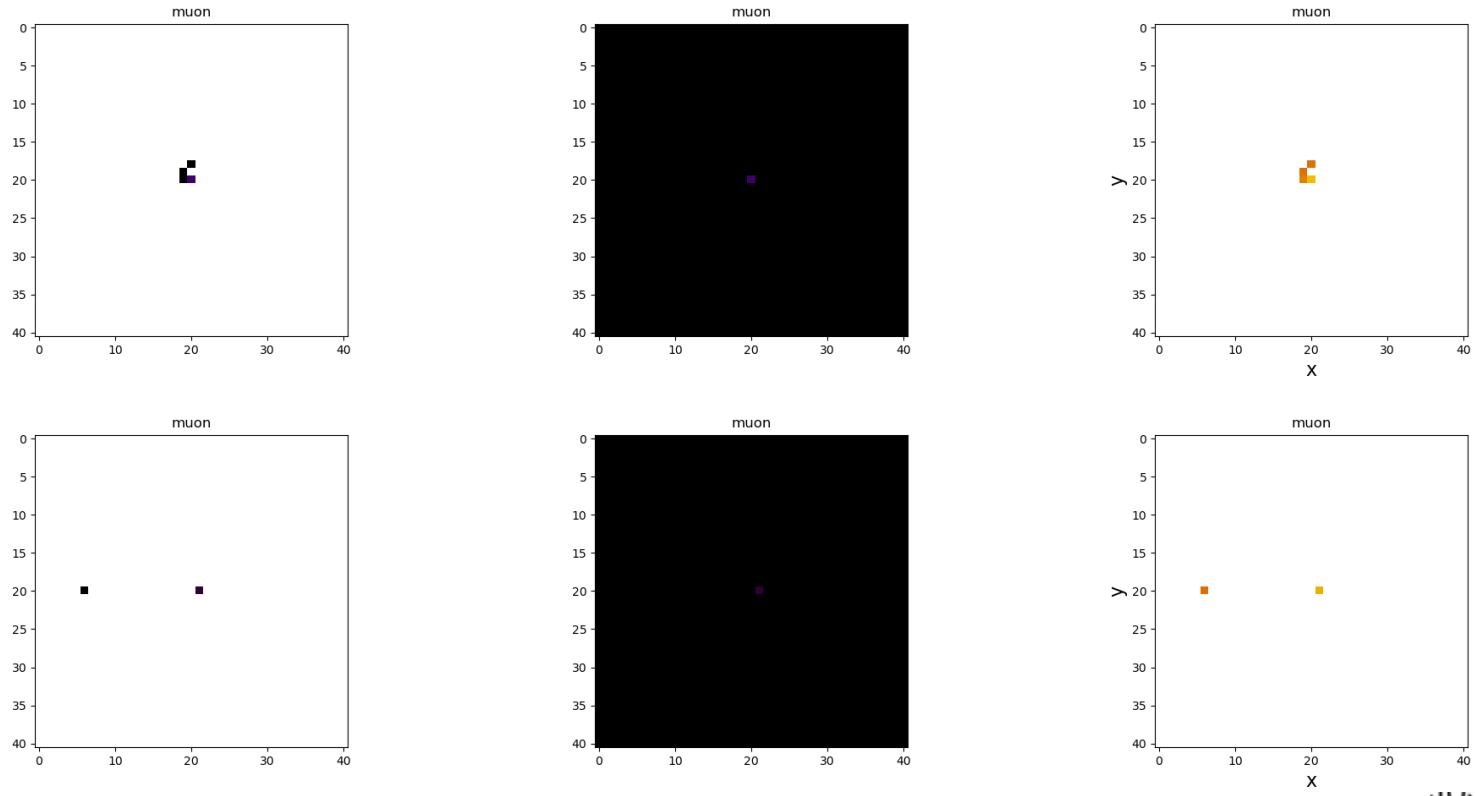
back



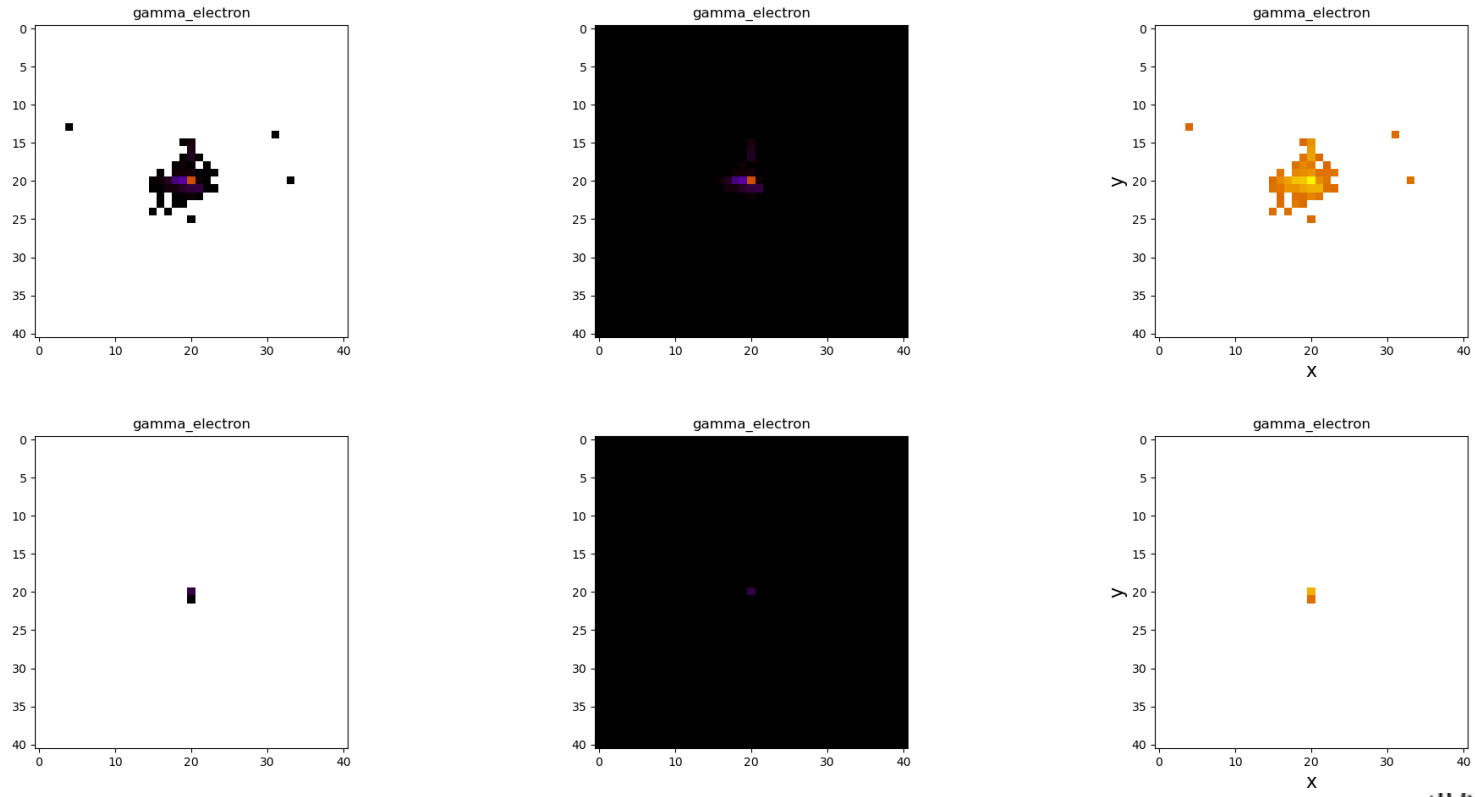
More Image Data - pi^+/pi^-



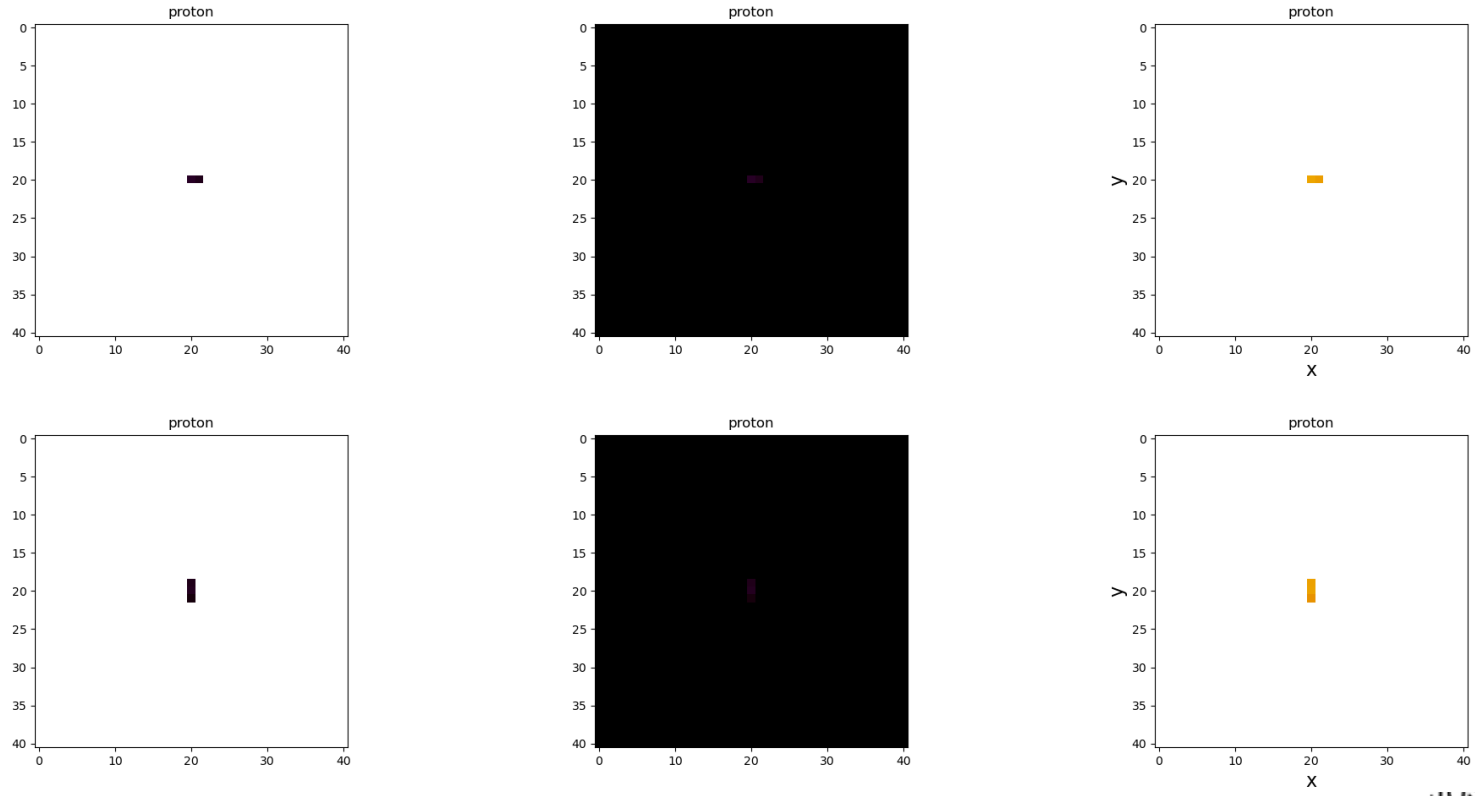
More Image Data - μ^+/μ^-



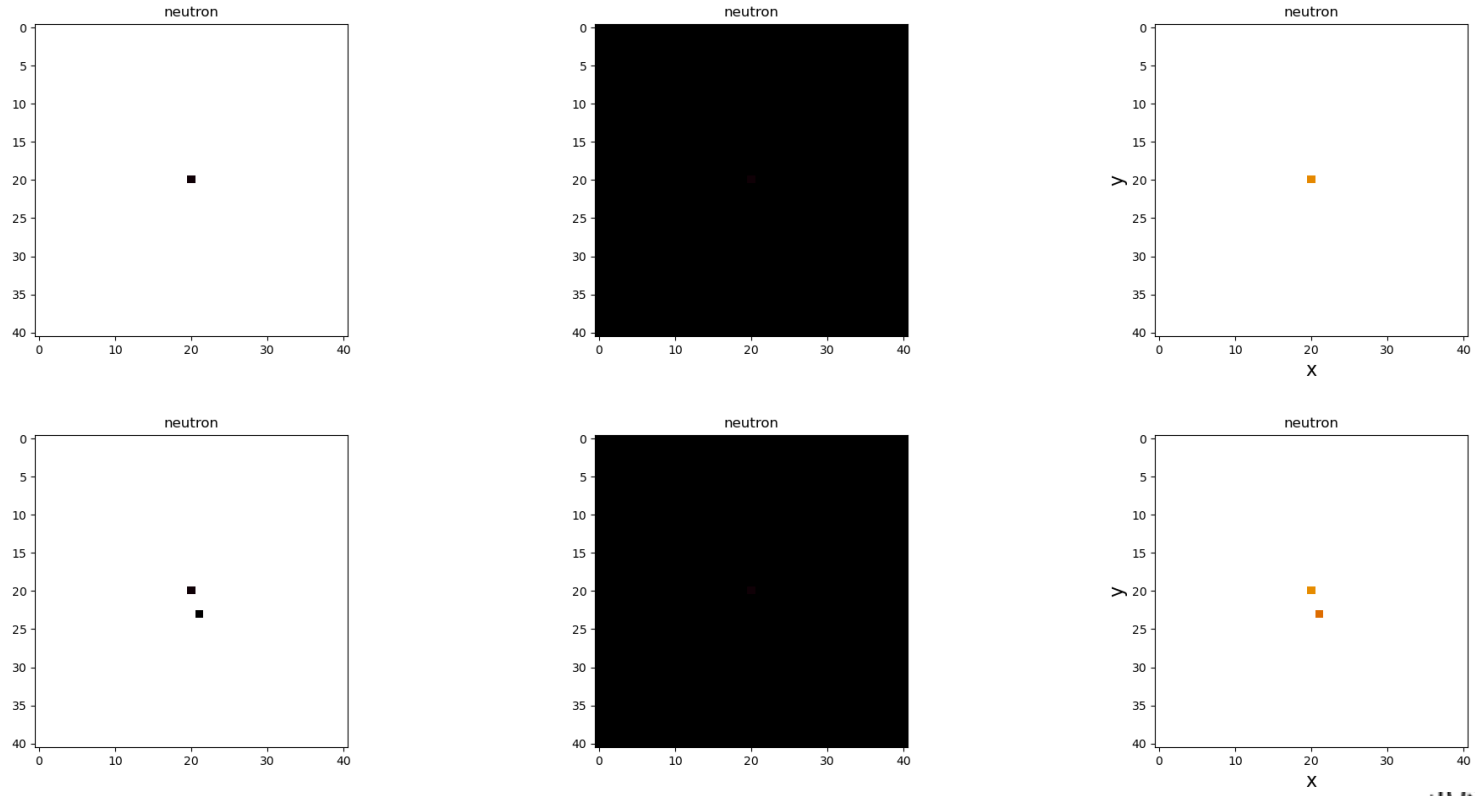
More Image Data - $\gamma/e^+/e^-$



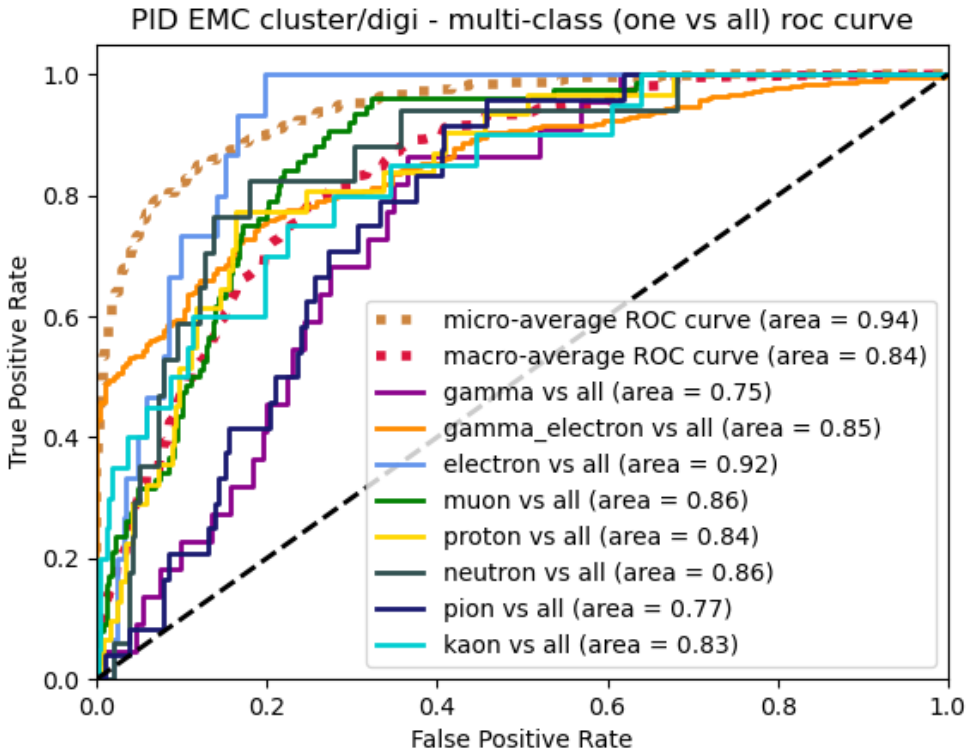
More Image Data - p/$\bar{p}$



More Image Data - n



Other Dataset Results



Other dataset! $p\bar{p} \rightarrow \phi\phi$

Information about each simulated file:

pandaroot_box_generator_muon_pion

filename	file	gamma	gamma_electron	electron	muon	proton	neutron	pion	kaon	background	n_final	n_start	n_dropped	primary	momentum
00.csv	0	12	1949	32	7432	0	0	0	0	0	9425	9501	76	muon	0-2
01.csv	1	18	1975	30	7360	0	0	0	0	0	9383	9456	73	muon	0-2
02.csv	2	10	1992	28	7372	0	0	0	0	0	9402	9475	73	muon	0-2
03.csv	3	8	1974	29	7383	0	1	0	0	0	9395	9473	78	muon	0-2
04.csv	4	7	1988	29	7348	0	1	0	0	0	9373	9450	77	muon	0-2
05.csv	5	5	2471	48	7508	0	0	0	0	0	10032	10074	42	muon	2-4
06.csv	6	8	2517	54	7458	0	0	0	0	0	10037	10079	42	muon	2-4
07.csv	7	9	2536	60	7450	0	0	0	0	0	10055	10102	47	muon	2-4
08.csv	8	5	2545	44	7423	0	0	0	0	0	10017	10068	51	muon	2-4
09.csv	9	6	2497	45	7489	0	0	0	0	0	10037	10076	39	muon	2-4
10.csv	10	10	2734	64	7265	0	0	0	0	0	10073	10116	43	muon	4-6
11.csv	11	6	2731	64	7252	0	0	2	0	0	10055	10107	52	muon	4-6
12.csv	12	8	2774	68	7188	0	0	0	0	0	10038	10102	64	muon	4-6
13.csv	13	2	2769	78	7228	0	0	1	0	0	10078	10121	43	muon	4-6
14.csv	14	7	2789	77	7186	0	0	0	0	0	10059	10128	69	muon	4-6
15.csv	15	9	2796	68	7196	0	0	0	0	0	10069	10115	46	muon	6-8
16.csv	16	10	2898	88	7081	0	0	0	0	0	10077	10135	58	muon	6-8
17.csv	17	5	2853	68	7134	0	0	0	0	0	10060	10119	59	muon	6-8
18.csv	18	7	2899	73	7105	1	0	0	0	0	10085	10127	42	muon	6-8
19.csv	19	11	2815	88	7180	0	0	0	0	0	10094	10143	49	muon	6-8
20.csv	20	343	6265	87	109	282	223	2834	1	0	10144	12089	1945	pi+	0-2
21.csv	21	320	6195	85	116	276	200	2796	1	0	9989	11942	1953	pi+	0-2
22.csv	22	319	6351	85	140	293	212	2837	1	0	10238	12191	1953	pi+	0-2
23.csv	23	303	6392	77	124	240	226	2758	1	0	10121	12094	1973	pi+	0-2
24.csv	24	303	6292	80	113	260	216	2762	1	0	10027	12047	2020	pi+	0-2
25.csv	25	307	9160	101	50	210	326	2147	7	0	12308	16080	3772	pi+	2-4
26.csv	26	386	9204	135	51	227	311	2151	1	0	12466	16303	3837	pi+	2-4
27.csv	27	408	9214	136	40	205	320	2158	2	0	12483	16335	3852	pi+	2-4
28.csv	28	407	9189	144	52	249	301	2164	4	0	12510	16344	3834	pi+	2-4
29.csv	29	376	9110	121	53	210	285	2158	2	0	12315	16081	3766	pi+	2-4
30.csv	30	363	10021	207	41	156	349	2000	2	0	13139	17822	4683	pi+	4-6
31.csv	31	358	10143	198	42	166	301	2006	3	0	13217	17864	4647	pi+	4-6
32.csv	32	337	10132	185	34	171	322	2014	3	0	13198	17824	4626	pi+	4-6
33.csv	33	315	10224	200	51	177	337	2055	3	0	13362	18014	4652	pi+	4-6
34.csv	34	362	10291	192	36	179	326	1987	6	0	13379	18000	4621	pi+	4-6
35.csv	35	366	11042	219	26	138	350	1889	2	0	14032	19419	5387	pi+	6-8
36.csv	36	318	11100	246	25	158	299	1896	1	0	14043	19372	5329	pi+	6-8
37.csv	37	320	11028	243	30	150	323	1882	3	0	13979	19398	5419	pi+	6-8
38.csv	38	372	11150	232	32	171	359	1889	2	0	14207	19600	5393	pi+	6-8
39.csv	39	359	11265	246	35	137	344	1832	2	0	14220	19626	5406	pi+	6-8
all		7105	234270	4354	147238	4056	5932	44218	48	0	447221	527412	80191		

Information:
generated particles in the angle ranges: (phi=[170°,190°], tht=[70°,100°])
background is always zero (would contain stuff if there were other particles generated not otherwise specified)
rows C-K are numbers of clusters associated with particles from monte carlo truth information