ML Status and Plans

Coordinating overall ML for ZTF Help (and problems) always wanted Thursday 2 PM meetings

Ashish Mahabal <<u>aam@astro.caltech.edu</u>> ZTF ML Lead Lead Computational Scientist, CD3, Caltech ZTF Team Meeting, UW, 20190905

rb -> drb ('braai')

rf -> CNN



Duev et al. Next talk

Deep Learning with AStreaks



DeepStreaks

CNNs



Figure 2. Decision logic used by DeepStreaks to identify plausible streaks. The problem is split into three simpler sub-problems, each solved by a dedicated group of classifiers assigning real vs. bogus ("rb"), short vs. long ("sl"), and keep vs. ditch ("kd") scores. At least one member of each group must output a score that passes a pre-defined threshold. See Section 2.1 for details.

Duev et al. Next talk

Crowded field (shown by Frank)



Deconvolution



SURF project Shubhranshu Singh (IIT-G) Mentor: A Mahabal Co-mentor: D Duev

If the PSF is known:

- Wiener Filtering
 - Based on Fourier transform of the signal, the PSF, and the noise
- CLEAN
 - Approximate point sources by delta-functions
- Maximum Entropy
 - Fits the data with maximum entropy
- Richardson-Lucy
 - Iterative method Bayesian methodology

If the PSF is unknown - Deep Learning



Image pairs of convolved and clean images required

Clean images: HST's WFC3 (normalised to [0, 1]) Convolved images - use ZTF PSF

HST data (For training) 2K by 4K FITS images UV filter with 32 bits per pixel ZTF data (For testing) 3K by 3K FITS images g/r filter with 32 bits per pixel During training - cropped images used

Simulated data - Using PyRAF mkobjects and starlist commands

1000, 256x256 images with 5-15 gaussian sources Background with poisson noise



Input images - 64x64, 256x256 and 512x512 for different experiments. 500 epochs, Adam optimizer, learning rate of 10⁻⁴ (for 256x256 images)

	Training set	Validation set	Test set
HST data	1450 images	162 images	180 images
Simulated data	810 images	90 images	100 images

Typically used - L2-norm and L1-norm

Modified loss function - mixture of L1-norm and multi-scale structure similarity (MS-SSIM) MS-SSIM - Calculated using mean and variance of the data, at various scales.

Zhao, Hang, et al. "Loss functions for image restoration with neural networks." Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment."

$$\mathcal{L}^{MS-SSIM}(I) = 1 - \text{MS-SSIM}(I) \qquad \qquad \mathcal{L}^{Mix} = 0.8 \cdot \mathcal{L}^{MS-SSIM} + 0.2 \cdot \mathcal{L}^{\ell_1}$$

	Convolved Image	Original model	Added layer	Deeper Model
Average PSNR	43.99 dB	48.5 dB	51.74 dB	47.83 dB
Max. PSNR	46.80 dB	51.23 dB	55.03 dB	51.84 dB
Min. PSNR	40.11 dB	37.87 dB	40.06 dB	33.6 dB

Results on Simulated Data

PSNR - Peak Signal-to-Noise Ratio



Shubhranshu Singh

HST data + ZTF PSF

deconvolved

Autoencoder results





ZTF image

Generative Adversarial Network (GAN)



ISOIA et al. CVPK 2017)

deblurGAN

ResNET



He, Kaiming, et al. "Deep residual learning for image recognition."



Total parameters - 261,761

deconvolved





Remove artifacts near bright sources Use a better PSF model Improve the ResNet architecture Make a pipeline for processing 3k x 3k ZTF images

RNN - Vinu (ZTF)



SURF project Vinu Sankar (IIT-G) Mentor: A Mahabal Co-mentor: M Graham



Image courtesy: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Wavenet

Dilation Long range But needs to be deep. Seq2seq



Figure 3: Visualization of a stack of *dilated* causal convolutional layers.

Error associated in mag. Irregular temporal gaps, sparse data. Padding inputs for RNNs. RNNs tend to forget long lcs.

Tackling challenges



dt as input to RNN. Use GRUs or LSTMs^[1] instead of vanilla RNN cell. Drop points with dt > 4 months. Stitching instead of padding.



Static RNN models Zero-padding Stitching GRUCell, FastGRNNCell, various hidden units 16, 32, 64, 128

Dynamic RNN models

Bucketing in batches

Masking

Various inputs

[dt, mag], [dt, dm], [t-t₀, dt, dm, mag, magerr], [dt, mag, dm/dt, t-t₀]... Inputs tried with/without normalizing Drop points with dt > 4 months

Best model:

Static RNN with stitching, drop dt>120, input [dt, mag] (normalized)

Vinu Sankar

RNNs and delta-ts

using for multiple filters



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

X : Inpu	ut tim	e seri	ies (2	variał	oles);		M: Mask						
s : Tim	estan	nps fo	or X ;				Δ : Time	inte	rval f	or X .			
$\boldsymbol{X} = \begin{bmatrix} 47\\ NA \end{bmatrix}$	49 15	<i>NA</i> 14	40 NA	NA NA	43 NA	55 15]	$M = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$	1 1	0 1	1 0	0 0	1 0	$\begin{bmatrix} 1\\1 \end{bmatrix}$
<i>s</i> = [0	0.1	0.6	1.6	2.2	2.5	3.1]	$\mathbf{\Delta} = \begin{bmatrix} 0.0 \\ 0.0 \end{bmatrix}$	0.1 0.1	0.5 0.5	1.5 1.0	0.6 1.6	0.9 1.9	0.6 2.5

Figure 2. An example of measurement vectors x_t , time stamps s_t , masking m_t , and time interval δ_t .

Che et al. 2018 Nature DOI:10.1038/s41598-018-24271-9

Naul et al. (encoder-decoder)

Best model: Static RNN with stitching, drop dt>120, input [dt, mag] (normalized) 96% for easy classes

Should certain delta-ts be ignored?



Using transformers^[1] Combining CNN + RNN Train on ZTF + CRTS Test on ZTF

Vinu Sankar

Plans

- Classification: Variable sources [+architecture] (JKM, UW)
- Visualization to understand/improve transient classification
- BTS (Adam Miller)
- Deep Coadds (Danny Goldstein)
- Zooniverse with light curves (Richard Walters)
- RNN (visiting postdocs)
- Transfer learning (UNC)
- Asteroid light curves? (Rex)
- Babamul [broker]
- Gaia/ALERCE synergy

Help always needed Thursday 2 PM meetings Coordinating overall ML for ZTF

aam@astro.caltech.edu

10 fields (Jan)

Total number of observations (R+g) on 2019-06-08

																			f879														f878	
80		f87 28	7			f876 29				f875 5			f8	874 6				f873 6	0			f872 6			f871 0	1			f870 0			f8 1	23 69 18	
80	f{	92	_	f867 182		f866 182		f8 5	65 94	1	864 549	f	863 356		f862 376	2		f861 426		f860 377	0 7	f	359 360	f8 4	358 104	1	f857 443		f856 463		f855 189		f854 176	_
	1853		f852 194	ft	851 195	f850 181		f849 781	f84 86	18 33	f847 803	f84 66	6 3	f845 573		f844 626		f843 605	fe 4	42 00	f84 38	1 5	f840 332	f839 379	9	f838 403	f83 42	37 25	f836 253	f8 2	35 03	f834 189	f83 40	3
60	f832 410	f83 42	1 ft 9 4	330 432	f829 187	f828 809	f8 81	27 77	826 1013	f825 1128	f824 1078	f823 943	f8 8	322 369	f821 748	f82 70	20)3	f819 864	f818 481	f8 4	17 67	f816 398	f815 360	f814 399	f813 392	8 f8: 2 2:	12 f 18	f811 204	f810 256	f809 378	f80 39	8 <u>f8</u> 7 9	07 f8 96 10	06)77
	f805 645	f804 662	f803 821	f802 461	f801 495	f800 414	f799 632	f798 1140	f797) 1236	f796 125	5 f795 2 1220	f794 1215	f793 1073	f792 3 986	2 f79 5 81	1 f7 2 7	90 47	f789 932	f788 491	f787 693	f786 680	f785 495	f784 377	f783 374	f782 414	f781 406	f780 256	f779 466	f778 466	£777 404	f776 413	f775 f 438	f774 f7 197 20	73 08
	f772	f771 541	f770 777	f769 1292	f768 2 736	f767 484	f766 1325	f765 1401	764 f7 1245 1	763 f7 269 1	762 f761 260 1194	f760 4 1191	f759 1093) f758 3 1163	f757 8 870	f756 823	f75 51	5 f75 5 54	4 f753 7 458	f752 634	f751 596	f750 537	f749 f7 331 4	48 f74 15 42	7 f746 8 416	1745 276	f744 300	f743 501	1742 f 474	741 f7 421 2	40 f7. 18 24	39 f73 14 229	8 f737 9 376	f736 679
40	f735 f 384	734 1 366 2	33 f73 05 18	2 f731 8 347	l f730 7 504	f729 f7 698 4	28 f7 37 44	27 f72 47 142	6 f725 9 1245	f724 f 1246]	723 f722 252 121	f721 71121	f720 f 1057	f719 f7 971 8	18 7 845 8	17 f7 15 8	16 f7 95 5	15 f7 16 5	14 f713 24 448	f712 425	f711 356	f710 f7 307 3	709 f708 333 347	523 f707	f706 f7 372 2	05 f704 55 291	4 f703 7 505	f702 706	f701 f7 273 2	00 1699 23 222	698 338	f697 f69 224 2	96 f695 18 727	f694 361
ec	f693 214	f692f6 656 6	591 <mark>f69</mark> 577 <u>33</u>	0 f689 8 600) f688 f) 488 4	687 f68 403 63	6 f685 8 426	f684f	583168 38014	2 f681 91169	f680f679 1150182) f678 f 2 978	677 fe 880 1	676 f67 016 33	5 f674 8 335	f673 499	f672 505	f671 f 523	670 f66 426 42	9 f668 7 391	f667 335	f666 f6 479 4	65 f664 12 321	f663 f66 554 36	52 f661 53 819	f660 f6 286 5	59 (658 06 658	8 f657 8 248	f656 f65 273 23	55 f654 32 222	f653 f6 327 5	52 f651 99 394	f650 f64 213 37	9 f648 2 395
	f647 177	f646 f6 200 2	45 f644 00 213	f643 691	642 f64 173 43	41 f640 86 651	f639 f6 644 4	538 fd3 134 37	7 f636 f 8 353 1	535f63 06910	4f633 f6: 50 954 18	32 f631 88 233	f630 f 208	f629 f6 214 2	28 f627 57 271	f626 624	f625 464	f624 f6 645 3	523 f622 380 328	f621 f 291	f620 f6 300 4	19 f618 10 308	3 f617 f61 3 245 39	l6 f615 95 287	f614 f61 296 29	.3 f612 9 484	f611 f6 516 2	10 f60 99 229	9 f608 f6 9 274 2	07 f606 89 212	f605 f6 186 2	04 f603 98 194	f602 f60 202 20	1 f600 1 219
20	- f599 f 326	598 f59 175 20	7 f596 2 203	f595 f5 350 2	594 1593 210 192	f592f5 410 6	91 (590 26 \416	0 f589 f 5 211 (388 f58 342 88	7f586 f 3 336	585 f584 f 199 213	583 f58 249 20	32 f581)8 217	1 f580 f 7 216 2	579 f57 242 40	78 f577 08 384	7f576 4 399	f575 f 280 2	574f573 237 280	f572 f 233 2	571f57 255 38	0f569f 4 266	568 f567 211 370	f566f56 278 24	5f564f 4 247	563 f56: 422 50:	2 f561 f 2 322	560 f5 281 3	59 f558 f 86 409	557 f55 202 19	5 f555 f 7 238	554 f553 271 293	3 f552 f55 3 167 27	1 f550 8 190
	f549f5 142 1	48 f547 50 176	7 f546 f5 5 185 1	645 f54 .92 21	4 f543 f 6 230	542 f541 1 72 363	f540 f 555 (539 f53 546 42	8 f\$37 f5 8 304 4	36 f53 42 34	5f534f533 4 183 203	3 f532 f5 L 217 1	31f53 81 23	30f529f 30 237	528 f5/ 304 24	27 f52 48 25	6 f525 5 368	5 f524 f 3 218	f523 f52 310 28	2 f521 6 243	f520 f5 298 4	519 f518 20 344	3 f517 f51 4 295 39	16 f51/5 f 92 207	514 f513 371 82	3 <i>f</i> 512f5 0 554 2	511 f510 215 1/77	0f509f 7263	508f507 224 271	f506f50 230 16	5f504f 6 218	503 f502 242 251	2 f501 f50 L 251 24	0f499 B 282
1	f498f49 272 248	7 f496 f4 3 251 1	495 f494 158 162	4f493f 2 174	f492 f49 172 16	1f490 f4 5 192 2	489 f48 296 43	14871 313	86 f485 59 321	f484f4 249 2	83 f482 f4 15 173 1	81f480 35 199	f479f4 164 1	478f47 173 20	7f476f 5 179	475f4 184 1	74f47 77 31	3f472 0 164	f471f47 160 16	0f469f 3 149	f468f4(251 32	67f466f 20 264	f465 f464 174 157	f463f46 420 44	2 f461 f4 1 518 3	460f459 316 175) f458 f4 5 268 2	157 f45 20 17	6 f455 f4 7 221 1	54f453f 92 172	452 f45 266 25	1f450f4 5 254 2	449 f448 f 33 219	447 145
0	f446f4 119 1	45 f44 16 11	4f443f4 1 185 1	442f44 L59 15	1f440f 59 147	439 f48 197 16	8 f437 f L 364	436 f43 452 38	5 f434 f 0 286	433 f43 154 14	2 f431 f43 3 136 10	0f429f 9 112	428f4 132 3	27 f426 31 161	f425 f4 134 1	424 f4: 120 1:	23f42 29 13	2f421 2 129	f420f4 143 1	19f418 86 141	f417f4	416f41 351 16	5 f414 f41 5 205 1	3 f412 f 57 675 1	411 (410 .045 280) f409 f4) 125 2	08f407 29 165	7 f406 f 5 323 3	405 f404 252 208	f403f40 540 26	2f401f 1 268	400 f399 258 222	9 f398 f39 2 203 27	7f396 0171
f	f395f394 171 89	1f393f3 85	392 f39 147 13	1f390 6 132	f389f38 162 17	8 f387 f. 9 132 1	386 f38	85 f384 50 540	1383 f38 296 24	2 f381 f 8 99	380 f379 f 92 99	378f37 96 9	7f376 9 108	5f375f3 3 112 1	74 f37. 13 95	90 90	03	370f3 90 9	69 f368 97 112	f367f3 371 1	66f365 45 133	5 f 364 f 3 3 157 1	363f362f	106 57	245 245	58f357	f356f3 128 1	55 f35 54 170	4 f353 f3 B 262 2	52f351f 08 339	350 f34 249 26	9f348f3 8 225 2	847 f346 f 225 209 1	345 190
-20	f344 f. 151 1	343f34 123 10	2f341f 0 103	340 f33 102 12	39f338 29 124	115 96	6 f335 f 5 85	f334f33 487 43	13 1332 1 73 207	331 f3 574 3	0 f329 f32 2 68 7	28f327f 9 67	f326f3 78 f	825 f324 89 94	f323 f. 76	322 f3: 76 7	21f32 9 7	0f319 2 77	f318f31 82 11	7f316 9 115	f315f3 135 1	814f313 .42 105	3 f312 f31 5 /79 34	1 f310 f3 7 609 5	309 f308 562 439	f307f3	06 f305 66 159	f304f3 156 1	303f302 161 167	f301f30 164 18	0f299f 0 174	298 f297 165 17	1296 129 199 15	5f294 2 159
	f293 f 162	292 f29 109 11	91 f290 14 101	f289 f. 86	288 f28 85 80	7 f286 f) 74	285 f28 63 30	84 f283 06 424	f282 f2 399 1	81 f280 82 299	f279 f278 177 36	44 f277 f2	276 f2 47 4	75 f274 16 39	f273 f2 55	272 f2 47 6	71 f27 57 5	70 f269 0 69	f268 f2 76	67 f266 75 76	5 f265 f 74	f264 f26 73 4	53 f262 f2 5 385 4	261 (260 405 1000	f259 f2 5 529 5	58 f257 60 689	f256 f2 127 1	255 f25 L17 13	4 f253 f 3 129	252 f251 117 135	f250 f2 94	249 f248 79 122	f247 f24 136 14	6 f245 5 159
3	360		33	D		300			270		2	40		2	10			180 RA			150		1	20	í	90			60			30		0

features/variability -> classification

[Hackday possibility]

Timescale-Luminosity plot (Dan, Anna, ...)



advanced dynamic plotting -> filtering/discovering ...

[Hackday possibility]

Planned Zooniverse Campaign



Exactly 1 source within 30 arcsec

ZTF Brokering architecture



Data:

125M alerts, 2.5B light curves Brokers:

Alerce, ANTARES, Lasair, MARS, ... GROWTH, AMPEL, kowalski Followup:

SEDM, BTS TNS



Visualization for interpretability

- A. Activation Maximization
 - Initial layer filters easy to visualize
 - Generate input image that activates later filters
- B. Saliency Maps
 - Gradient of o/p category wrt input image
 - Understanding attention of the classifier
- C. Class Activation Maps
 - Gradients based on first dense layer
 - Spatial information still intact





Meet Gandhi