### WRITER IDENTIFICATION AND WRITER RETRIEVAL USING VISION TRANSFORMER FOR FORENSIC DOCUMENTS

Michael Koepf, Florian Kleber and Robert Sablatnig



# MOTIVATION

- The law states that a handwriting comparison is enough evidence to identify a person
- Is used in forensic analysis to conclude authorship of a questioned document, e.g. threat or ransom letters
- Carried out by handwriting experts
  - Time consuming
- Law enforcement agency have archives of writing samples



# WRITER IDENTIFICATION/RETRIEVAL

- Writer identification is the task of determining the author of a sample handwriting from a set of writers
- Writer retrieval is the task to obtain all documents of one writer out of a set of documents



 Atanasiu, V., Likforman-Sulem, L. & Vincent, N., "Writer Retrieval - Exploration of a Novel Biometric Scenario Using Perceptual Features Derived from Script Orientation"; ICDAR, 2011, 628 - 632

### WI/WR WITH AND WITHOUT ENROLLMENT

- "With enrollment": samples for each writer in the training set (similar to other classification tasks, e.g., image recognition) (cf. [Chr18])
- "Without enrollment": writers in the training set and test set are disjoint Thus writers cannot be predicted directly

 [Chr18] Christlein, V.: Handwriting Analysis with Focus on Writer Identification and Writer Retrieval. Ph.D. thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (2018)

# CHALLENGES

- Different writing styles (Fig. (a))
- Different scripts

• • •

- Variations in handwriting of one individual (Fig. (b))
- Similarities between different individuals
- Used material (pen, paper) (Fig. (c))
- Intentional modifications—obfuscation



(a) WRITE dataset (writer ID 13)(b) CVL Database (writer ID 2)(c) WRITE dataset (writer ID 16)

### CHALLENGES

Senater Allen Ellender, of Louisiana, sported off the opposition by telling a television anotience it was "current Washington gossip" that Weaves once had Communist altitiations. The Senate Banking Committee, which is headed by another Souther Senater -Willis Robertson, of Virginia - met today in closed session to discuss Weaver's appointment. Senator Robertson labe disclosed he had sent a welle to the Kennedy saying he had received several complaints about weaver's longality. Sendtor Allen Ellendes, of buisiana, spooled off the opposition by telling a television andience it was a current Washington gossip "that weaves ance had communist affiliations. The Senate Barling Committee, which is headed by andhes Souther Sender - within Robebon, of Virginica - net today in closed session to discuss weaves's appointment. Sender Roboton lates disclosed he had sent a lelles to Mr. Kennedy saying he had neceived

There was a computed 8.4 per cent. owing towards her Nathanaiists parts of Dr. Vowoerd, her Price print star. This implies here are appeared of the Proquerrise party frem the Hame. Carty results in yesterday's generat election showed Nathanaists being netweet in their strangholds, the Bloemtentia and the Transocal, with shighty increased majorities. In the Cape, the velesar limited Party politicen, thr. Harry Lowence, staring a a Programice, was

### METHODOLOGY

State-of-the-art methods mainly use Convolutional Neural Networks (CNNs)

Vision Transformers (ViTs) – self-attention-based Deep Neural Networks

- current state of the art for image classification on ImageNet

Research question:

How do Vision Transformers, trained from scratch, perform on WI/WR tasks?

# VISION TRANSFORMER (VIT)

#### Used architecture: ViT-Lite-7/4 [HWS+21]

- Input image size: 32x32
- Patch size: 4x4
- Embedding space dim: 256
- Encoder: 7 layers

[HWS+21] Ali Hassani, Steven Walton, Nikhil Shah, Abulikemu Abuduweili, Ji- achen Li, and Humphrey Shi. Escaping the Big Data Paradigm with Compact Transformers. arXiv:2104.05704 [cs], June 2021. arXiv: 2104.05704.



### Writer identification with enrollment



"With enrollment": samples for each writer in the training set (similar to other classification tasks, e.g., image recognition)

ViT acts as a feature-extraction-classifier-combination

# EXPERIMENTAL SETUP

Dataset: CVL [KFDS13;KFDS18]

- CVL w/ enrollment
  - Latin script (English and German reference texts)
  - 310 writers (2 pages for training, 1 for validation, 2–4 for testing)
- CVL w/ enrollment–English: subset of CVL w/ enrollment (only 1 page for testing)

Procedure:

- Training: Hyperparameter optimization—Grid Search
- Evaluation of the test set with the top-k criterion

#### CVL sample

And fortune, on his dammed quartel smiling, Show'd like a rebel's whore: but all's too weak: For brave Normas - well he deserves that name – Disdening fortune, with his brandish'd steel, Which smoled with bloody excention, Lake valour's minion carved out his passage Till he faced the slave:

And fortune, on his demand donned quarrel miling, Show'd like a cebel's inhore: but all's too wook: For broke domacswell he deserves that mane - Disdaining fortune, with his brandish'd sheel, which moded with bloody execution, hile valant's minima carved and his pressage till he faced the share;

13-2

kud forhure, on his demand donned quarrel miliug, Show'd like a rebel's whore: but all's too weak: For brave Mamacswell he deserves that name - Disdaining forhure, with his brandish'd sheel, which moked with bloody execution, hike velour's minion canved and his pressage till he faced the shave;

Segmented page-level image

Original scan

# RESULTS

### CVL w/ enrollment

	Top-1	Top-5
He and Schomaker [HS20]	0.991	0.994
He and Schomaker [HS21]	0.994	0.997
Proposed	0.990	0.999

### CVL w/ enrollment–English

	Top-1	Top-5
Khan et al. [KKTB19]	0.990	n/a
Kumar and Sharma [KS20]	0.994	1.000
Javidi and Jampour [JS20]	0.962	n/a
Proposed	0.990	0.997

### Writer identification without enrollment



"Without enrollment": writers in the training set and test set are disjoint

- Writers cannot be predicted directly

ViT only acts as a *feature-extractor* (generic—trained for domain) Classifier: k-nearest neighbors (kNN)

# EXPERIMENTAL SETUP

#### Datasets:

- CVL w/o enrollment
- ICDAR 2013: Latin and Greek script
- WRITE: Latin script, different writing styles (cursive handwriting, block letters)

		Training	Training Set		Test Set	
Dataset	# Writers	# Writers Pages	#	# Writers	# Pages	
CVL w/o enrollment [KFD\$13;KFD\$18]	310	27	189	283	1,415	
ICDAR 2013 [LGSP13]	350	100	400	250	1,000	
WRITE	16	n/a	n/a	16	202	

#### ICDAR 2013 samples

If we desire to avoid insult we must be able to repel it. If we desire to secure peace one of the most powerful instruments of our hising prosperity it must be known that we are at all times ready for war.

The willingness with which our young people are likely to serve in any war no matter how justified shall be directly proportional to how they perceive veterans of early wars were treated and appreciated by our nation.

Περιφρονείτε τα βιβλία εδείς που όλη βας η Jush είνου βουτημένη μέδα στη ματαιότητα μαι 6τη φιλο δοβία μαι στην αναβήτηση των πδονών ή μέβα στην οκνηρία. Αλλά σμεφτείτε πως όλος ο μηωστός κόβμος δεν μυβερνιέται παρά μόνο με τα βιβλία εκτός απ' τα άχρια έθνη.

Ο άνθρωπος που δίνει καλή δυμβουλή χτίζει με το ένα κέρι. Ο άνθρωπος που δίνει καλή δυμβουλή και κόλό παράδαμια χτίζει και με τα δύο που πέρια. Αλλά εκείνος που δίνει και κοική δυμβουλή και καικό παράδειμοι με το ένα χέρι χτίζει και με το άπο βκρεμίζει.

### WRITE sample

ZU- UND VORNAME: HARTIGOVÁ BXRBOCA.

#### SCHRIFTPROBE

Der folgende Text ist <u>einmal in Schreibschrift</u> (Fichte, Tanne, Kiefer ...) und <u>einmal in</u> Druckschrift (Fichte, Tanne, Kiefer ...) abzuschreiben:

Sh

Fichte, Tanne, Kiefer, Lärche, Ahorn, Rotbuche, Ulme und Weide sind Bäume unserer Heimat. Citrus, Gummibaum, Ingwer, Vanille, Jasmin, Myrte und Zeder hingegen tropische Pflanzen. Nerz und Otter liefern Petzwerk. Saphir und Diamant sind Edelsteine, Quarz ist ein Mineral. 1 2 3 4 5 6 7 8 9 0

Fielde, Tame, Kiefer, därde, Ahorn, Kotbrede, Henre und Weide sind Baume Unterer Heidred. Citrus, Summibaum, Ingner, Verille, Josunin, Myrte was und Feder Hingsgen tropische Pfleuriten. New und Feder Lingsgen tropische Pfleuriten. New und Feder Liegen Polynerk, Sophin und Diamant sind Eduskine, Querz ist ein Rinenel. A 2 3 4 5 6 7 8 8 0 Fichte, Tanne, Kiefer, Lärche, Ahorn, Rotbuche, Ulme und Weide Sind Bäume unverer Heimet. Citrus, Gummibaum, Ingwer, Vanille, Josmin, Myrte und Zeder hingegen tropische PFlanzen. Nerz und Otter liefern Pelzwerk. Saphir Und Diamant sind Edelsteine, Querz ist ein Minerel. 1.23456789 20

19.03. 1939 Barbora Hel

Fielde, Tame, Kiefer, darde, Ahorn, Kottorche, liene und Weide sind Baume Unker Heichet. Citaus, Summibann, Ingnier, Venille, Jesunin, Myrte was und Feder Winggen tiopische Pfleurzen. Nerzund Otter eigen Forzuerk, Sophin und Biamant sind Edelskeine, Querz ist ein Rinenel. A 2 3 4 5 6 7 8 9 0 Fichte, Tanne, Kiefer, Larche, Ahorn, Rotbuche, Ulme und Weide Sind Baume unverer Heimet. Citrus, Gummibaum, Ingwer, Vanille, Josmin, Myrte und Zeder hingegen tropische Pflanzen. Nerz und Otter liefern Pelzwerk. Saphir Und Diamant sind Edelsteine, Querz ist ein Minerel. 1.23456789.2

Segmented page-level image

Fichte, Tanne, Kiefer, Larche, Ahorn, Rotbuche, Ulme und Weide Sind Baume unverer Heimat. Citrus, Summibourn, Ingwer, Vanille, Josmin, Myrte und Zeder hingegen tropische Pflanzen. Nerz und Otter liefern Pelzwerk. Saphir Und Diamant sind Edelsteine, Quarz ist ein Minerel. 1.234567890

Fielde, Tanne, Kiefer, dorde, Ahorn, Kotbide, hene und Weide sind Beume Unkerer Heidret. Citrus, Summibann, Inguer, Varille, Josunin, Myrte see und Jeder hinggen tropische Pfeurzen. Nerzund otte eigen Potzverk, Sophin und Diamout sind Echeskine, Querz ist cir Riccol. 1 2 3 4 5 6 7 8 9 0

<sup>ge</sup> Segmented text regions (cursive handwriting and block letters)

Scan

### EXPERIMENTAL SETUP

#### • Training: Hyperparameter optimization—Grid Search

• 2 models (CVL w/o enrollment, ICDAR 2013)

In total 3 available models (feature extractors):

- CVL w/ enrollment
- CVL w/o enrollment
- ICDAR 2013

# EXPERIMENTAL SETUP

- Evaluation procedure [LGSP13]:
- Every document is used once as a query
- Calculation of distance to every other document in the test set
  - Ranking of documents for each document by distance/similarity
- Evaluation measures:
- Soft-top-k
- Hard-top-k
- Mean average precision (mAP)
- Best results are presented

# CVL W/O ENROLLMENT

	Top-1	H-Top-2	S-Top-5	mAP
Fiel and Sablatnig [FS15]	0.989	0.976	0.993	n/a
Christlein et al. [CBMA15]	0.994	0.988	n/a	0.978
Tang and Wu [TW16]	0.997	0.990	0.998	n/a
Christlein et al. [CBH+17]	0.992	0.984	0.996	0.980
Liang et al. [LWH21]	0.990	0.982	0.993	0.970
Proposed (ICDAR 2013 + transformer encoder + std. Euclidean)	0.974	0.950	0.984	0.928

# ICDAR 2013

	Top-1	H-Top-2	H-Top-3	S-Top-5	mAP
Fiel and Sablatnig [FS15]	0.885	0.405	0.158	0.960	n/a
Christlein et al. [CBMA15]	0.989	0.832	0.613	n/a	0.886
Tang and Wu [TW16]	0.990	0.844	0.681	0.992	n/a
Christlein et al. [CBH+17]	0.997	0.848	0.635	0.998	0.894
Keglevic et al. [KFS18]	0.989	0.779	0.564	n/a	0.861
Liang et al. [LWH21]	0.950	0.452	0.369	0.981	0.650
Proposed (ICDAR 2013 + transformer encoder + Canberra)	0.970	0.747	0.543	0.986	0.844

# WRITE DATASET

Method (feature extactor + feature vector + metric)	Top-1	S-Top5 S	-Top-10	mAP
CVL w/ enrollment + MLP head + std. Euclidean	0.851	0.911	0.921	0.534
CVL w/o enrollment + transformer encoder + std. Euclidean	0.777	0.881	0.906	0.514
ICDAR 2013 + transformer encoder + std. Euclidean	0.817	0.906	0.921	0.533

### WRITE DATASET

#### WRITE-Cursive handwriting (184 text regions)

Method (feature extactor + feature vector + metric)	Top-1	S-Top-5
CVL w/ enrollment + transformer encoder + std. Euclidean	0.859	0.908
CVL w/o enrollment + MLP head + Cityblock	0.810	0.880
ICDAR 2013 + transformer encoder + Cosine	0.832	0.902

#### WRITE-Block letters (145 text regions)

Method (feature extactor + feature vector + metric)	Top-1	S-Top-5
CVL w/ enrollment + MLP head + std. Euclidean	0.848	0.931
CVL w/o enrollment + MLP head + std. Euclidean	0.779	0.917
ICDAR 2013 + transformer encoder + std. Euclidean	0.800	0.897

# CONCLUSION

#### Writer identification with enrollment Results compete with state of the art

Writer identficiation/Writer retrieval without enrollment

- Future research necessary to improve results
- Better encoding (e.g., Vector of Linearly Aggregated Descriptors (VLAD))

Additionally analyzed:

- Influence of different writing styles
- Influence of different scripts (Latin and Greek)

### FUTURE RESEARCH DIRECTIONS

ViTs with convolutional layers

Pre-trained ViTs (transfer learning)

Word-level images—explainability through visualization of Attention (e.g., Attention Rollout [AZ20])



### WRITER IDENTIFICATION AND WRITER RETRIEVAL USING VISION TRANSFORMER FOR FORENSIC DOCUMENTS

Michael Koepf, Florian Kleber and Robert Sablatnig



thank you