



TINAJO 2022 27 de octubre *October 27*





Colabora



AI in the Middle Age: Parchment Study and Arrangement by Appearance-Based Recognition





Prof. Pierluigi Feliciati Prof. Emanuele Frontoni

University of Macerata

TEAM

Archivio di stato di Milano:

Benedetto Luigi Compagnoni - *Director*; Carmela Santoro - *Archivist, Deputy director*; Antonella Cesarini - *Archivist*; Davide Bruno De Franco - *Archivist, Director Archivio di Stato di Novara*.

University of Macerata & Marche Polytechnic University:

Emanuele Frontoni, *Professor - Computer Science and Computer Vision*Pierluigi Feliciati - *Professor, Records and Information Science*Marina Paolanti - *PhD, Assistant professor*Andrea Felicetti - *PhD, Postdoc*

University of Pavia:

Luca Fois - PhD, Postdoc, Dipartimento di studi umanistici

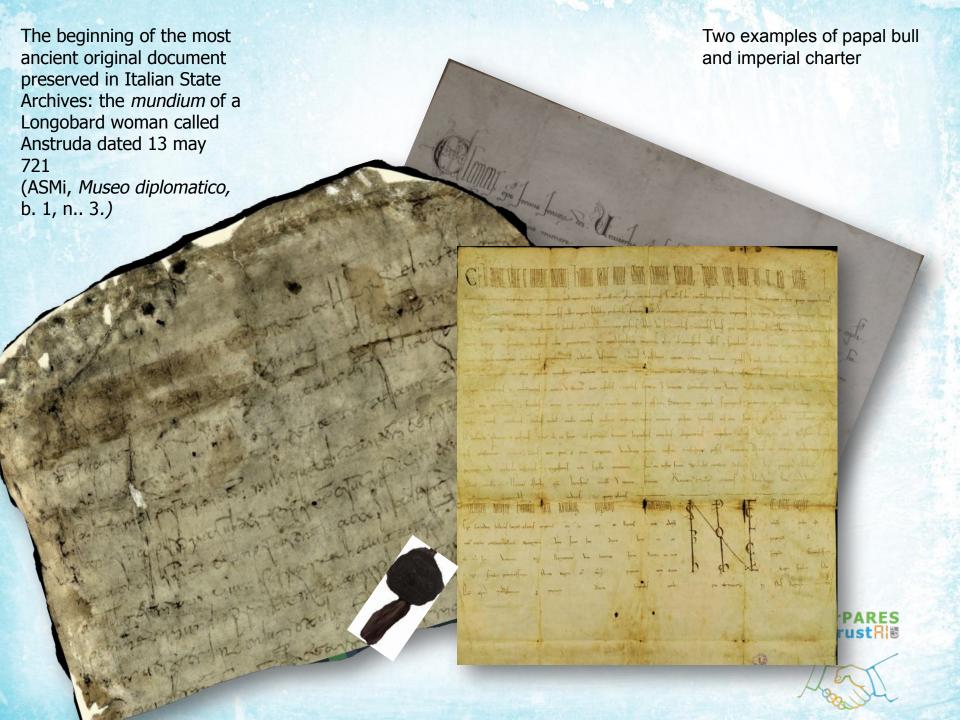
Washington University:

Joseph T. Tennis - *Professor, Associate Dean for Faculty Affairs, Executive Director of Administrative Services*



INTRODUCTION

- Ancient written private documents on parchment were an important communication mean to humankind and have, for that motive, an invaluable historical value to our culture heritage and at the same time they are authentic deeds among people, valid till nowadays.
- Therefore, it becomes imperative that they are preserved and perpetuated in order to extend their life span in the interests of population and future generations.
- Currently, digitization of historical parchments is extraordinarily convenient as it allows easy access to the documents from remote positions and removes the need for possible adverse physical management.



AN OUTSTANDING HERITAGE

- The Archivio di Stato of Milan (ASMi) preserves more than 130.000 documents on parchment, covering nearly a thousand years of history.
- The main collections of parchment documents are the Museo diplomatico and the Pergamene per fondi.
- The Museo diplomatico holds the most ancient Italian original private deed, dating back to the year 721 and a large number of documents preceding the year 1000.
- The Pergamene fondi holds imperial and royal charters, papal bulls, chancery documents, and private deeds coming from a huge variety of ecclesiastical and secular institutions of northern Italy (12th-18th centuries).





anno mice vicar mille uno conteximo vigerino teto mile aplintic pma. No inti no ambrolinta ariatori gerni unti fili qua perri agila mat a fili arq; lantrane influt st qua adam a dorada mat ei a bellunda ama cos a reluta qua ambrolin omi stoco udulucto qui presti fum lege unere longobardos consenembi nobil qua sup feminis pliches und salois ment y dentum marchis a obitonis gornos entere es suflos arq; p. daram licentia ponetil inful ab albro undice e millo dom regil apo durin In omipl ac redemptor ne anunal qual conder as thudul alutil Comp inuitar e do not qu'ambroliul a arrabul a lantrane unt le agila a soraba ar q' bellunda solum a undicam l'au phoe mum munolabile undicam etir main un campul und no quem habé un sum un tro loco a fundo udulucto ad locu ubi de un connano choera es ama sce marce amer se stefant aser sobil amon se prasis e et ple campul q miura unta pric quinq; omil ple camp cumbrub; e accollionib; sul inin plem die e ora Suemat in une e prie tare monarii lei ambrohi ubi ei seum conti est corp un tan or nos qu'insti e un broes teneam pseum campi ad laborandu e ad una unilitare facienda una tin ut not e un brose solumnul at qu'sm om anno infestimitate sei mar un fucti adiplium monarium & formito modui unu ad mura uphul loci a babbas usua monaris a es successores bent dare gustare comuni qui upsum such conducera pane & formito a carne a una a poc un dua conducera pane & formito a carne a una a poc un dua conducera plum monaris a es successores de carne a una a poc un dua conducera plum monaris un superior de carne a una a poc un dua conducera plum monaris de plum monaris un superior de carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una apoc un dua conducera pane & formito a carne a una carne a pane de carne a una carne a una carne a pane de carne a una carne a pane de carne a carne a carne a una carne a carne a carne a carne a una carne a carne m' ur tinol aut ver broef non dorun plum hetu omi armo al uplum monarui ur tinol zune uplu monarui babeat a tenear uplum campu qua dui der m ur anolaut urs brod non sor of a lipostophul monarii fe cert sup unpostam & uplo campo nobil aut ur il broth Tune n plaluam uplu ficta mulipsu hetu adupsu monarus ut sups a si postophul monarii se cert sup unposta a nord broth; a postoqua dumiser males mul iplu hetu as unu monar u mon ari non reliquera nobil iplam sup impostra a nord brow a polique domilera nobil iplam sup impostra a nord brow car indican uno conore sa como supostra nune prolum monaru ad uplum monarus at tup & qua sie & creut nea bona nolumas und due car motocan uno conore sarpre sunt actu Molos Mont III manua illos ambrosis a arraldo gernos a lanstrance un for a gotte a dorado arqi bellunde ghoc midicara un sup herrogamer a spis unta posetal fement esenter un sup

If allered where a musting some regret possed inflat becomes deduce sup a St

Thu Hmanus antelme arealde landrance perer arnalde while pagane landrance refles;

to (go attiul noral lac palacu leruph polt tout copt in to Soi;

American lister de milleure conqueros sons sons montes montes reportes de la la monte de la film de la company de carefa relata from about a stability come for some restable of the some many of the some and the some of the some ple amizone abasses of manufacture of the sole regular a mandally mile to a language of prenent and the notice of property of parents and the property of parents of to come of a minima one a mande of the sole regular of munded la minima of the sole and the sole of th comer on a universe, a middle sume colerante for some und abo hoic milla par in unolonia in mallona portanea voluntate has car world for a the form a milland par in unolonia in mallona portanea voluntate has car world for a the form a milland for com manufatratione or of about mention of manufatra of the learning of the lea possib protes or les buted our dies reacount in long about monier of que not only in latter a some fante in fully a possib all protes due of influence protection of the protest of the protest of the protection of the protest of the protection of points print a corbit & bufet sura me record for corner and no rotal garrar to enful the a another fence in ful to a point all print que are pitet rott as flux all med matterns to infer rord for a sufer rord for enful for infer roading rotation of the rotation of the rotation as pare of monor and matterns. de mayer & prose mo us in buse & monare en super ior de sencifinde a recession for inin to insuper the monare of monare of insuper of monare of monare of the fuer a series of contract who post monare in the first series of the formare of of emerge prose me of in the monator in furth he form north and chie a oralism lette who plu monather in well at not began nee lotal tre of monator in falls to plu monather in a lette who plu monather in a lette who plu monather in a lette in the formate of the formate of the monather in the formate of the formate of the monather in in chem querere coffaciament et montres felle spacina fuert falla de esta montre praca ul trà niam restaurare de da tapais que con montre praca ul interes coffaciament esta montre de la tapais esta de la m in chem querre of an un of money to make a partie of money to antique a money of antique and a producer makerya Co antique and an antique and a producer makerya Co antique and an antique and a plumon of application of application of application of application of applications of appli atuere que de la compare de puilla engle e rescundo us abuci manore. Compla monagnie de los facere fossare monte de un la abor monagnie de los facere fossare monte de un la abor monagnie de los facere fossare monte de un la abor monagnie de los faceres fossare monagnie de la compla monagnie de los faceres fossare monagnie de la compla monagnie de los faceres fossare monagnie de la compla monag por of part of monary to one of the second o gredeunds cu carra gen frais and arababa fropal fact que a zarete entabelle a amyon de beluste I hacker would ut fut for requiers a sple will ed mater fue a col sume com The first about manh & ord enrole marriage of the focus more and about congresse of the about manh man' popular to a small lake maple goes of in aber l'ograno a unifred geriog pe aune a rafalle gern qu'ilabelle a ber ria affatti et or loroni genuire ip beluite q car introgaucit a ratie car ad cofu mand man popular to be bennom source marnery graffing or pagar Lanfracia of buzalem alter of the curitor octain guarula genalo Spaule avails baftard Johannom & pozel loonar be talleaferm andre tel tie Go to the dender of crips post cridea copten seded of radica believed nepota med ou polite lorers quere se introquer Hat plant plat with leder wasta poult with a arrest new a of quanto magnified a albe gerne infante file of a cre aller habiter to legima crate forester fuere haber of face aller co put meder land amore. ad parte of monara. I con por post tan omit orin p fto pero recepto le si official infraction face habent face harm of afficial and popular of Sulfore this bennon. Frante in pora & labor of propagation der to the soll debre and the propagation of monarm & fue por I better rous of police and fidualise the beamon's feater in pena labor cour . Canter that qual a placet of recept habone our en laudaines under ab allis unland aghi emer's april prace of dare habene upla car emperore est morians . use after both marrie and possible prace of dare habene upla car emperore est morians . use after both marrie and possible prace of dare habene upla car emperore est morians . use after both marrie and possible prace of the both marries and possible prace to fit willow the got live inpera line income

Inne drice anear milletime contre fimo quadragelimo tercio: less teat madi unic telt. Placet ang consient un doni quitredis monateris foi ambrolis subi el seion regeler 33 th comp cir consiste monachy fuou on landally got graful aucent appeal quice got hope watally of monaris the atomatic husant plant confirmant let no moraris appeal que green na zamerani gdi donco giulicione gdi d'enere arq caramera ad parte urilitare i plou romin scola ecote sei peri qui ad lala que sie quanta erair. It indi noe d'hear due sie iplem som sple abt. Ente inter ambrolo gamerano leu guilleron any cammero astabenda romenti leu centa reddenda libellario noe ula asamo migner mone expleto dans maia ula impremi De modol tre l'Strent Diffure medal fieth one anno dani d'emfiguari incaneur iplini monari le ambroli. La midelet racione un adil inalle collecte ple fedarel Buch heli " fice eves page sphi feel with feele unte fuert. E- plot une event Mon soft feelavel Cisé monaro fi regient fuert conti fragul anul argi, si, boi, uni - cantella una la har spie att fine feet line feet line feet sin hete candella to punter spie att fileer puritire poul mounts fie hie infer selacabrus poultu fiet som aune sare spil feelatib fie sup le cu suit sucher Illia sup imposite en int ad no fire pend i int posici in que inplis at con tiste in succession so shac concenera libelli monere plumser - inca omia ur suf l'no primisera te con ponar illa part que los no confusiore pares fise fuarre pene nos arqui. Se lond: mesto, libret: Ses - octo - intup in case conseniera libelli primer una ipie abbacceper abiptir seolarib place libello mon di bonon libra none quel ipre abb Mor ad parte iprint monari formero file qu'a porti que bene medial permeto que permeto monare interest - funds domergates ; cadelle marrano asquage, interes mandrones les sie un el come acti els manaros of goo we freder attr and facto ff ego pomer nur m 4 + Ego 2015 m 7 sac fre sun somit Tomi HH man illi langulli auscari que fur y firmaner; Tout I I man ich gungern gunge nafallou grit fuer 7 homaner. Tout HI II manuit quarmeris caquelles sources quidantolis malmatini que statemant perfacti stagnante scornoli arq: p ambrosis que stomaco telle i: arad now fact palaces int fue progue ffig. siste les perrul nord septi poli poli poli - si

InterPARES Trust



NEW PATHS OF RESEARCH

- Though studied and edited since the eighteenth century, this
 extensive material has not been investigated systematically in
 order to identify common features recurrent in homogeneous
 groups of documents.
- Processing automatically a large number of scanned documents can lead to new understandings of general cross-cutting issues, not related to a single fund.







AI & DEEP LEARNING FOR THE PARCHMENT DOCUMENTS

- One of these issues is defining the number and the activity of the notaries in the city of Milan and its contado.
- The AI deep learning approach can automatically investigate a large number of private documents and give a substantial contribution to the knowledge of Milanese notaries in the Middle Ages.
- The basic feature chosen to be identified for the research is the signum, an authentication element put by notaries.
- The signum appears in about all private deeds and has specific and easily recognisable charateristics (shape, position, etc.)



InterPARES Trust⊞



THE SIGNUM: A USEFUL MARKER

- The signum or Notarial sign is a specific and personal drawn mark used by a single notary in the protocollo and before his signature.
- Identifying the signum means that every notary could be recognised and tracked in a virtually infinite series of documents.
- The AI will contribute either in create a library of signa: a virtual matricula of notaries and a basis to investigate their less visible features.













PERGANET: A LIGHTWEIGHT DL-BASED SYSTEM

- The AI will be tested for an appearance-based recognition of ancient parchments and documents, implementing the use of AI in the archival science: a method that could be reproduced by many other Archives and for different types of documents.
- It is proposed PergaNet a lightweight DL-based system for historical reconstructions of ancient parchments.
- The aim of PergaNet is the automatic analysis and processing of huge amount of scanned parchments.





BUILDING UP A NOTARIAL MATRICULA BEFORE AD 1350

- One of the aims of the project is to use the AI to build up a Matricula of Milanese notaries of the twelfth and thirteenth centuries (in separate steps).
- The progressive investigation of all the parchment documents of the ASMi will lead to a complete list of notaries and their identification.
- The process will also generate a list of all the surviving documents related to every single notary.



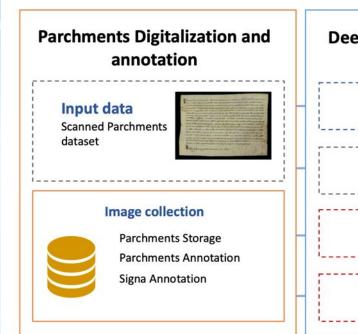
BUILDING UP A NOTARIAL MATRICULA BEFORE AD 1350

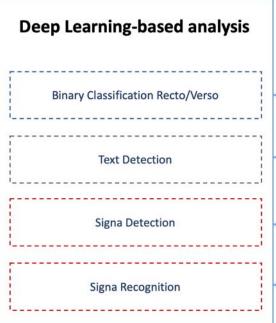


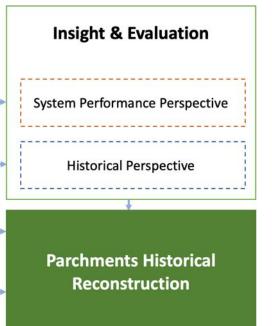
terPARES Trust | | | | |



PERGANET







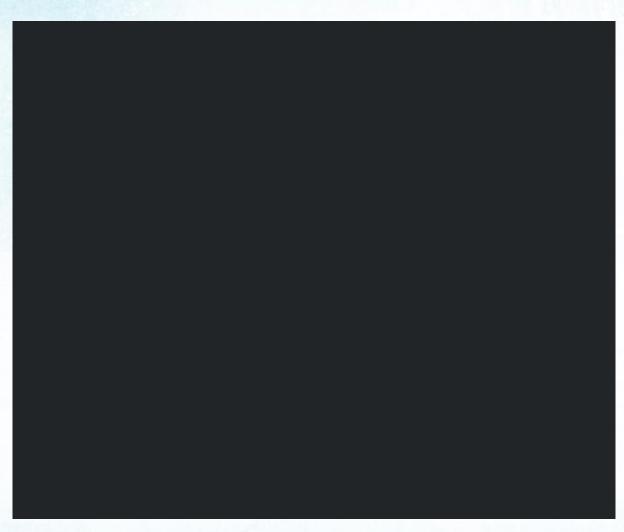




DATASET COLLECTION

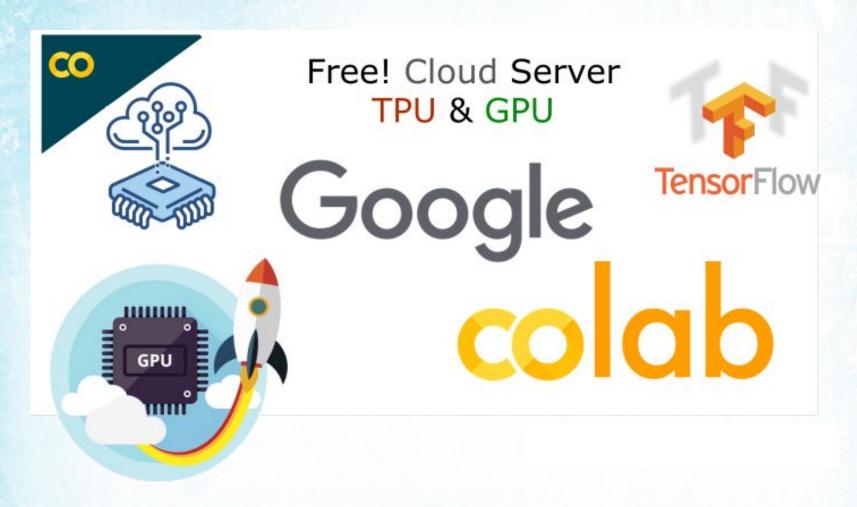
- Data collection is moving forward and actually we have the following archives cooperation together and collecting data for the study:
 - Archivio di Stato di Milano
 (220 documents already processed, nearly 1200 scanned and ready to be processed)
 - Archivio di Stato di Novara (91 documents ready to be processed)
 - Archivio di Stato di Ascoli Piceno
 (27 documents ready to be processed)





InterPARES TrustRI®



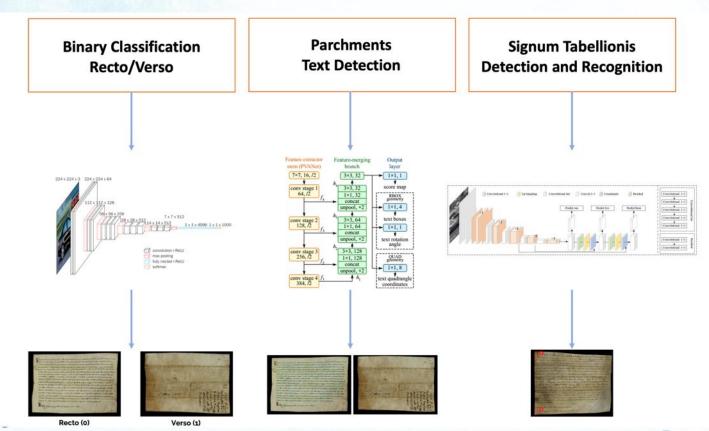


PARCHMENTS CLASSIFICATION & SEGMENTATION https://drive.google.com/drive/folders/1lj8lj VTr4wTtgyBHIGO9TmKK0qJ-xxV





DEEP LEARNING PIPELINE







BINARY CLASSIFICATION RECTO/VERSO

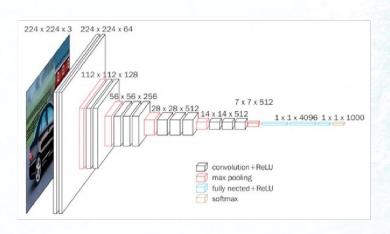
VGG16 DEEP NEURAL NETWORK



Recto



Verso

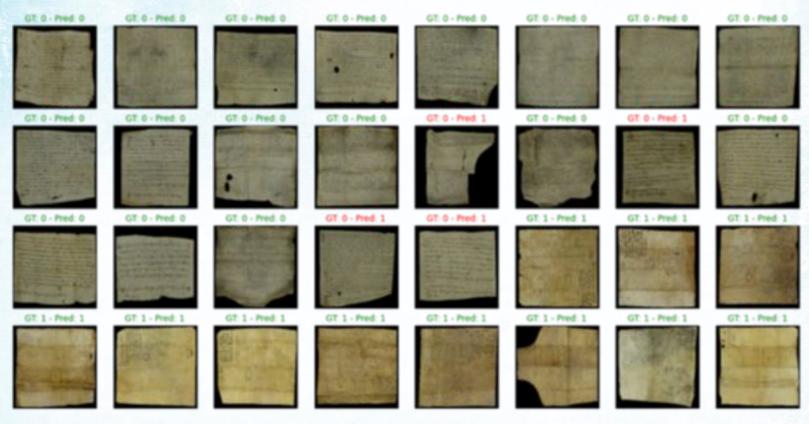


Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.





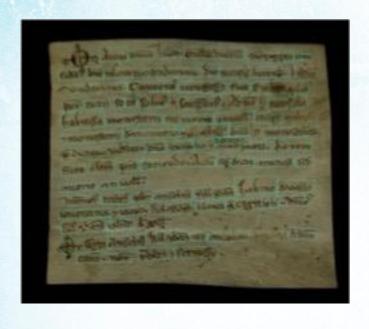
BINARY CLASSIFICATION RECTO/VERSO: RESULTS

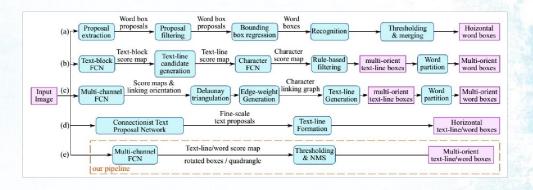


InterPARES Trust R ₪



PARCHMENTS TEXT DETECTION





Xinyu Zhou., et all, A. (2017). EAST: An Efficient and Accurate Scene Text Detector. arXiv:1704.03155v2.

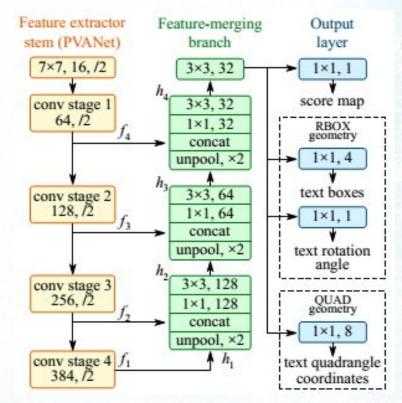




PARCHMENTS TEXT DETECTION: EXPERIMENTS

Word detection

- DNN Model: EAST
- Use of trained model (tested over datasets: ICDAR 2015, MSRA-TD500, COCO-Text)



InterPARES Trust RI®



PARCHMENTS TEXT DETECTION: RESULTS

The second of the second place of the second place of the second place of the second o

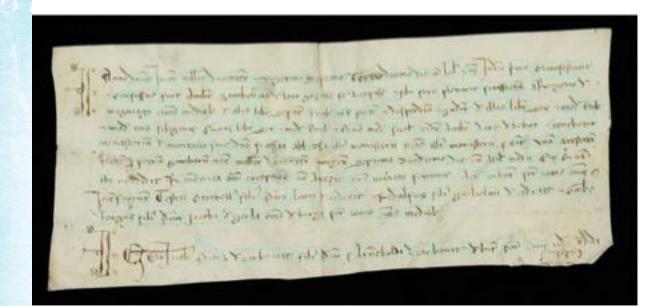
They got good - cont was did.







PARCHMENTS TEXT DETECTION:

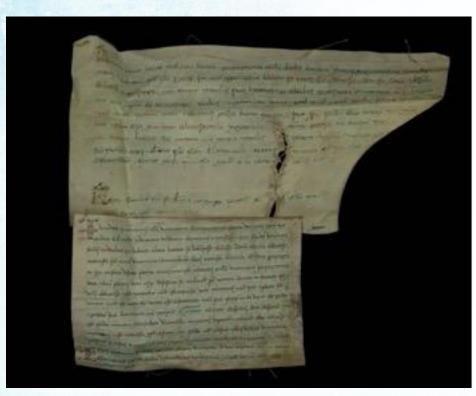


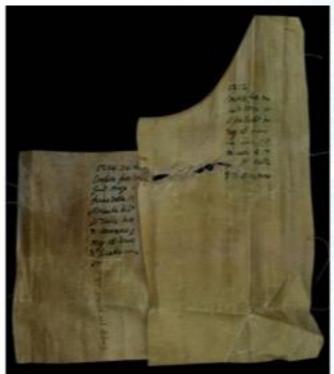


InterPARES Trust | | | | |



PARCHMENTS TEXT DETECTION: RESULTS

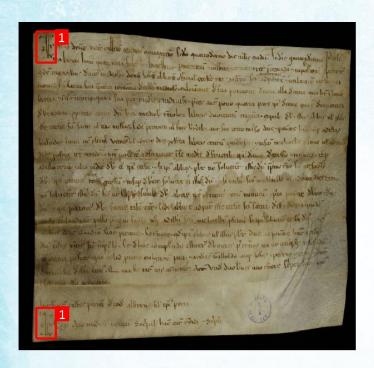


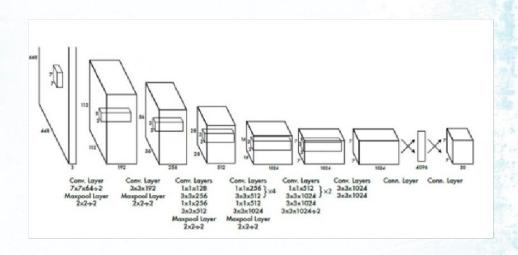






SIGNA TABELLIONIS DETECTION AND RECOGNITION





Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

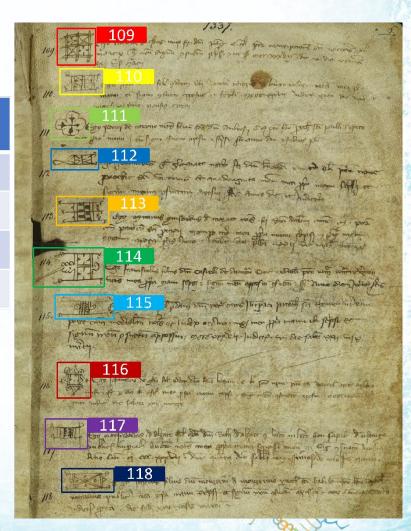




NOTARIAL MATRICULA ANNOTATION

2768 signa

ID	NAME	HISTORICAL INFORMATION
	•••	
109	notary 109	
110	notary 110	



SIGNA TABELLIONIS DETECTION AND RECOGNITION: RESULTS

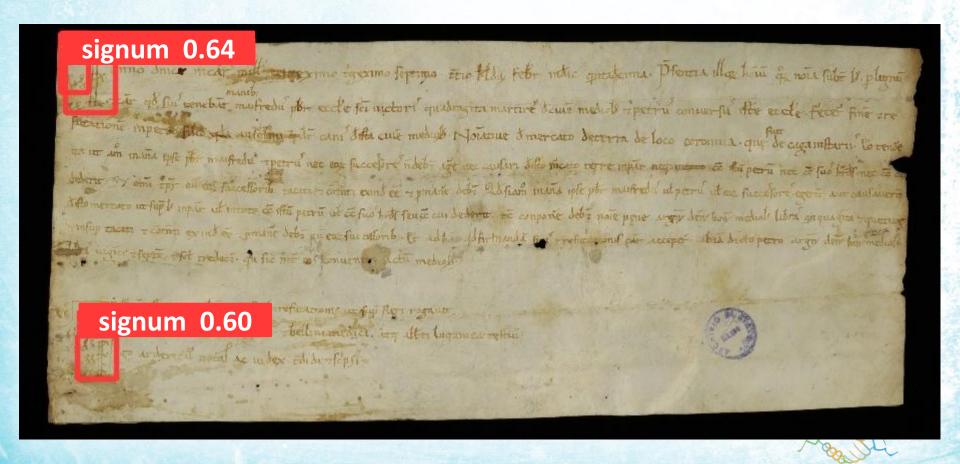
The above how quez note have a muchur processes suffer or out to secured motione person gor aspentius Some meduly. Sona lobe albare official coolie se water for sparce unturne coolie or nonce la latte her que somme South medoli Asianut Stus porcione deven ille Sound que beiloco berto ta renvero que e las por modie prudinto pirer sur pono quetes par qui dame que desnomen & bourano ppierro uses du bou medioli combis liber Suveenti ragion orquet Storthis al las ut plus the cocke for farme at are mellion live persons at him liedels and his cores mile due parties in the adultion halinder hum no pland vemo let sour despotter liber course quider soule machaelle plane utasses Sie poten it water up porter reference it medie d'incerte un diene d'accerte un denne descrite minute inter adella rema sales madie & be ex upi coche . fr upi abbay phe no foluere 1th, du tymo eno he ur lapte Il the period torn grotte nely & but folices en the Do to make for mobile the power det ten no Polisette ille su he ar lapte Small St abert we perwer were minure por moure Shire its In supe perroce to facere tale our trocalbare adpar the cutie for facer. Sills Some quate more et landent pulle piro ur lapte use adelhi for mechaelle plant le que Edward or the die I have some trade too perwer to trace of weather utillar the due acquire hat elis du come time he importe le shoe somplado chece stucies permos us acumanas a fal de Daciner politic que acted proces outgrown pure vardice tallolde any letter perrue of the Attended Salle law. I me he me on a went. Dow Vad due ling un chore lepoplate son Tollers die menter. In land rolling part of allery fil up pers. when also modern morgan sadged but ant itom "soph





SIGNA TABELLIONIS DETECTION AND RECOGNITION: RESULTS

For each prediction is showed class (signum) and confidence

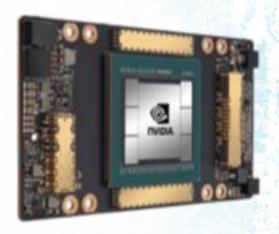


SIGNA TABELLIONIS DETECTION AND RECOGNITION: RESULTS

walia loli q de unlan some module laçane q de moner da as parre monarie of sed el curamalle france an maria un in prosignum 0.73 reportequementemente facere debes ple all a facere des facere report fuel in about in monarie d'infine mille calé ein quale ruier in monarie laudancent p per librar ochracient and I omile practi a ger bet a bujed, que sole who after report deriver ut about money at little reported for more in the information of more in the contract of the practical o amoline what garrard infurte of amoline arm fante infubr , of fin willam at ali willam & matroniane habens grate of uph willam to note y fe ut & alief fuel femores a use of the capite spears of whit ally in butto & monaire splendil adgrees debat plu practi abillo and & of facer for facer of process of the facer debat sple will over a facer debat face on split fue reported in laudative author or monafterin pollet formare a habe cluse inferior of fue neports of a abalia reparentable a unado inferior of light al The new lette als spie monared ucharet state it toller of lotal fre needigna afterpla chiefa inquerganeret protes al fre phi while fluor reports in delic querore references and menares to fight chule fuere faits abopte munde in fubr. sengurgament proses ut fra con retaining the segretaining of menaria. Or infug facere bette sin angle menarial felo media of fuer with a spin mount of court motion freducts after open of a contract of frence of the spin carril et en om en urdrate altrapai actibia paccofia act of rie of pirit ip wood at fair reports a public sundo aff at acti manore a retembre. E rif perit oper ar breear factor of mente Toffara fig tra if who get neporta ab una par use or de das Co ur ople who or fur neport it debeit contribuere est monarro levare ponte fup wellows uto monathered habet rope ab una pir of de das Co signum 0.36 signum 0.36 analos ur fup fure regaun signum 0.32 or secure w fue of has an absolut manda manu popul signum 0.76 artini galadi arealis lasterisi landidi peris caralis st arealis palaris cestisi signum 0.72 Co so stood of under freque has breue forept

EXPERIMENT SETTINGS

- DNN Model: VGG16
- Image reshape to 224x224
- Data normalization: mean subtraction, std scaling
- Augmentation with image rotation and flip
- Batch size: 16
- Adam optimizer
- Learning Rate: 10e-4 (from 0th to 10th epoch)
- 10e-5 (from 10th to 20th epoch)
- Cross-Entropy loss
- Training from scratch
- 20 epochs







OTHER ARCHIVAL APPLICATIONS

Once refined and fully developed, the AI could be used in a wide range of applications:

- Recognize the peculiar system of writing of every single notary.
- Analyze archival notes on the back of the documents and retrace previous archival arrangements or uses of groups of documents.
- Recognize recurring images or other features in huge series of documents.
- Identify common patterns in manuscript maps or drawings.
- Make publicly available original & relevant AI Data Set.
- And many others...









AI in the Middle age Arrangement of the documents via appearance-based recognition

Prof. Emanuele Frontoni emanuele.frontoni@unimc.it

VRAI homepage: http://vrai.dii.univpm.it

Publications: http://vrai.dii.univpm.it/publications

Software & Datasets: http://vrai.dii.univpm.it/datasets

YouTube: shorturl.at/htvN2

LinkedIn: https://it.linkedin.com/in/emanuelefrontoni



