Multimodal multi-output ordinal regression for discovering gravitationally-lensed transients

Nicolò Oreste Pinciroli Vago
Department of Electronics, Information and Bioengineering
Politecnico di Milano
Milan, Italy
nicolooreste.pinciroli@polimi.it

Piero Fraternali
Department of Electronics, Information and Bioengineering
Politecnico di Milano
Milan, Italy
piero.fraternali@polimi.it

Abstract

Gravitational lenses are caused by massive bodies that distort space-time, bending light. They can distort transients, such as Supernovae (SN), which are being studied extensively. Gravitationally-lensed supernovae (LSN) are rare, so only a few have been discovered. Future astronomical surveys will collect huge amounts of data, calling for automated and accurate discovery techniques to find them. Still, only a few works aim to discover LSN, most use only a few classes to characterize candidate observations, and only a few exploit spatial and temporal information. This work introduces AstroCountNet (ACoNet), an ensemble of multimodal neural networks that takes in input spatio-temporal data and, for each observation, counts the occurrences of 7 astronomical bodies. ACoNet achieves, on average, more than 85% macro F_1 score on four datasets. The network is then adapted into AstroClassNet (AClaNet) to address classification problems, achieving macro F_1 scores between $\approx 59\%$ and $\approx 93\%$.

1 Introduction

Gravitational lensing is an astrophysical phenomenon that consists of the bending of light due to massive bodies that distort space-time. It provides insights into the distribution and properties of dark matter [1] and allows the study of transient phenomena [2]. Transient phenomena are time-dependent and last from milliseconds to years [3–5]. Transient phenomena include, e.g., SN explosions and pulsating stars, extreme phenomena that can give origin to new physical theories or models [6]. The study of gravitationally-lensed transients is an emerging field with applications in astronomy and cosmology [7, 8]. Their rareness and the abundance of data requires automated discovery [9].

This work proposes a novel approach that addresses two tasks: counting and classifying central and off-central objects in multimodal data containing transient phenomena and gravitational lenses. We use the DeepGraviLens (DGL) datasets [10, 11], which contain spatio-temporal observations, each comprising multiple astronomical bodies (stars, galaxies, SN and lensed galaxies) observed in the *griz* electromagnetic field regions. For counting, the output is the number of bodies of each type in each observation. For classification, the output is the body types in each observation.

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

Several works look for gravitational lenses in unimodal data (e.g., images or tabular data) [12–17], but neglect the temporal component and do not account for lensed transients. The works in [18, 19, 10] address this issue with multimodal neural networks that take in input spatio-temporal data. Still, they are limited to a specific classification problem (i.e., classifying gravitational lenses and LSN) and are not designed for counting nor for classifying other objects. Moreover, they focus only on the central objects in images.

This paper presents AstroCountNet (ACoNet), the first architecture for counting and classifying astronomical bodies in multimodal data comprising gravitational lenses and transients. The contributions can be summarized as follows: (1) we introduce ACoNet, which takes in input spatio-temporal astrophysical observations and outputs the counts for each astronomical body in each observation, (2) demonstrate the effectiveness of soft-label and hard-label ordinal regression over classification for counting astronomical bodies in the DGL datasets [10, 11], and (3) introduce AstroClassNet (AClaNet), which uses task adaptation to classify transients and gravitational lenses. AClaNet obtains significant improvements compared to the classes extracted from ACoNet counts (up to $\approx$ 11% macro F_1 score). The code is made available at `https://anonymous.4open.science/r/AstroNets-F84D/`.

2 Datasets and method

2.1 Datasets

Table 1: The astronomical bodies of the four datasets and the possible count values for each of them

	Star	Galaxy	SNIa	SNCC	Lensed galaxy	LSNIa	LSNCC
Count values	0, 1, 2	0, 1, 2	0, 1	0, 1	0, 1	0, 1	0, 1

This work uses the four datasets introduced in DGL [10] and DeepZipper [18, 19] (DESI-DOT, LSST-wide, DES-deep and DES-wide), each simulating a cosmic survey with different characteristics. Each dataset contains $\approx$ 20,000 observations and is divided into a train set ($\approx$ 70%), a validation set ($\approx$ 15%) and a test set ($\approx$ 15%).

The inputs are astrophysical observations consisting of four 45 $\times$ 45-pixel images (one for each band of the *griz* photometric system) and four brightness time series (one for each band of the *griz* photometric system). The input samples can be labelled with multiple classes (i.e., astronomical bodies), each associated with a count (i.e., the number of instances of that astronomical body in the observation). The values of the count range from 0 to 2, as shown in Table 1. The datasets are imbalanced with respect to the astronomical bodies (e.g., $\approx$ 3% of each dataset contains a Supernova of type Ia (SNIa)) and the respective counts (e.g., observations with 2 stars are less frequent than the ones with 1 star).

2.2 Tasks, targets and outputs

This work addresses the tasks and search targets summarized in Table 2. Counting aims at determining the number of instances of all the types of bodies and classification aims at determining the types of bodies. The considered types of bodies are: *star*, *SNIa*, *Core-collapse supernova (SNCC)*, *galaxy* and their lensed counterparts *LSNIa*, *LSNCC*, and *Lensed galaxy*. The class *Other* denotes the observations containing none of the objects belonging to the classes considered in a task.

Binary classification aims to find SN without discerning their type (Ia and CC) or the presence of lenses. Multi-class classification aims at distinguishing lensed and non-lensed objects into different classes: SNIa, SNCC, Gravitationally-lensed SNIa (LSNIa), Gravitationally-lensed SNCC (LSNCC), Lensed galaxy (only for *SN and grav. lenses*) and Other. This work focuses on transient and/or lensed objects, so stars and galaxies are neglected in the multi-class classification task.

We first address the counting problem, which is more general than classification, because counting aims at finding both the body types and the number of bodies of each type, while classification considers only the body types. We compare the three alternative representations of the output proposed in [20]: Multi-Class Classification (MCC), Soft-label Ordinal Regression (SOR), and Multi-label (hard) Ordinal Regression (MOR).

Table 2: Summary of the tasks. In the "Type" column, "C" stands for "counting", "BC" stands for "binary classification", and "MC" stands for multi-class classification. The astronomical body types considered for each task are marked with •.

Task – Target		**Lensed**			**Non-lensed**				**Other**
Task	**Target**	**LSNIa**	**LSNCC**	**LGalaxy**	**SNIa**	**SNCC**	**Galaxy**	**Star**	
C	All objects	•	•	•	•	•	•	•	
BC	Gravitational lenses	•	•	•					•
BC	SN (coarse)	•	•		•	•			•
MC	SN (fine)	•	•		•	•			•
MC	SN and grav. lenses	•	•	•	•	•			•

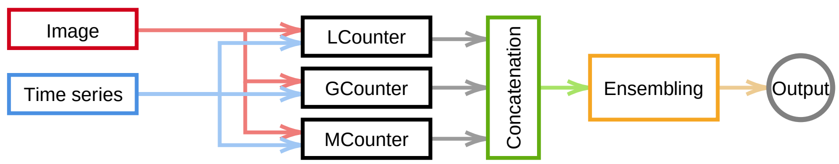

Figure 1: AstroCountNet (ACoNet) architecture

2.3 Metrics

The labels for all the tasks are imbalanced, so we use the macro F_1 score. Computing the macro F_1 in classification is straightforward. In counting, there are 7 possible counts, as shown in Table 2 (one per body type). Each count is treated as a class in a multi-class single-label classification. We apply the following procedure: (1) compute the macro F_1 score for each astrophysical body type, weighting each possible counting output equally and (2) compute the average of the macro F_1 scores, weighting each body type equally.

2.4 Architecture

The tasks listed in Table 2 are addressed with the two networks ACoNet for counting and AClaNet for classification.

ACoNet is a multimodal multi-output inference pipeline computing one count per body type. Figure 1 illustrates its architecture, which is formed by three multimodal multi-output subnetworks (LCounter, GCounter and MCounter) connected to an ensembling module, which takes in input the outputs of the three subnetworks and estimates the count for each object type. LCounter, GCounter and MCounter respectively focus on local features, global features and both local and global features. Each subnetwork has 7 output branches, one for each astronomical body type. For each branch, the network estimates the number of instances of that body type in that observation. Ensembling algorithms leverage the complementarity of the sub-networks to improve the count and classification predictions. For counting, we use Random Forest (RF)[1] and 3 of the ordinal regression algorithms implemented in the `mord` library[2]: Logistic All-Threshold (AT) [21], Ordinal Ridge (OR), and Least Absolute Deviation (LAD). We use the default hyperparameters, as defined in `mord` and `scikit-learn` [22, 23]. The ensembling pipeline takes in input the encoded representation of the counts for each sub-network and output encoding. It outputs the counts for each ensembling algorithm (and hyperparameters combination). The best ensembling algorithms and hyperparameters are the ones that maximize the average macro F_1 score on the validation set.

AClaNet has a structure similar to that of ACoNet. The ensembling module is replaced by a task adaptation module, with the same inputs as the ensembling module. The output is the predicted class for each sample. For this purpose, we use the three pre-trained multimodal multi-output subnetworks (LCounter, GCounter and MCounter), concatenate their outputs and apply to the resulting vector the classification algorithms of 4 families for task adaptation: Support Vector Classifier (SVC) (a

[1] As implemented in `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html` (as of November 2024)

[2] Version 0.3, available at `https://pythonhosted.org/mord/` (as of November 2024)

Table 3: Comparison of the macro F_1 scores for counting on the 4 datasets obtained by ACoNet, sorted by the mean of the F_1 scores across the datasets. The 1σ confidence intervals are obtained using bootstrapping with 50 samples.

	DES-deep	**DES-wide**	**DESI-DOT**	**LSST-wide**	**Mean**
SNCC	$0.6550^{+0.0215}_{-0.0349}$	$0.7901^{+0.0271}_{-0.0229}$	$0.9011^{+0.0138}_{-0.0130}$	$0.8180^{+0.0208}_{-0.0181}$	0.7910
SNIa	$0.7330^{+0.0243}_{-0.0301}$	$0.7533^{+0.0271}_{-0.0213}$	$0.9310^{+0.0139}_{-0.0189}$	$0.8684^{+0.0169}_{-0.0213}$	0.8214
LSNCC	$0.6811^{+0.0112}_{-0.0099}$	$0.8920^{+0.0059}_{-0.0080}$	$0.8998^{+0.0042}_{-0.0081}$	$0.8672^{+0.0086}_{-0.0075}$	0.8350
Lensed galaxy	$0.7726^{+0.0083}_{-0.0126}$	$0.8829^{+0.0073}_{-0.0084}$	$0.8713^{+0.0067}_{-0.0084}$	$0.9008^{+0.0066}_{-0.0072}$	0.8569
Galaxy	$0.8434^{+0.0107}_{-0.0063}$	$0.8756^{+0.0115}_{-0.0109}$	$0.8704^{+0.0062}_{-0.0081}$	$0.8820^{+0.0093}_{-0.0109}$	0.8678
LSNIa	$0.7899^{+0.0068}_{-0.0090}$	$0.8847^{+0.0099}_{-0.0081}$	$0.9373^{+0.0025}_{-0.0054}$	$0.9068^{+0.0066}_{-0.0055}$	0.8797
Star	$0.9461^{+0.0045}_{-0.0088}$	$0.9680^{+0.0066}_{-0.0089}$	$0.9476^{+0.0074}_{-0.0075}$	$0.9285^{+0.0081}_{-0.0139}$	0.9476
Avg. macro F_1	$0.7744^{+0.0125}_{-0.0159}$	$0.8638^{+0.0136}_{-0.0127}$	$0.9083^{+0.0078}_{-0.0099}$	$0.8817^{+0.0110}_{-0.0121}$	0.8571

kernel-based method), Logistic Regression (LR) (a linear model), Decision Tree (DT) (a tree-based model), and Multi-Layer Perceptron (MLP) (a neural network), as implemented in `scikit-learn`.

3 Evaluation

Table 3 summarizes the results for the counting task after ensembling LCounter, MCounter and GCounter and using the MOR and SOR output representations, which outperform the MCC representation ($+10\%$ macro F_1 score). Each row represents an astronomical body type, each column is a dataset and each cell contains the macro F_1 score for the counting task, where each counting value (0, 1, 2) is regarded as a different class. The results show that the F_1 scores are consistently high (> 0.65) across all datasets and objects. The most significant variations in results arise for transients, both lensed and not lensed, because of significant differences in the number of acquired samples across datasets [18]. Some objects are observed often (e.g., stars) or have peculiar and recognizable characteristics (e.g., LSN brightness variations are visible because of the gravitational lens). On average, ACoNet can find and count them better than other objects. In DESI-DOT, the non-lensed SN F_1 score is similar to that of other bodies, because the higher quality of observations mitigates the faintness and compensates for the scarcity of samples.

Table 4 presents the results of the four classification tasks. The table also includes a column labeled "Co", which shows the results obtained by ACoNet considering that an object is present if its count is at least 1. A prediction by ACoNet is considered correct if all the classes considered in the task are predicted correctly. Task adaptation, on average, improves the F_1 score ($+1\%$ to $+6\%$) with respect to the ACoNet-based prediction. The results are better for binary classification tasks ("Gravitational lenses" and "SN (coarse)"). The two multi-class tasks have similar results, indicating that the presence of lensed galaxies does not affect performances significantly. The lowest F_1 scores are observed for DES-deep. The results on the other datasets are similar because the quality of their images is comparable and better than that of DES-deep, as shown in [18, 19].

4 Conclusions and future work

This work introduces ACoNet, an ensemble of multimodal neural networks that counts bodies belonging to different classes in astronomical observations. We show that, for this task, ordinal regression with MOR and SOR is more effective than MCC (more than +10% macro F_1 score on average). We also show that ACoNet is effective at classifying gravitational lenses and SN but adapting ACoNet into AClaNet for classification tasks improves the macro F_1 score up to +11%.

Despite promising results, some limitations should be considered. This approach must be evaluated on a large set of real data and could be extended to more classes of gravitationally-lensed transients and to redshift estimation, which is otherwise time-consuming. Future work will concentrate on applying ACoNet to real observations, further extensions of the network to address redshift estimation and using explainability techniques.

Table 4: Comparison of the macro F_1 scores for classification. "Co" indicates the results obtained by ACoNet, "Cl" indicates the results obtained by AClaNet and "I" is the improvement. The 1σ confidence intervals are computed using bootstrapping with 50 samples.

	Gravitational lenses			SN (coarse)			SN (fine)			SN and grav. Lenses		
Dataset	**Co**	**Cl**	**I**	**Co**	**Cl**	**I**	**Co**	**Cl**	**I**	**Co**	**Cl**	**I**
DES-deep	$0.7527^{+0.0070}_{-0.0082}$	$0.8585^{+0.0046}_{-0.0063}$	0.1058	$0.7955^{+0.0042}_{-0.0075}$	$0.8566^{+0.0077}_{-0.0076}$	0.0611	$0.5505^{+0.0228}_{-0.0201}$	$0.5907^{+0.0146}_{-0.0135}$	0.0402	$0.5339^{+0.0125}_{-0.0158}$	$0.5935^{+0.0189}_{-0.0112}$	0.0596
DES-wide	$0.8519^{+0.0079}_{-0.0046}$	$0.9176^{+0.0058}_{-0.0059}$	0.0657	$0.8615^{+0.0040}_{-0.0060}$	$0.8852^{+0.0039}_{-0.0087}$	0.0237	$0.7245^{+0.0142}_{-0.0087}$	$0.7274^{+0.0128}_{-0.0224}$	0.0029	$0.7163^{+0.0153}_{-0.0128}$	$0.7297^{+0.0121}_{-0.0113}$	0.0135
DESI-DOT	$0.8780^{+0.0054}_{-0.0067}$	$0.9191^{+0.0069}_{-0.0047}$	0.0411	$0.9047^{+0.0068}_{-0.0046}$	$0.9113^{+0.0056}_{-0.0045}$	0.0065	$0.8615^{+0.0073}_{-0.0154}$	$0.8701^{+0.0076}_{-0.0104}$	0.0085	$0.8308^{+0.0075}_{-0.0114}$	$0.8526^{+0.0081}_{-0.0076}$	0.0219
LSST-wide	$0.9036^{+0.0088}_{-0.0039}$	$0.9270^{+0.0045}_{-0.0055}$	0.0234	$0.9200^{+0.0047}_{-0.0054}$	$0.9308^{+0.0038}_{-0.0054}$	0.0108	$0.7917^{+0.0171}_{-0.0141}$	$0.7972^{+0.0115}_{-0.0167}$	0.0055	$0.7833^{+0.0148}_{-0.0055}$	$0.7974^{+0.0133}_{-0.0134}$	0.0140
Mean	0.8456	0.9056	0.0590	0.8704	0.8956	0.0255	0.7321	0.7467	0.0143	0.7161	0.7456	0.0272

References

[1] L. A. Moustakas and R. B. Metcalf. Detecting dark matter substructure spectroscopically in strong gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 339(3): 607–615, March 2003. ISSN 1365-2966. doi: 10.1046/j.1365-8711.2003.06055.x. URL http://dx.doi.org/10.1046/j.1365-8711.2003.06055.x.

[2] Masamune Oguri. Strong gravitational lensing of explosive transients. *Reports on Progress in Physics*, 82(12):126901, November 2019. ISSN 1361-6633. doi: 10.1088/1361-6633/ab4fc5. URL http://dx.doi.org/10.1088/1361-6633/ab4fc5.

[3] A. Papitto, E. Bozzo, C. Ferrigno, T. Belloni, L. Burderi, T. Di Salvo, A. Riggio, A. D'Aì, and R. Iaria. The discovery of the 401 hz accreting millisecond pulsar igr j17498-2921 in a 3.8 h orbit. *Astronomy & Astrophysics*, 535:L4, November 2011. ISSN 1432-0746. doi: 10.1051/0004-6361/201117995. URL http://dx.doi.org/10.1051/0004-6361/201117995.

[4] G. Bélanger. On detecting transient phenomena. *The Astrophysical Journal*, 773(1):66, July 2013. ISSN 1538-4357. doi: 10.1088/0004-637x/773/1/66. URL http://dx.doi.org/10.1088/0004-637X/773/1/66.

[5] Jian-Min Wang, Jun-Rong Liu, Luis C. Ho, and Pu Du. Accretion-modified stars in accretion disks of active galactic nuclei: Slowly transient appearance. *The Astrophysical Journal Letters*, 911(1):L14, April 2021. ISSN 2041-8213. doi: 10.3847/2041-8213/abee81. URL http://dx.doi.org/10.3847/2041-8213/abee81.

[6] S. Charpinet, G. Fontaine, P. Brassard, and B. Dorman. The potential of asteroseismology for hot, subdwarf b stars: A new class of pulsating stars? *The Astrophysical Journal*, 471 (2):L103–L106, November 1996. ISSN 0004-637X. doi: 10.1086/310335. URL http://dx.doi.org/10.1086/310335.

[7] Dan Ryczanowski, Graham P Smith, Matteo Bianconi, Sean McGee, Andrew Robertson, Richard Massey, and Mathilde Jauzac. Enabling discovery of gravitationally lensed explosive transients: a new method to build an all-sky watch list of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 520(2):2547–2557, January 2023. ISSN 1365-2966. doi: 10.1093/mnras/stad231. URL http://dx.doi.org/10.1093/mnras/stad231.

[8] Patrick L. Kelly, Steven Rodney, Tommaso Treu, Masamune Oguri, Wenlei Chen, Adi Zitrin, Simon Birrer, Vivien Bonvin, Luc Dessart, Jose M. Diego, Alexei V. Filippenko, Ryan J. Foley, Daniel Gilman, Jens Hjorth, Mathilde Jauzac, Kaisey Mandel, Martin Millon, Justin Pierel, Keren Sharon, Stephen Thorp, Liliya Williams, Tom Broadhurst, Alan Dressler, Or Graur, Saurabh Jha, Curtis McCully, Marc Postman, Kasper Borello Schmidt, Brad E. Tucker, and Anja von der Linden. Constraints on the hubble constant from supernova refsdal's reappearance. *Science*, 380(6649), June 2023. ISSN 1095-9203. doi: 10.1126/science.abh1322. URL http://dx.doi.org/10.1126/science.abh1322.

[9] Edward Karavakis, Wen Guan, Zhaoyu Yang, Tadashi Maeno, Torre Wenaus, Jennifer Adelman-McCarthy, Fernando Barreiro Megino, Kaushik De, Richard Dubois, Michelle Gower, Tim Jenness, Alexei Klimentov, Tatiana Korchuganova, Mikolaj Kowalik, FaHui Lin, Paul Nilsson, Sergey Padolski, Wei Yang, and Shuwei Ye. Integrating the panda workload management system with the vera c. rubin observatory. *EPJ Web of Conferences*, 295:04026, 2024. ISSN 2100-014X. doi: 10.1051/epjconf/202429504026. URL http://dx.doi.org/10.1051/epjconf/202429504026.

[10] Nicolò Oreste Pinciroli Vago and Piero Fraternali. Deepgravilens: a multi-modal architecture for classifying gravitational lensing data. *Neural Computing and Applications*, 35 (26):19253–19277, June 2023. ISSN 1433-3058. doi: 10.1007/s00521-023-08766-9. URL http://dx.doi.org/10.1007/s00521-023-08766-9.

[11] Nicolò Oreste Pinciroli Vago and Piero Fraternali. Deepgravilens, 2023. URL https://zenodo.org/record/7860294.

[12] CE Petrillo, CRESCENZO Tortora, S Chatterjee, G Vernardos, LVE Koopmans, G Verdoes Kleijn, NICOLA ROSARIO Napolitano, G Covone, P Schneider, ANIELLO Grado, et al. Finding strong gravitational lenses in the kilo degree survey with convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, 472(1):1129–1150, 2017.

[13] CE Petrillo, CRESCENZO Tortora, S Chatterjee, G Vernardos, LVE Koopmans, G Verdoes Kleijn, NICOLA ROSARIO Napolitano, G Covone, LS Kelvin, and AM Hopkins. Testing convolutional neural networks for finding strong gravitational lenses in kids. *Monthly Notices of the Royal Astronomical Society*, 482(1):807–820, 2019.

[14] R Cañameras, S Schuldt, SH Suyu, S Taubenberger, T Meinhardt, L Leal-Taixé, C Lemon, K Rojas, and E Savary. Holismokes-ii. identifying galaxy-scale strong gravitational lenses in pan-starrs using convolutional neural networks. *Astronomy & Astrophysics*, 644:A163, 2020.

[15] D Stern, SG Djorgovski, A Krone-Martins, Dominique Sluse, Ludovic Delchambre, C Ducourant, R Teixeira, Jean Surdej, C Boehm, J Den Brok, et al. Gaia gral: Gaia dr2 gravitational lens systems. vi. spectroscopic confirmation and modeling of quadruply imaged lensed quasars. *The Astrophysical Journal*, 921(1):42, 2021.

[16] G. Angora, P. Rosati, M. Meneghetti, M. Brescia, A. Mercurio, C. Grillo, P. Bergamini, A. Acebron, G. Caminha, M. Nonino, L. Tortorelli, L. Bazzanini, and E. Vanzella. Searching for strong galaxy-scale lenses in galaxy clusters with deep networks: I. methodology and network performance. *Astronomy & Astrophysics*, 676:A40, August 2023. ISSN 1432-0746. doi: 10.1051/0004-6361/202346283. URL `http://dx.doi.org/10.1051/0004-6361/202346283`.

[17] Cameron Lemon, Frédéric Courbin, Anupreeta More, Paul Schechter, Raoul Cañameras, Ludovic Delchambre, Calvin Leung, Yiping Shu, Chiara Spiniello, Yashar Hezaveh, Jonas Klüter, and Richard McMahon. Searching for strong gravitational lenses. *Space Science Reviews*, 220(2), February 2024. ISSN 1572-9672. doi: 10.1007/s11214-024-01042-9. URL `http://dx.doi.org/10.1007/s11214-024-01042-9`.

[18] Robert Morgan, Brian Nord, Keith Bechtol, SJ González, E Buckley-Geer, A Möller, JW Park, AG Kim, S Birrer, M Aguena, et al. DeepZipper: A novel deep-learning architecture for lensed supernovae identification. *The Astrophysical Journal*, 927(1):109, 2022.

[19] Robert Morgan, B Nord, K Bechtol, A Möller, WG Hartley, S Birrer, SJ González, M Martinez, RA Gruendl, EJ Buckley-Geer, et al. Deepzipper ii: Searching for lensed supernovae in dark energy survey data with deep learning. *arXiv preprint arXiv:2204.05924*, 2022.

[20] Coen de Vente, Pieter Vos, Matin Hosseinzadeh, Josien Pluim, and Mitko Veta. Deep learning regression for prostate cancer detection and grading in bi-parametric mri. *IEEE Transactions on Biomedical Engineering*, 68(2):374–383, February 2021. ISSN 1558-2531. doi: 10.1109/tbme.2020.2993528. URL `http://dx.doi.org/10.1109/TBME.2020.2993528`.

[21] Jason Rennie and Nathan Srebro. Loss functions for preference levels: Regression with discrete ordered labels. *Proceedings of the IJCAI Multidisciplinary Workshop on Advances in Preference Handling*, 01 2005.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[23] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.