Detection of Hand Drawn Electrical Circuit Diagrams and their Components using Deep Learning Methods and Conversion into LTspice Format

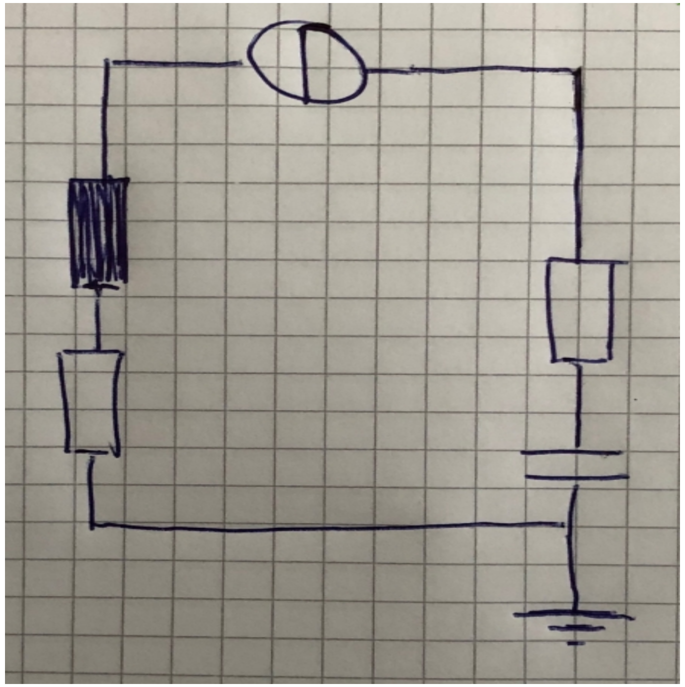
Master Thesis: Final Talk

Dmitrij Vinokour

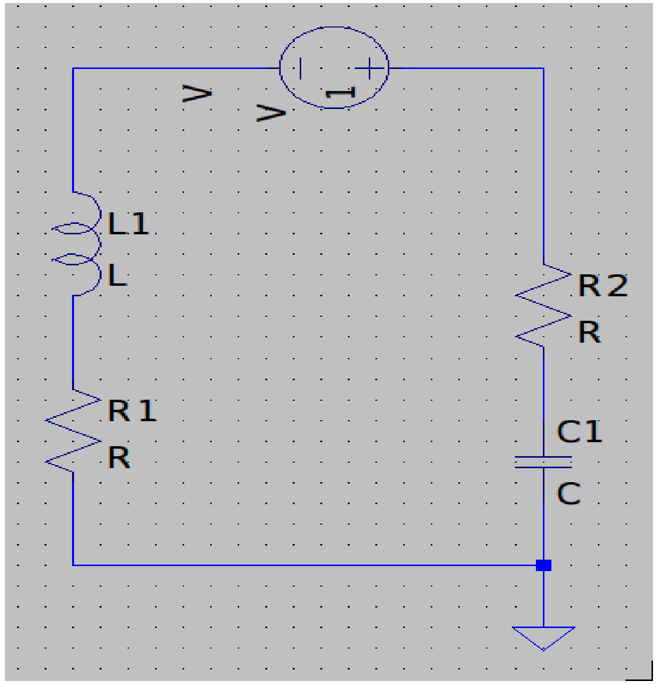
Supervised by Florian Thamm M. Sc., Felix Denzinger M. Sc., Prof. Dr.-Ing. habil. Andreas Maier,
Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nuremberg

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Introduction



Hand-Drawn Electrical Circuit



Model in LTspice

Motivation

- Drawing speed of hand-drawn diagrams ~90% faster than drawing in a tool (UML, flow charts, Markov chains etc. vs. Microsoft Visio) [1]

- Take the best out of both worlds

	Real World	**LTspice World**	**This Thesis**
Drawing	Fast	Slow	Fast
Calculating	Slow	Fast	Fast

[1]: K. Refaat et al., A New Approach for Context-Independent Handwritten Offline Diagram Recognition using Support Vector Machines. 2008.

Goals

1. Create a dataset of hand-drawn electrical circuit diagrams

2. Conversion Pipeline

    - Detect circuit components and their annotations

    - Operate on various paper backgrounds (white, checkered)

    - Generate Topology

    - Conversion into LTspice format

3. Evaluation algorithm

Dataset

- Collection of electrical circuit diagrams in German notation
- 31 contributors (21 train / validation, 10 test)
- On average 7.6 circuits / person
- Labels:
    - Bounding boxes labeled with orientation
    - Circuit masks for segmentation
    - Topology labels
    - Matching labels

	Images	Background	Annotated	train ratio	valid ratio	test ratio
	110	white		74.66%	6.69%	18.65%
	17	checkered		17.64%	11.77%	70.59%
	89	white	✓	78.65%	11.24%	10.11%
	21	checkered	✓	9.52%	19.05%	71.43%
Total	239			45.12%	12.19%	42.69%

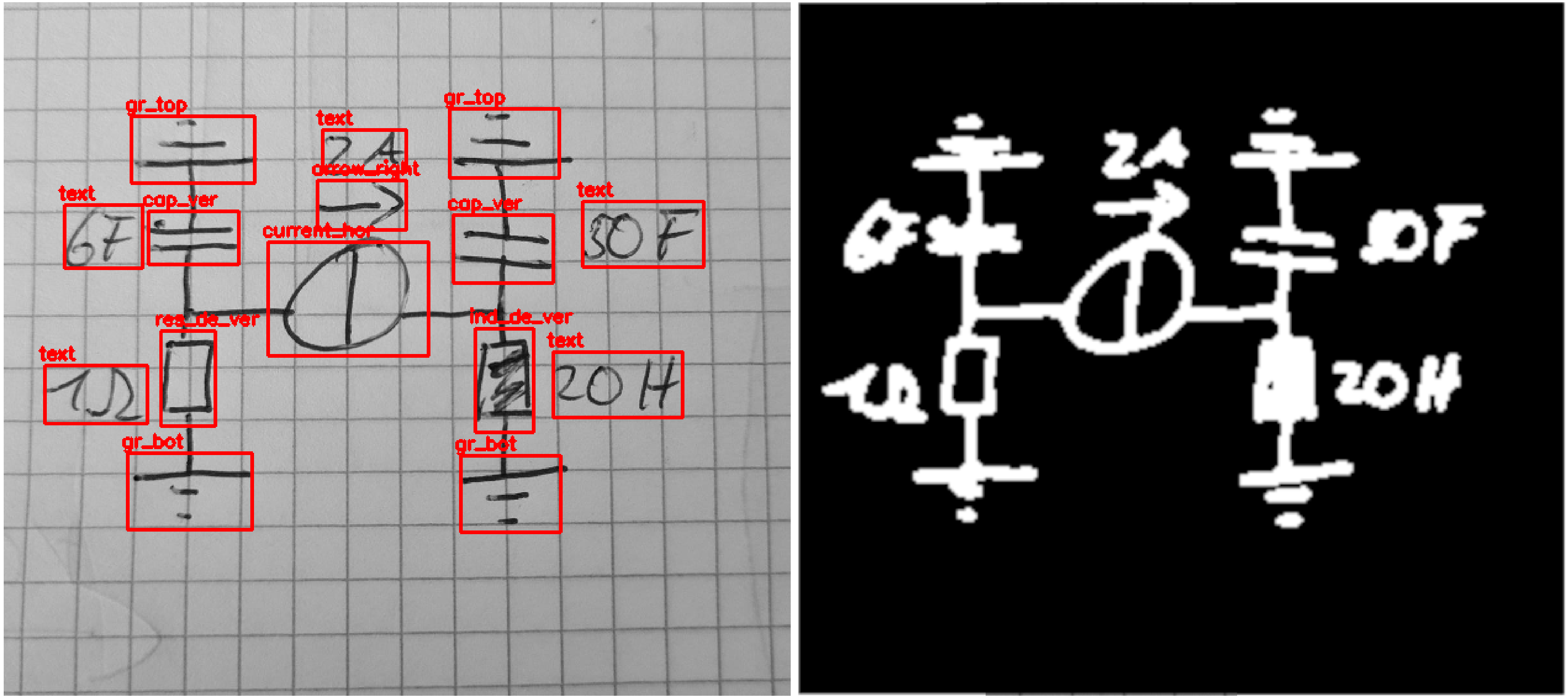
Overview of the circuits based on their background, annotation type and the train / validation / test split

Dataset

class	total	train ratio	valid ratio	test ratio
diode left	156	83.33%	8.97%	7.69%
diode top	210	82.38%	6.19%	11.43%
diode right	150	82.00%	12.67%	5.33%
diode bottom	102	67.65%	15.69%	16.67%
resistor horizontal	318	71.38%	6.92%	21.70%
resistor vertical	350	66.00%	6.57%	27.43%
capacitor horizontal	405	85.68%	4.94%	9.38%
capacitor vertical	268	65.30%	10.45%	24.25%
ground left	137	72.99%	10.95%	16.06%
ground top	137	81.02%	13.87%	5.11%
ground right	116	78.45%	14.66%	6.90%
ground bottom	178	73.60%	14.04%	12.36%
inductor horizontal	251	76.89%	8.37%	14.74%
inductor vertical	290	73.45%	9.31%	17.24%
source horizontal	188	77.66%	11.17%	11.17%
source vertical	238	64.71%	14.29%	21.01%
current horizontal	202	77.72%	9.41%	12.87%
current vertical	220	75.00%	12.73%	12.27%
text	877	61.92%	16.76%	21.32%
arrow left	57	70.18%	19.30%	10.53%
arrow top	77	64.94%	23.38%	11.69%
arrow right	105	70.48%	16.19%	13.33%
arrow bot	104	71.15%	15.38%	13.46%
total	5136	73.65%	12.27%	14.08%

Classwise data overview

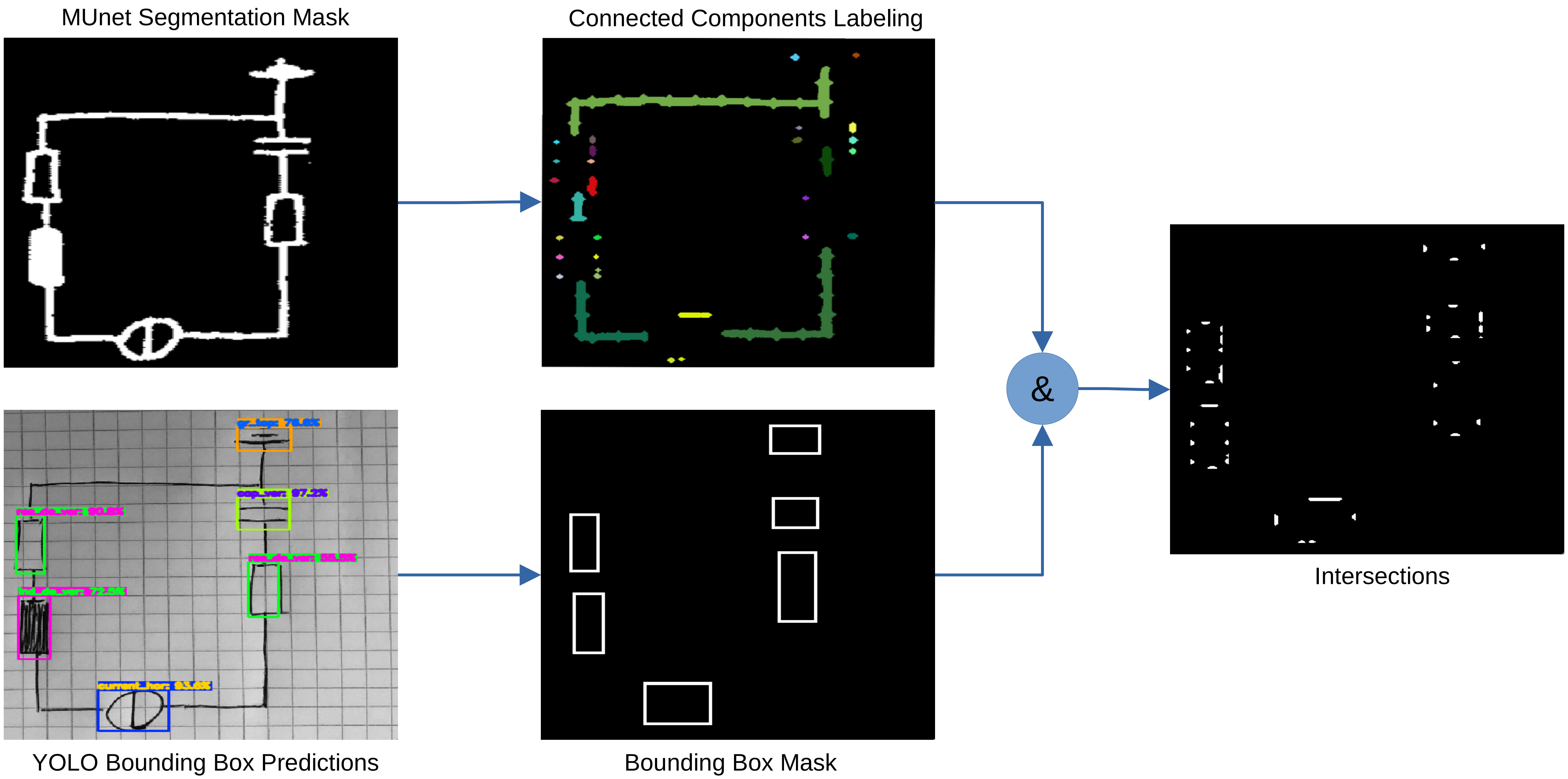
Dataset



Example labels for object detection (left) and segmentation (right)

Pipeline: Overview

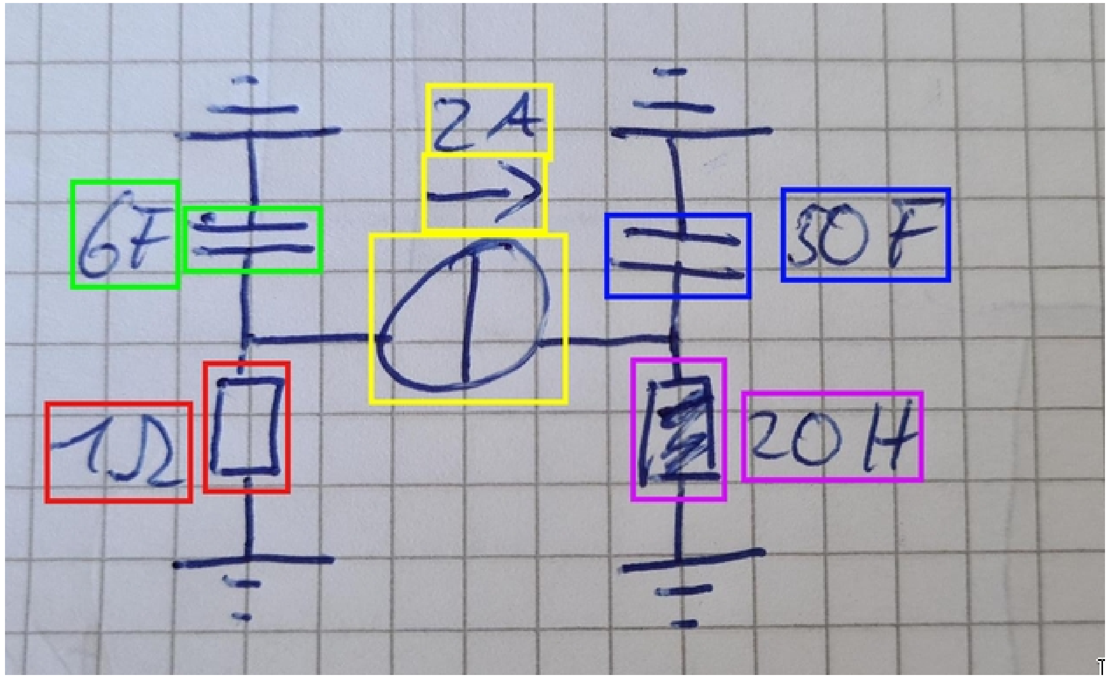
Pipeline Stage 3: Topology Creation



Pipeline Stage 4: Annotation Matching

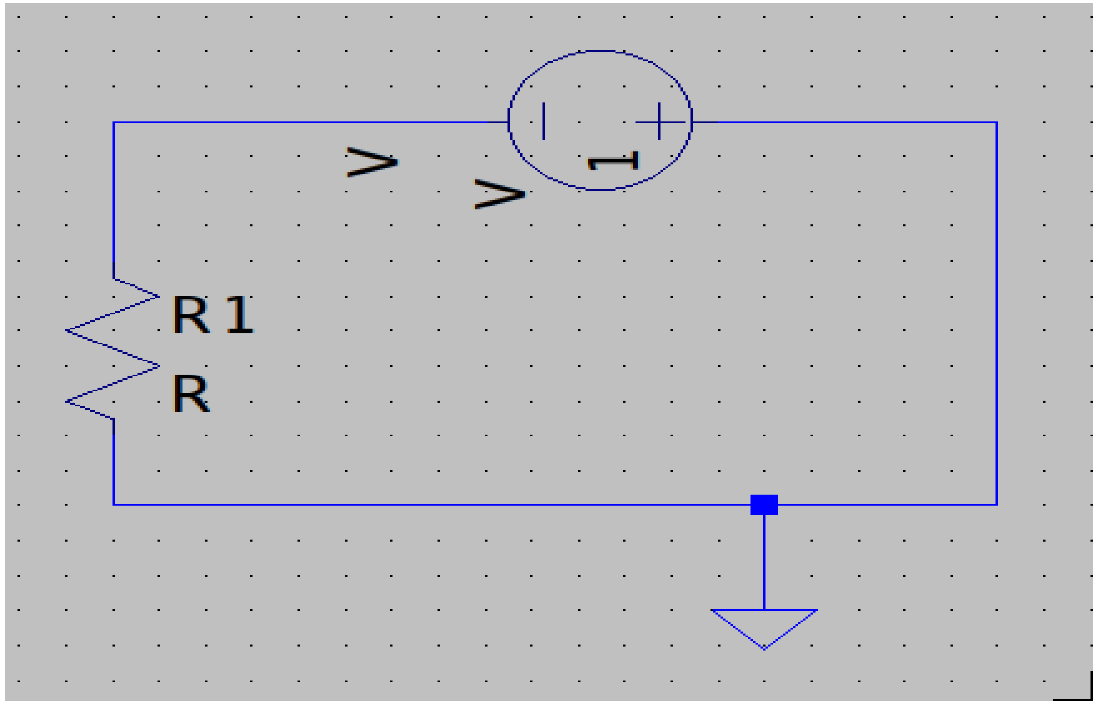
- Annotation matching based on nearest neighbor

- Brute force without considering multiple matches



Matched Annotations

Pipeline 6: LTspice Conversion



Model in LTspice

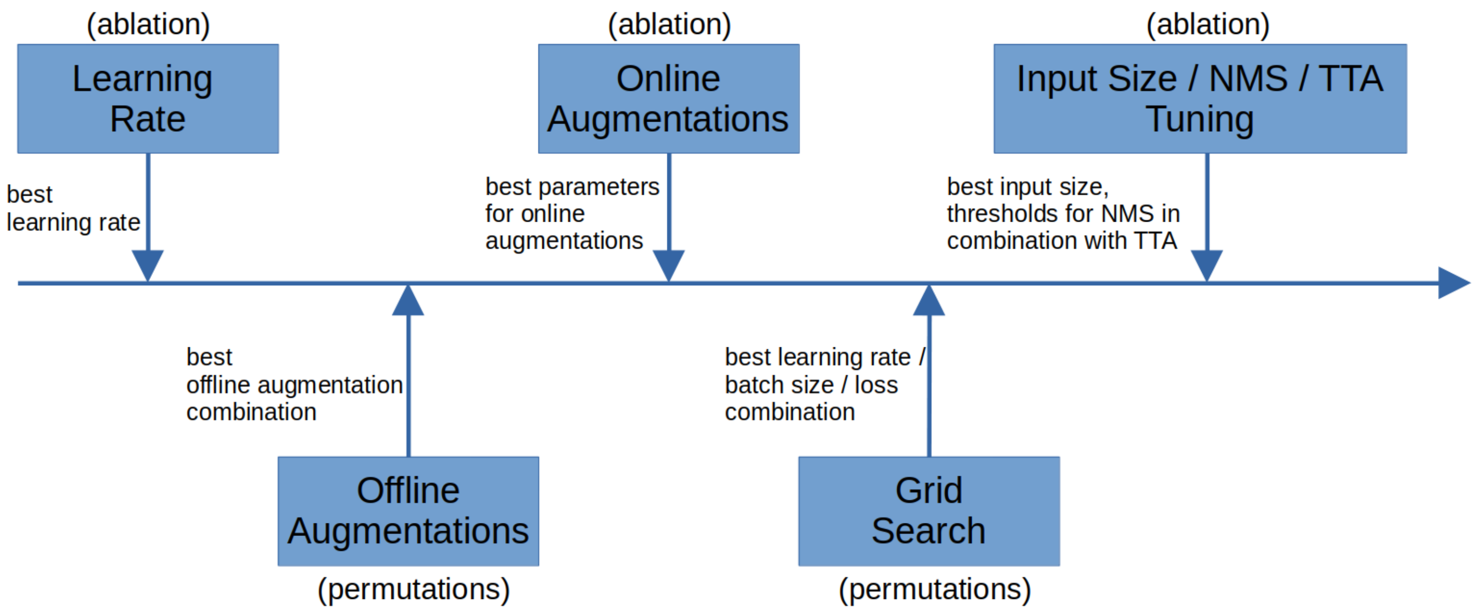
```
Version 4
SHEET 1 880 680
//    x1   y1  x2   y2
WIRE 416 80 288 80
WIRE 592 80 496 80
WIRE 288 144 288 80
WIRE 288 256 288 224
WIRE 512 256 288 256
WIRE 592 256 592 80
WIRE 592 256 512 256
WIRE 512 304 512 256
// ground
//    x  y
FLAG 512 304 0
//      comp     x   y  rot
SYMBOL voltage 512 80 R90
SYMATTR InstName V1
SYMBOL res 272 128 R0
SYMATTR InstName R1
```

Corresponding .asc

Training: YOLO

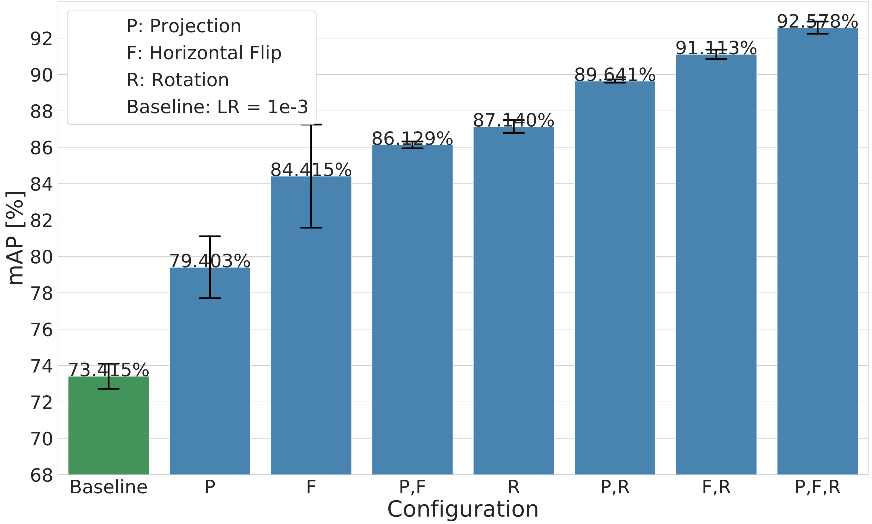
Configuration:

Batch Size: 64
Loss: CIoU
Optimizer: SGD with Momentum
Learning Rate Scheduler: 1000 batches exponential ramp-up
Input Size: 608x608x1
Metric: mAP@0.5:0.75:0.05



Training: YOLO Offline Augmentation

- Augmentations: Rotation 90°, Horizontal Flip, Copy-Paste Augmentation (Projection) [2]
- Rotation 3x 90°
- If flip and rotation also 3x 90° of flipped image
- Copy-Paste Augmentation 6 different backgrounds (different grid size and color)



Results of the offline augmentation experiment

[2]: G.Ghiasi et al., Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation. 2020.

Training: YOLO Online Augmentation and Grid Search

Offline augmentation: 92.578%

Online augmentation configuration:
- Three parameters per augmentation
- Augmentations: rotation, scale, safe crop, color jitter
- Rotation best with: 95.368%

Grid search configuration:

Learning Rates: 1.0e-2, 5.0e-3, 2.5e-3, 1.0e-3, 5.0e-4, 2.5e-4, 1.0e-4
Batch Sizes: 32, 64
Loss: CIoU [3], EIoU [4], Focal-EIoU [4]

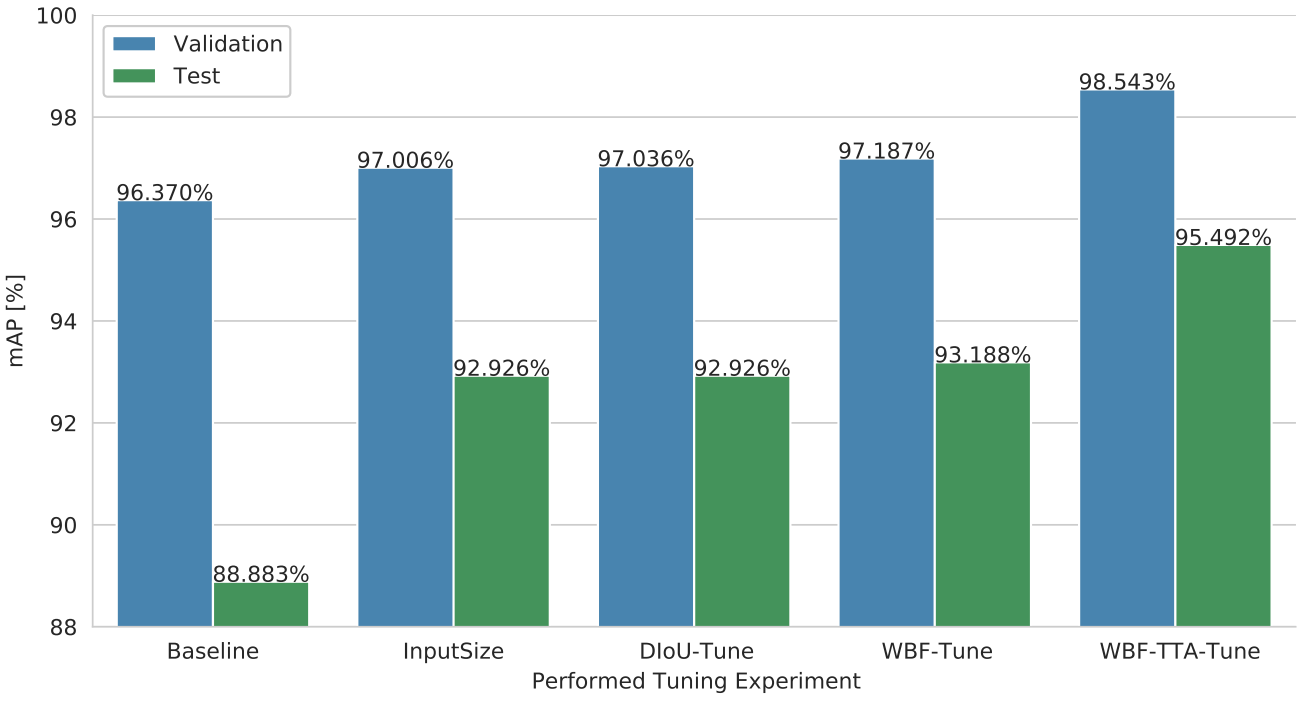
Results:
- Batch size of 32 consistently worse than batch size of 64
- EIoU > CIoU > Focal-EIoU (difference ~1%)
- Best combination LR 1.0e-2, BS 64, EioU with: 96.370%

[3]: Z. Zheng et al., Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. 2019.
[4]: Y. Zhang et al, Focal and Efficient IOU Loss for Accurate Bounding Box Regression, 2021.

Training: YOLO Final Results with Post-Training Fine-Tuning

- Baseline: Best performing grid search parameter combination

- All tuning was done on the validation dataset

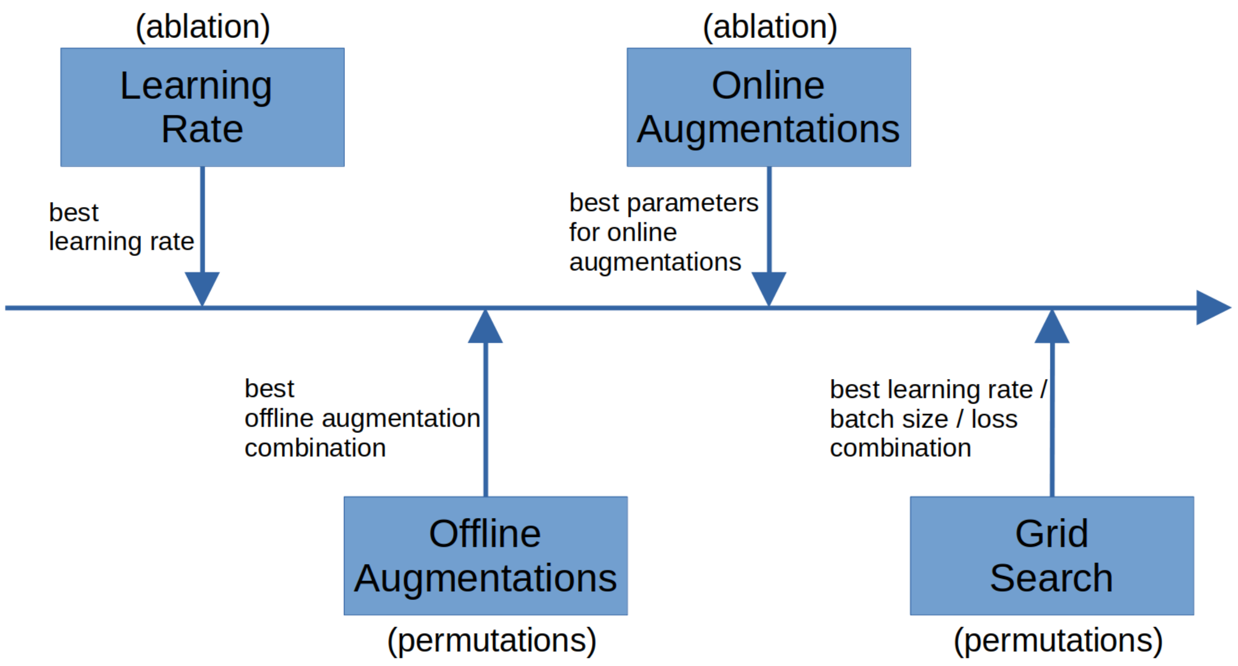


Results of the post-training fine-tuning on the validation and test set

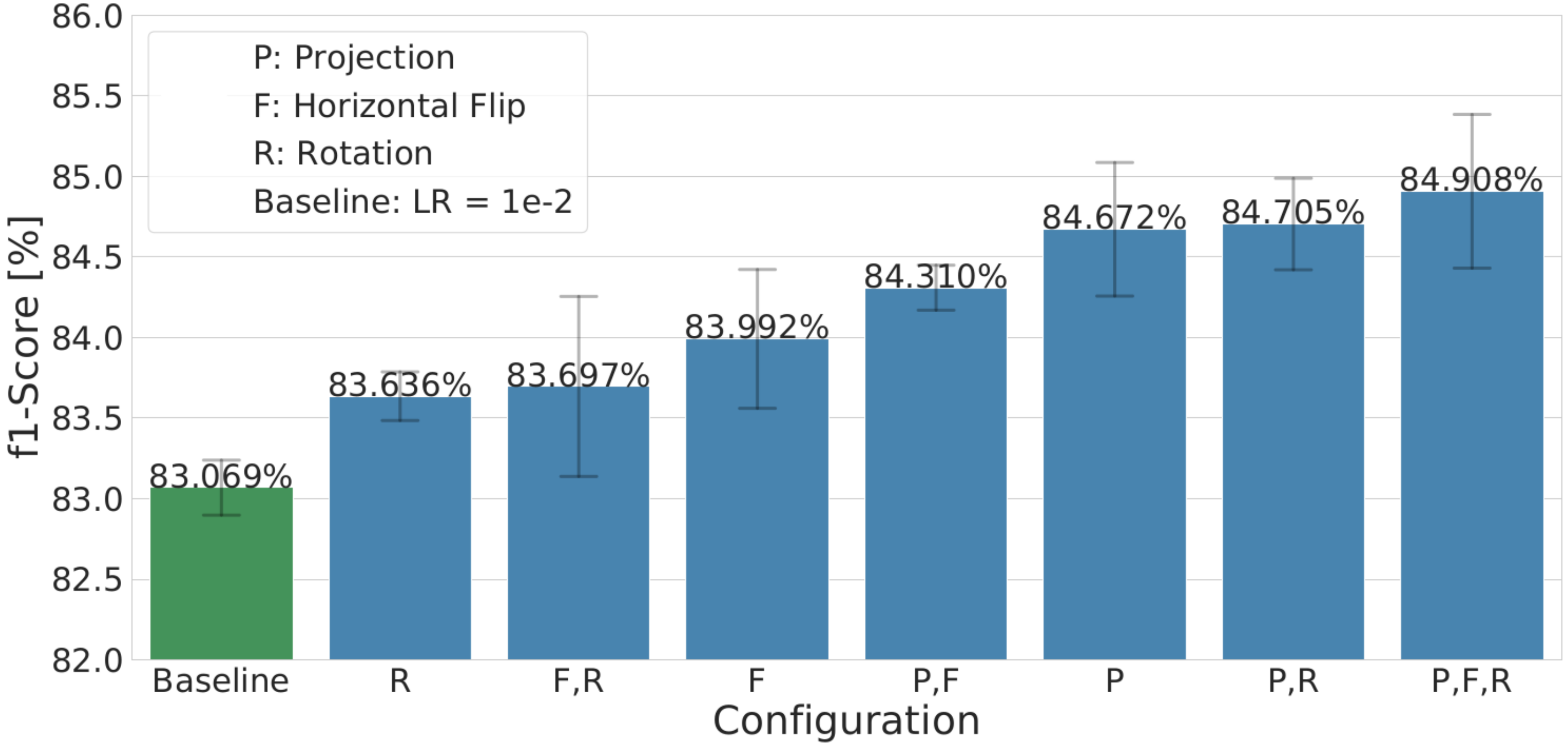
Training: MUnet

Configuration:

Batch Size: 64
Loss: Focal-Loss (α=0.8, γ=2.0)
Optimizer: AMSGrad
Input Size: 448x448x3 (gray scale, weights pretrained)
Metric: f1-score



Training: MUnet Offline Augmentation



Results of the offline augmentation experiment

Training: MUnet Online Augmentation and Grid Search

Offline augmentation: 84.908%

Online augmentation configuration:

- Three parameters per augmentation
- Augmentations: rotation, scale, crop, color jitter
- Crop best with: 85.707%

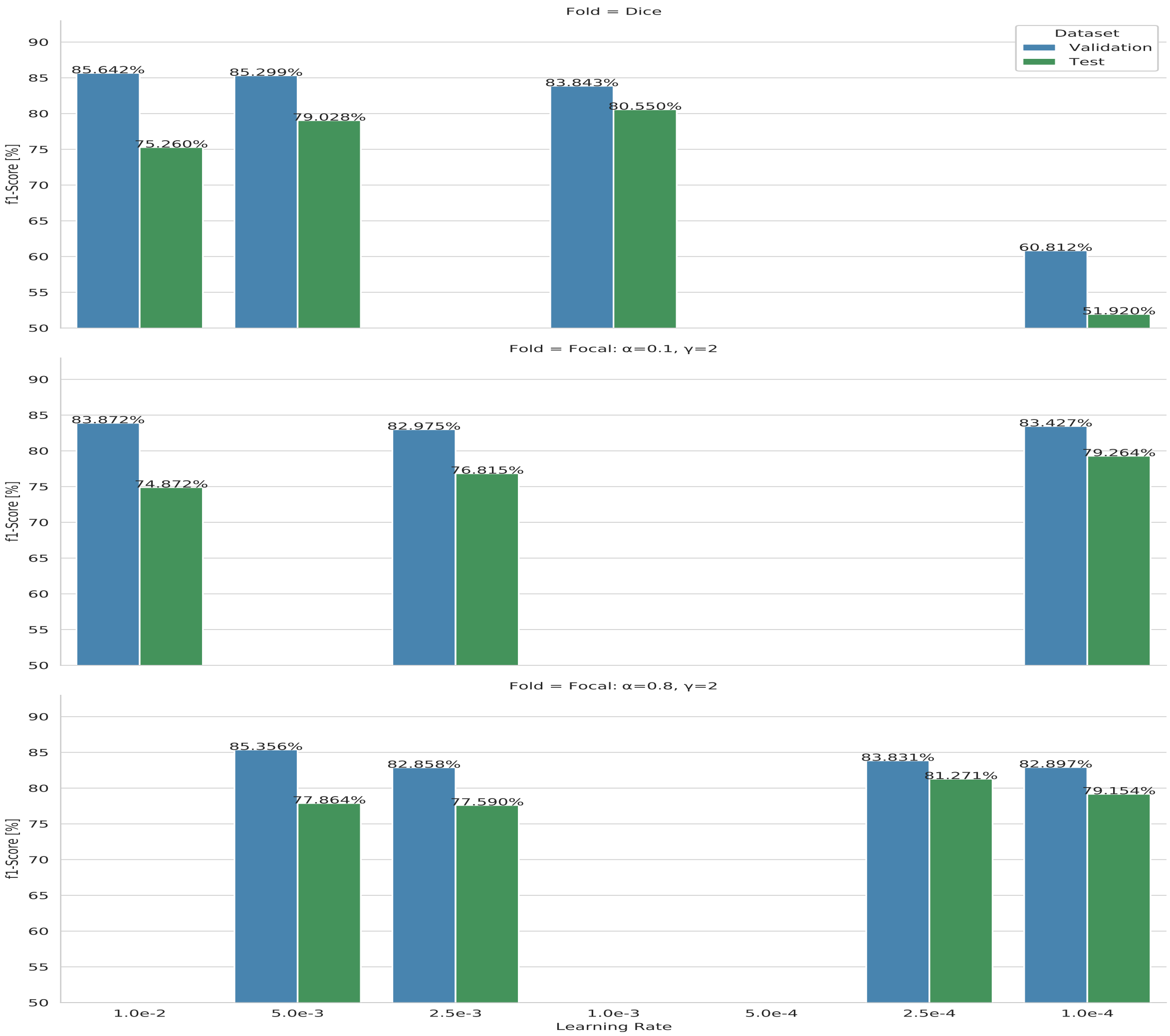
Grid search configuration:

Learning Rates: 1.0e-2, 5.0e-3, 2.5e-3, 1.0e-3, 5.0e-4, 2.5e-4, 1.0e-4
Batch Sizes: 32, 64
Loss: Dice, Focal-Loss (α=0.1, α=0.8)

Results:

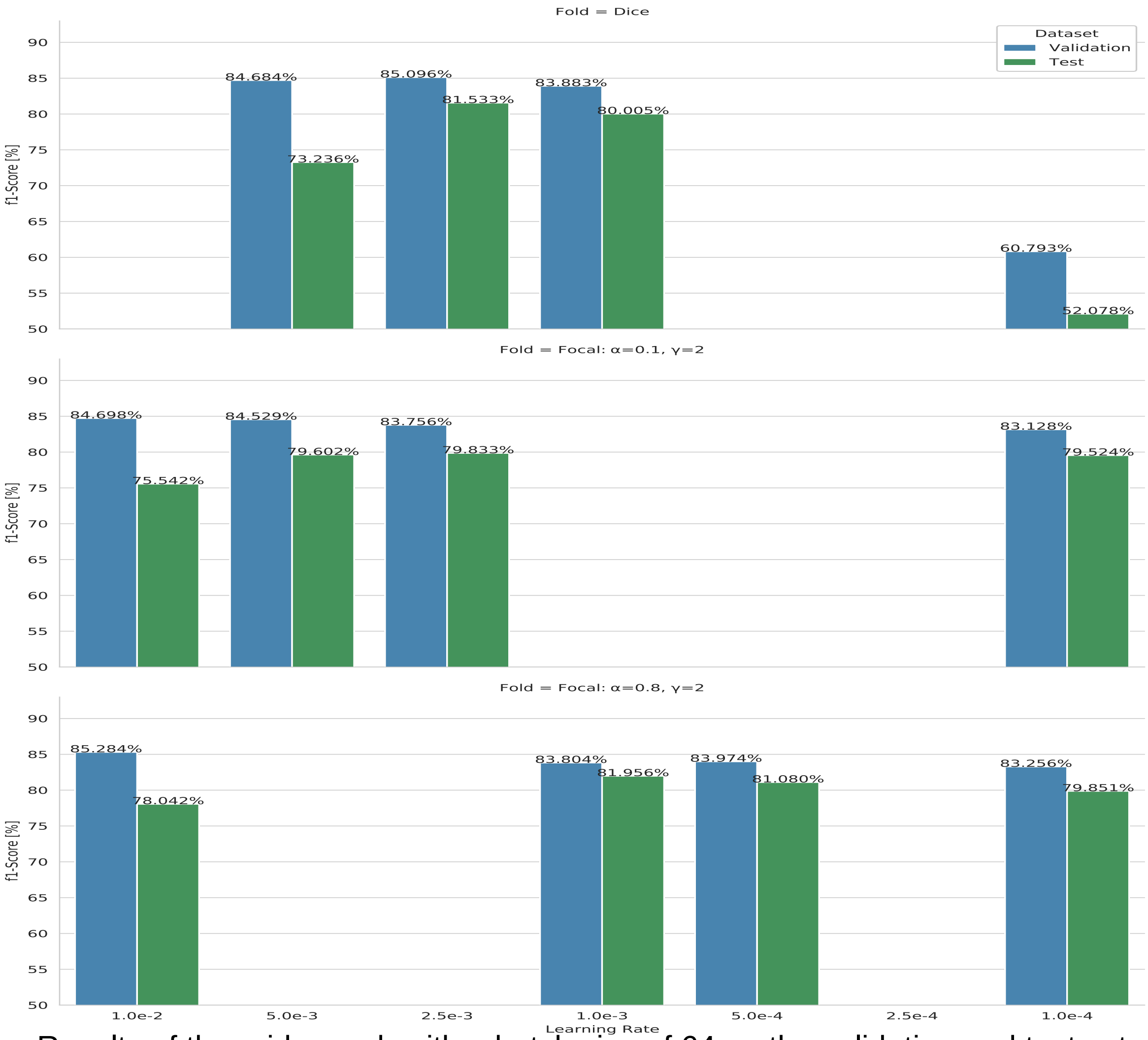
- High learning rates show bad performance on test dataset
- Low learning rates with Dice loss show bad performance

Training: MUnet Grid Search Batch Size 32



Results of the grid search with a batch size of 32 on the validation and test set

Training: MUnet Grid Search Batch Size 64



Results of the grid search with a batch size of 64 on the validation and test set

Evaluation

- Classification Evaluation

- Matching Evaluation

- Topology Evaluation

Pipeline: Overview

Evaluation: Classification

- Precision, recall, f1-score (0.35 IoU obtained from train and validation set as the maximum occlusion threshold)

- State-of-the-art performance in electrical circuit component recognition (prev. f1-score 94.85% [5])

NMS	ScoreThr.	IoUThr.	InputSize	Vote	TP	FP	FN	Precision	Recall	F1
DIoU	0.3	0.25	608 × 608	-	798	**6**	9	**99.25%**	98.88%	**99.07%**
DIoU	0.3	0.25	736 × 736	-	799	8	8	99.01%	99.01%	99.01%
DIoU	0.15	0.45	736 × 736	-	801	12	6	98.52%	99.26%	98.89%
WBF	0.15	0.25	736 × 736	-	801	12	6	98.52%	99.26%	98.89%
WBF-TTA	0.1	0.45	736 × 736	1	**804**	30	**3**	96.40%	**99.63%**	97.99%
WBF-TTA	0.1	0.45	736 × 736	2	**804**	21	**3**	97.45%	**99.63%**	98.53%
WBF-TTA	0.1	0.45	736 × 736	3	**804**	17	**3**	97.93%	**99.63%**	98.77%
WBF-TTA	0.1	0.45	736 × 736	4	**804**	13	**3**	98.41%	**99.63%**	99.01%
WBF-TTA	0.1	0.45	736 × 736	5	803	13	4	98.41%	99.50%	98.95%

Results of the classification

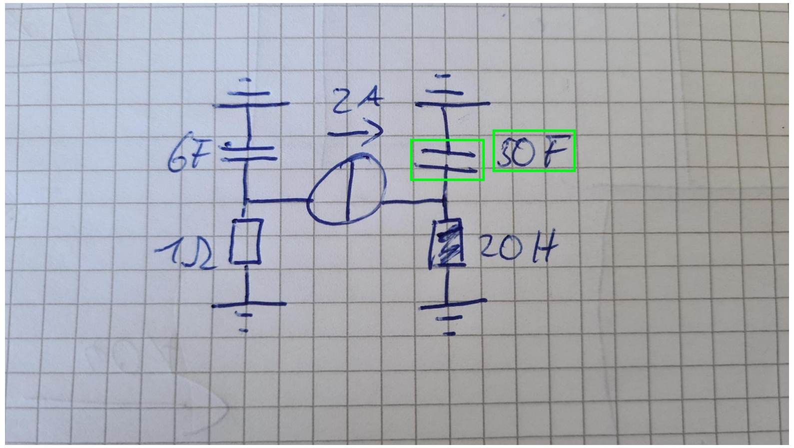
[5]: T. Dhanushika and L. Ranathunga, Fine-Tuned Line Connection Accompanied Boolean Expression Generation for Hand-Drawn Logic Circuits. 2019.

Evaluation: Annotation Matching General

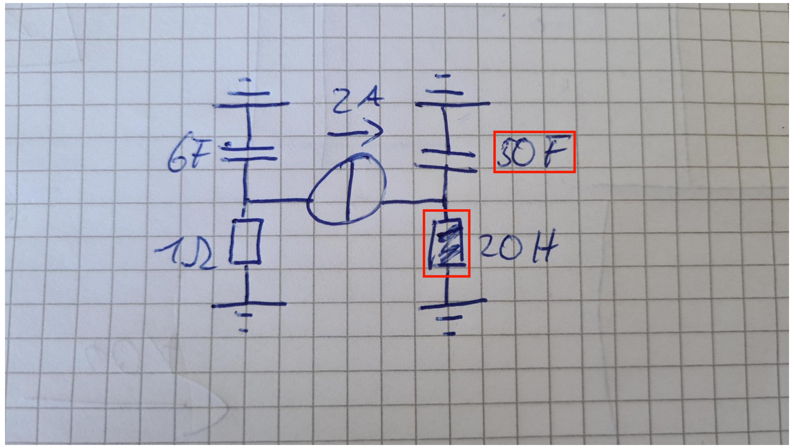
- True Positive (TP): correct match

- False Positive (FP): annotation matched to wrong component

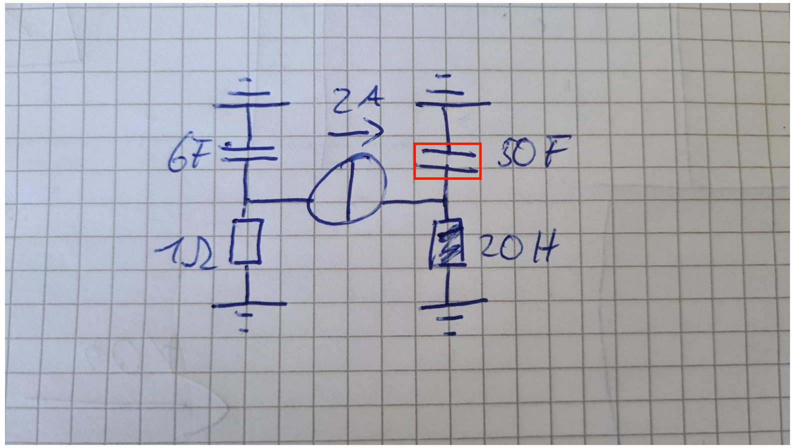
- False Negative (FN): annotation not detected, therefore no match at all (transient error from classification)



TP Case



FP Case



FN Case

Evaluation: Textual Annotation Matching

NMS	Score Thr.	IoU Thr.	InputSize	Votes	TP	FP	FN	Precision	Recall	F1
DIoU	0.3	0.25	608	-	169	9	0	94.94%	100.00%	97.41%
DIoU	0.3	0.25	736	-	169	9	0	94.94%	100.00%	97.41%
DIoU	0.15	0.45	736	-	169	9	0	94.94%	100.00%	97.41%
WBF	0.15	0.25	736	-	169	9	0	94.94%	100.00%	97.41%
WBF-TTA	0.1	0.45	736	1	169	9	0	94.94%	100.00%	97.41%
WBF-TTA	0.1	0.45	736	2	169	9	0	94.94%	100.00%	97.41%
WBF-TTA	0.1	0.45	736	3	169	9	0	94.94%	100.00%	97.41%
WBF-TTA	0.1	0.45	736	4	169	9	0	94.94%	100.00%	97.41%
WBF-TTA	0.1	0.45	736	5	169	9	0	94.94%	100.00%	97.41%

Results of the textual annotation matching

Evaluation: Arrow Annotation Matching

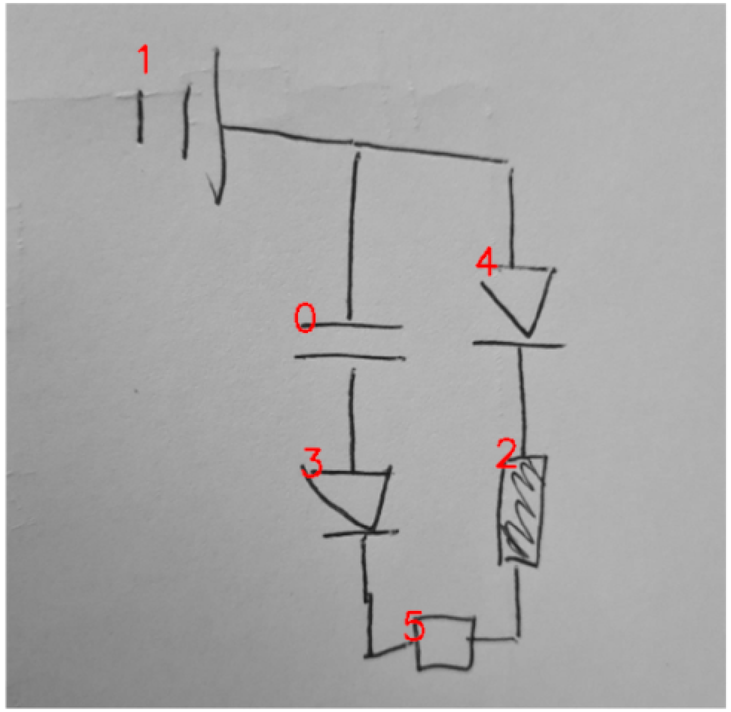
NMS	Score Thr.	IoU Thr.	InputSize	Votes	TP	FP	FN	Precision	Recall	F1
DIoU	0.3	0.25	608	-	40	0	0	100.00%	100.00%	100.00%
DIoU	0.3	0.25	736	-	40	0	0	100.00%	100.00%	100.00%
DIoU	0.15	0.45	736	-	40	0	0	100.00%	100.00%	100.00%
WBF	0.15	0.25	736	-	40	0	0	100.00%	100.00%	100.00%
WBF-TTA	0.1	0.45	736	1	40	0	0	100.00%	100.00%	100.00%
WBF-TTA	0.1	0.45	736	2	40	0	0	100.00%	100.00%	100.00%
WBF-TTA	0.1	0.45	736	3	40	0	0	100.00%	100.00%	100.00%
WBF-TTA	0.1	0.45	736	4	40	0	0	100.00%	100.00%	100.00%
WBF-TTA	0.1	0.45	736	5	40	0	0	100.00%	100.00%	100.00%

Results of the arrow annotation matching

Evaluation: Topology

- Hypergraph (generalization of a graph) [6]

- Represented as an adjacency matrix [7]



Circuit with indeces of bounding box labels

Index	0		1		2		3		4		5		
Edge 0	1	0	0	1	0	0	0	0	1	0	0	0	1r,0t,4t
Edge 1	0	1	0	0	0	0	1	0	0	0	0	0	0b,3t
Edge 2	0	0	0	0	0	0	0	1	0	0	1	0	3b,5l
Edge 3	0	0	0	0	0	1	0	0	0	0	0	1	5r,2b
Edge 4	0	0	0	0	1	0	0	0	0	1	0	0	4b,2t

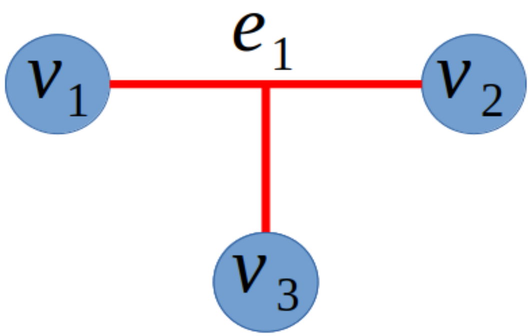
Adjacency matrix of the topology of the circuit

[6]: P. Valdivia et al., Analyzing Dynamic Hypergraphs with Parallel Aggregated Ordered Hypergraph Visualization. 2021.
[7]: E. Konstantinova and V. Skorobogatov, Application of hypergraph theory in chemistry, 2001

Evaluation: Topology TP Case

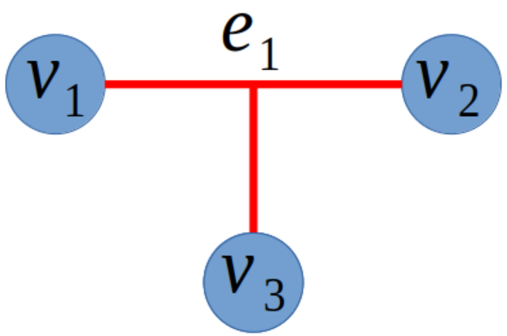
- "A perfect match"
- #(subedges) = #(vertices in ground truth hyperedge) – 1

→ $TP_h = 2$



	v_1	v_2	v_3
e_1	1	1	1

Ground Truth



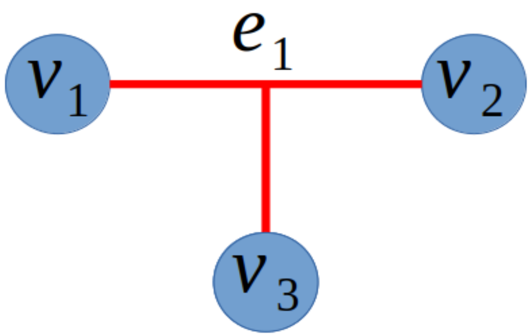
	v_1	v_2	v_3
e_1	1	1	1

Prediction

Evaluation: Topology FN Case

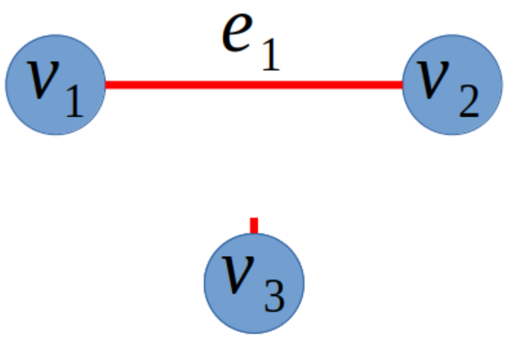
- "A subedge is missing"
- Perfect match can only be obtained through recombination

→ $TP_h = 1, FN_h = 1$



	v_1	v_2	v_3
e_1	1	1	1

Ground Truth



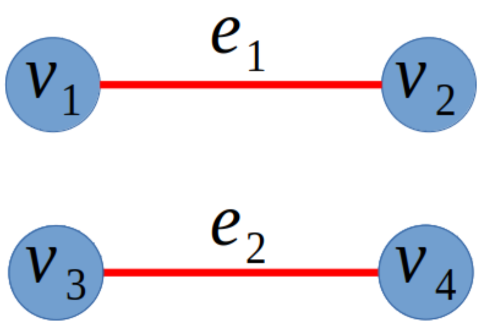
	v_1	v_2	v_3
e_1	1	1	0
e_2	0	0	1

Prediction

Evaluation: Topology FP Case

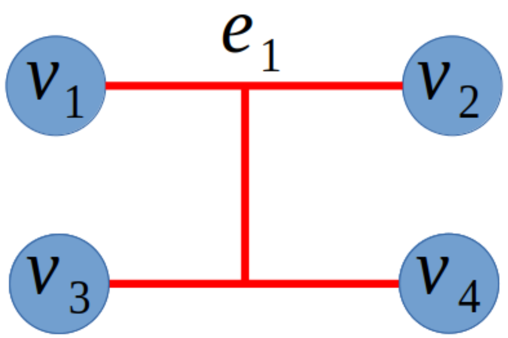
- “A subedge too much”
- Perfect match is obtained through splitting

→ $TP_h=2, FP_h=1$



	v_1	v_2	v_3	v_4
e_1	1	1	0	0
e_2	0	0	1	1

Ground Truth



	v_1	v_2	v_3	v_4
e_1	1	1	1	1

Prediction

Evaluation: Topology Algorithm

Algorithm (inspired by graph edit distance [8]):

1. Find perfect matches

2. Find combination matches

3. Find split matches

4. Do until no edges left

Drawbacks:

- Mixed cases of FP, FN not handled (worsed error assumed in such cases)

- Transient errors from classification not considered

[8]: X. Chen et al., An efficient algorithm for graph edit distance computation. 2018.

Evaluation: Topology Results

Batch Size	Loss	Learning Rate	$\mathbf{TP_h}$	$\mathbf{FP_h}$	$\mathbf{FN_h}$	Precision	Recall	F1
Ground Truth			620	0	47	100.0%	92.95%	96.35%
32	focal $\alpha = 0.1$	$1.0e^{-2}$	482	7	185	98.57%	72.26%	83.39%
32	focal $\alpha = 0.1$	$2.5e^{-3}$	488	**4**	179	99.19%	73.16%	84.21%
32	focal $\alpha = 0.1$	$1.0e^{-4}$	504	6	163	98.82%	75.56%	85.64%
64	focal $\alpha = 0.1$	$1.0e^{-2}$	464	7	203	98.51%	69.57%	81.55%
64	focal $\alpha = 0.1$	$5.0e^{-3}$	492	6	175	98.80%	73.76%	84.46%
64	focal $\alpha = 0.1$	$2.5e^{-3}$	491	**4**	176	99.19%	73.61%	84.51%
64	focal $\alpha = 0.1$	$1.0e^{-4}$	505	7	162	98.63%	75.71%	85.67%
32	focal $\alpha = 0.8$	$5.0e^{-3}$	485	5	182	98.98%	72.71%	83.84%
32	focal $\alpha = 0.8$	$2.5e^{-3}$	469	5	198	98.95%	70.31%	82.21%
32	focal $\alpha = 0.8$	$2.5e^{-4}$	498	5	169	99.01%	74.66%	85.13%
32	focal $\alpha = 0.8$	$1.0e^{-4}$	484	5	183	98.98%	72.56%	83.74%
64	focal $\alpha = 0.8$	$1.0e^{-2}$	492	5	175	98.99%	73.76%	84.54%
64	focal $\alpha = 0.8$	$1.0e^{-3}$	499	5	168	99.01%	74.81%	85.23%
64	focal $\alpha = 0.8$	$5.0e^{-4}$	511	**4**	156	99.22%	76.61%	86.46%
64	focal $\alpha = 0.8$	$1.0e^{-4}$	496	5	171	99.00%	74.36%	84.93%
32	dice	$1.0e^{-2}$	449	5	218	98.90%	67.32%	80.11%
32	dice	$5.0e^{-3}$	468	6	199	98.73%	70.16%	82.03%
32	dice	$1.0e^{-3}$	**531**	**4**	**136**	**99.25%**	**79.61%**	**88.35%**
32	dice	$1.0e^{-4}$	183	6	484	96.83%	27.44%	42.76%
64	dice	$5.0e^{-3}$	457	6	210	98.70%	68.52%	80.88%
64	dice	$2.5e^{-3}$	521	5	146	99.05%	78.11%	87.34%
64	dice	$1.0e^{-3}$	512	5	155	99.03%	76.76%	86.49%
64	dice	$1.0e^{-4}$	185	6	482	96.86%	27.74%	43.12%

Results of the topology evaluation

Summary & Outlook

Dataset

- Bounding Box Labels
- Segmentation Labels
- Topology Labels (only test dataset)
- Annotation Labels (only test dataset)

Pipeline

- Detects components and annotations
- Operates on various paper backgrounds (has not been done so far)
- OCR (not in the scope of this thesis)
- Generate topology
- Convert into LTspice format

Evaluation

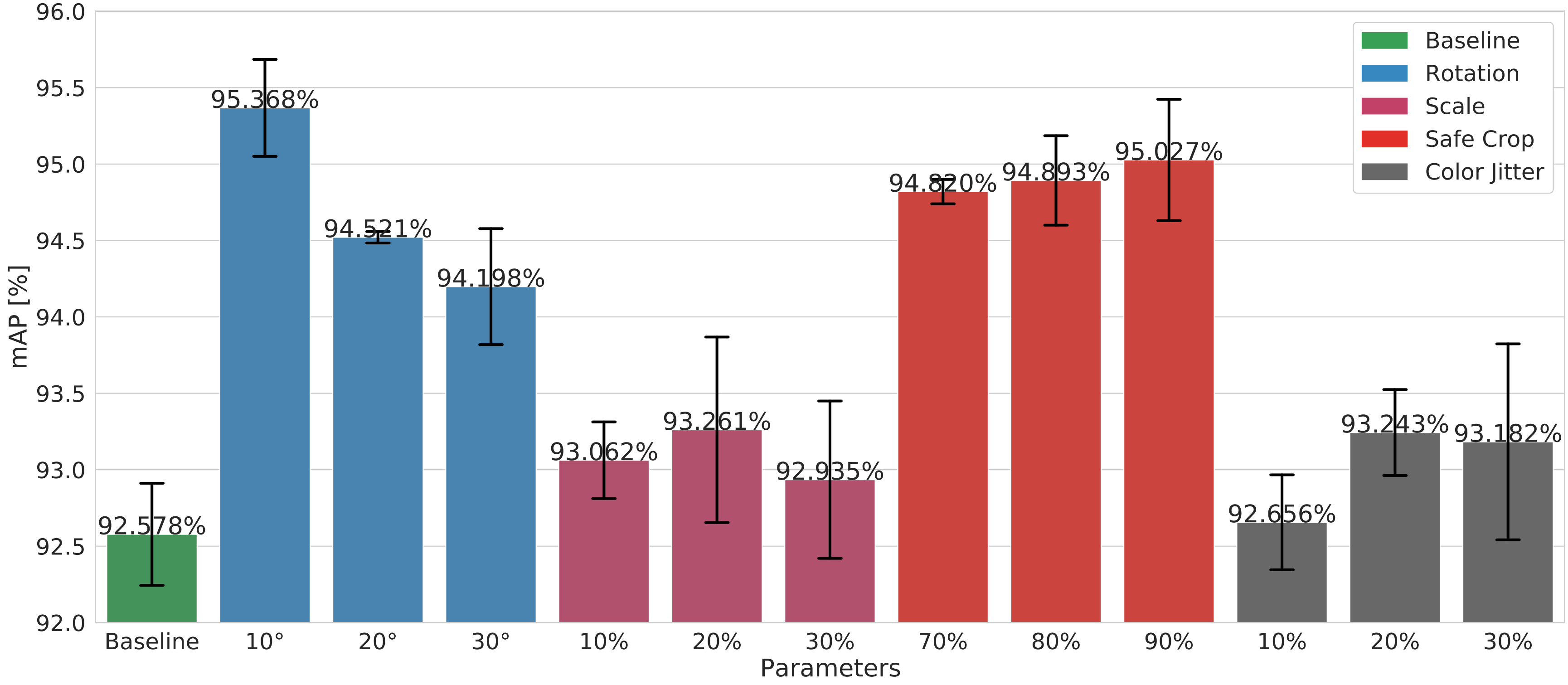
- Classification, annotation matching and topology evaluation
- Topology evaluation not fully functional

Results

- State-of-the-art performance in electrical circuit component classification (YOLO)

Questions?

Training: YOLO Online Augmentation, Grid Search and Fine-Tuning

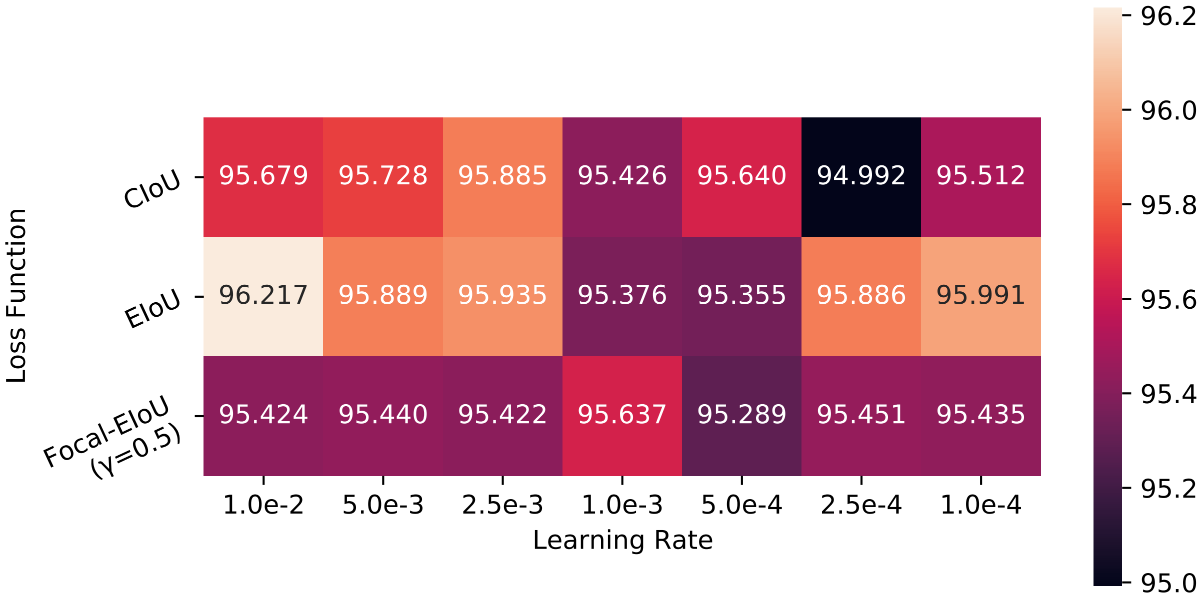


Results of the online augmentation experiment

Training: YOLO Grid Search

Configuration:

Learning Rates: 1.0e-2, 5.0e-3, 2.5e-3, 1.0e-3, 5.0e-4, 2.5e-4, 1.0e-4
Batch Sizes: 32, 64
Loss: CIoU [3], EIoU [4], Focal-EIoU [4]



Results of the hyperparameter grid search experiment for batch size 64

[3]: Z. Zheng et al., Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. 2019.
[4]: Y. Zhang et al, Focal and Efficient IOU Loss for Accurate Bounding Box Regression, 2021.