

Document Classification

Team Members: Hovsep Avagyan
Martun Karapetyan
Arman Deghoyan

Supervisor: Khachatur

RVL-CDIP Dataset

letter

memo

email

filefolder

form

handwritten

invoice

advertisement

RJR Interface Memo

MEMORANDUM FOR: [Name] DATE: April 26, 1982

FROM: [Name]

SUBJECT: [Subject]

The following specification changes were made to [Product]:

Item	Description
1.000	Change [Specification]
2.000	Change [Specification]
3.000	Change [Specification]
4.000	Change [Specification]
5.000	Change [Specification]
6.000	Change [Specification]
7.000	Change [Specification]
8.000	Change [Specification]
9.000	Change [Specification]
10.000	Change [Specification]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

MEMORANDUM

TO: [Name]

FROM: [Name]

SUBJECT: [Subject]

[Text]

Camel

Camel cigarettes are...
[Text]

budget

news article

presentation

scientific publication

questionnaire

resume

scientific report

specification

BUDGET

[Table with columns: Item, Description, Amount]

News Article

[Text]

Presentation

[Text]

Scientific Publication

[Text]

Questionnaire

[Form with questions and checkboxes]

Resume

[Form with fields for personal information and work history]

Scientific Report

[Table with data and text]

Specification

[Form with fields for product specifications]

Document Image Classification on RVL-CDIP

Leaderboard

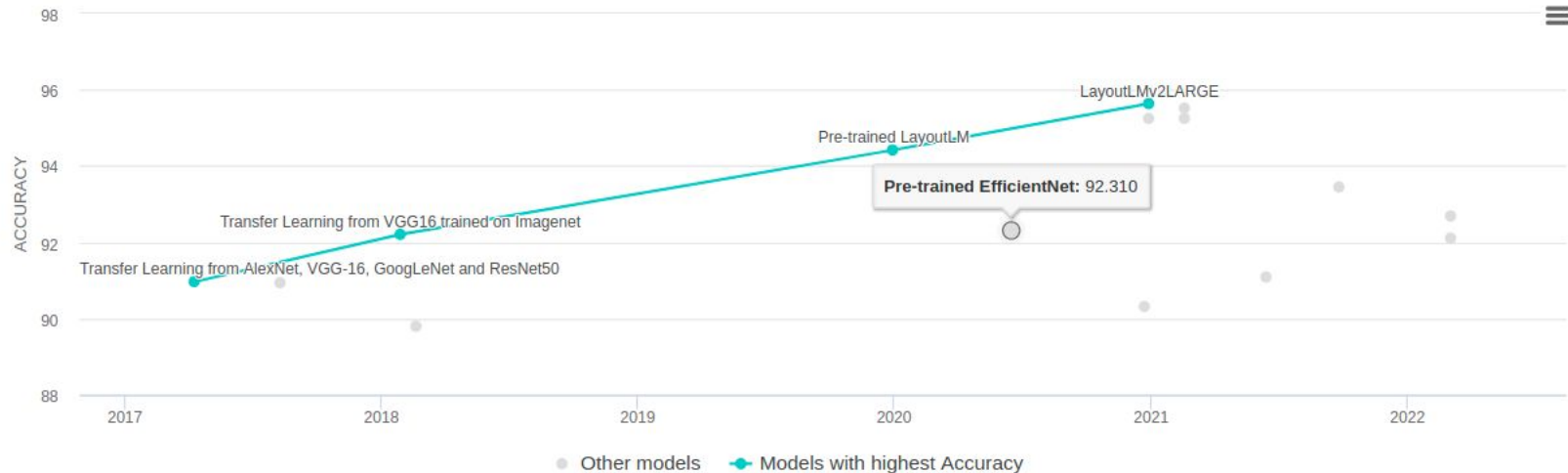
Dataset

View

Accuracy

by

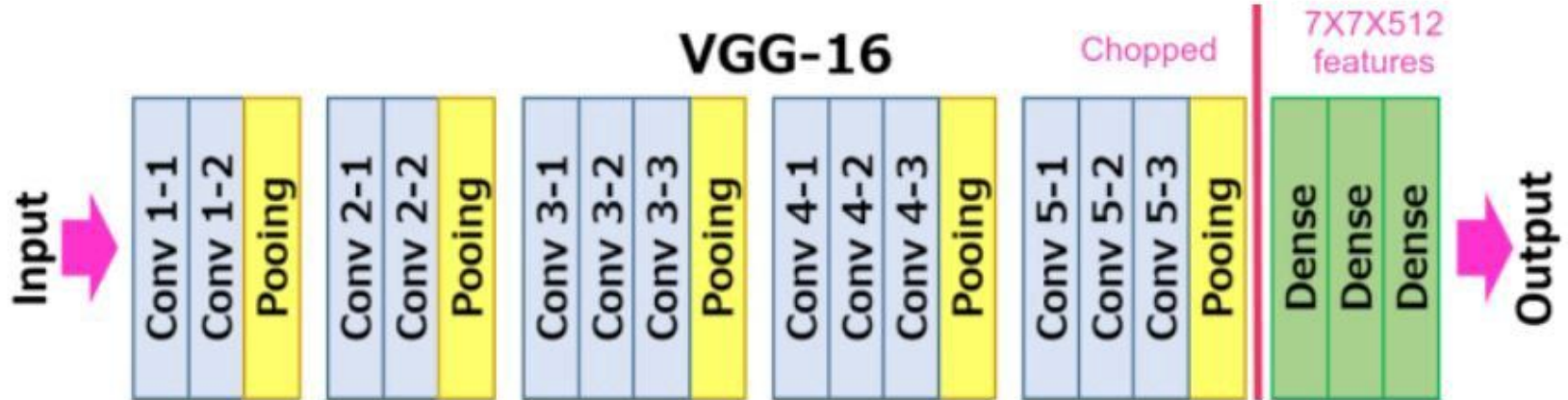
Date



Our Models

Model Description	Validation Accuracy	Model size	# of epochs
VGG with changed classifier	87.8%	~138 million params	12
VGG16 + Fasttext	92.27%	~138 million params	21
Layout_MV2	80%	~200 million params	1/3
OCR with NLP <ol style="list-style-type: none"> Countvectorizer with bigrams + logistic regression Fasttext with charngrams 5,7 	<ol style="list-style-type: none"> 0.796 0.735 	-	7 epoch for Fasttext

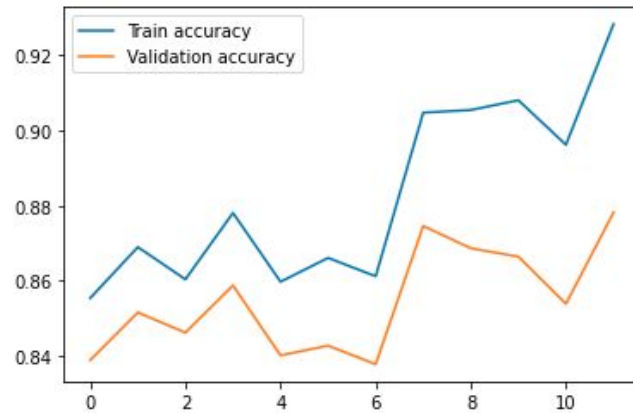
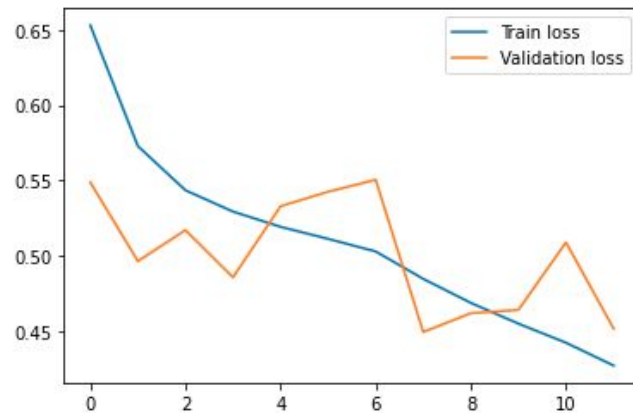
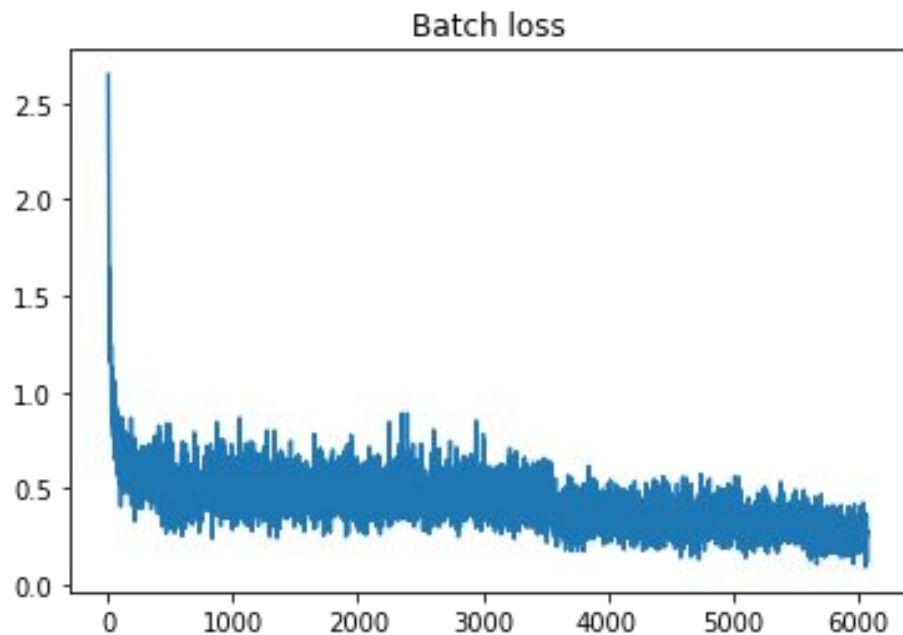
VGG with changed classifier



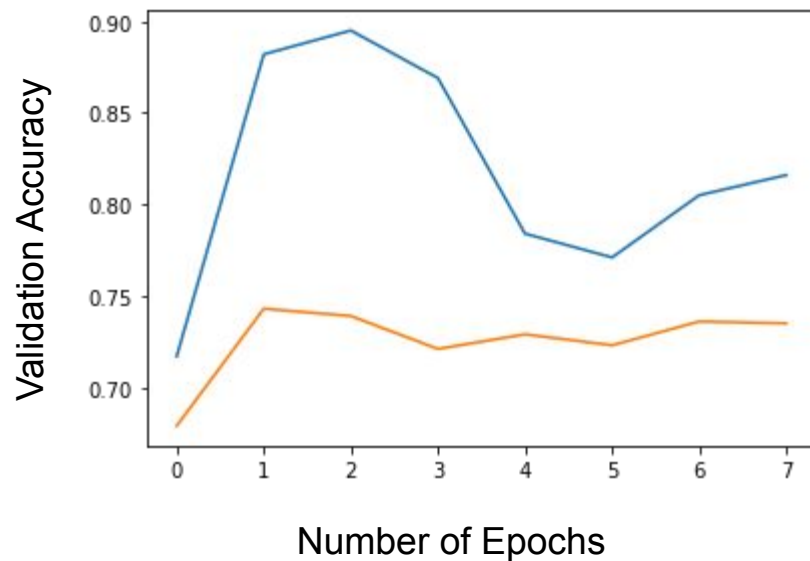
Training Parameters

```
num_epochs = 12
Batch_size = 128
criterion = nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(vgg16.parameters(), lr=0.0001,
weight_decay=0.015)
scheduler = ReduceLROnPlateau(optimizer, 'min', patience=2, factor=0.3)
```

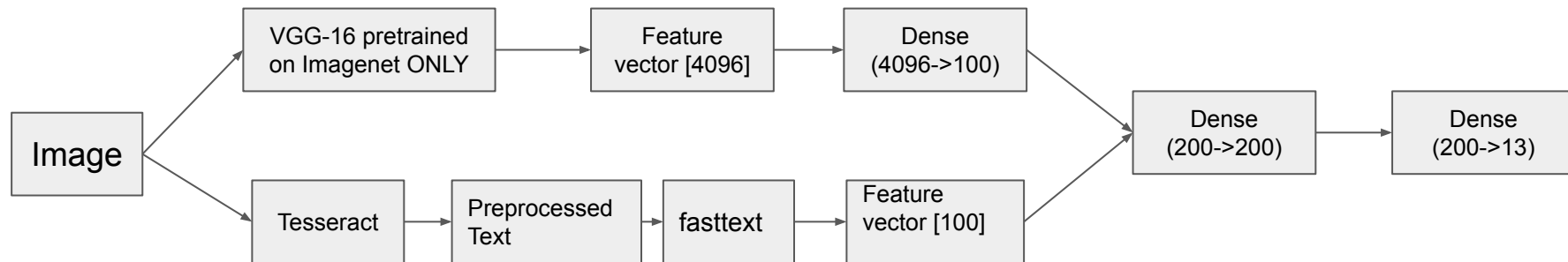
Losses and accuracies



Fasttext with charngrams 5,7



VGG16 + Fasttext



Sample preprocessed text

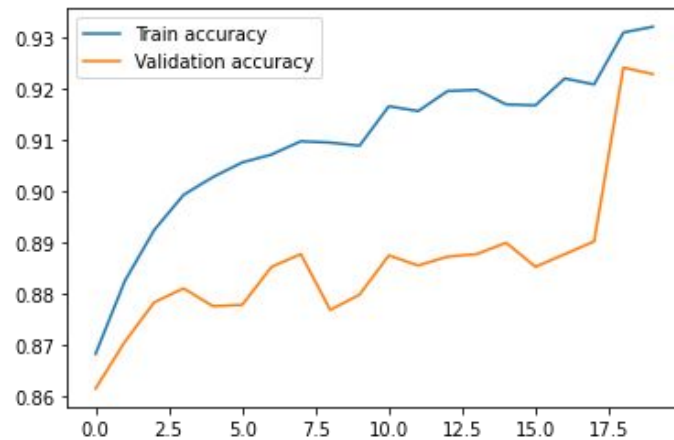
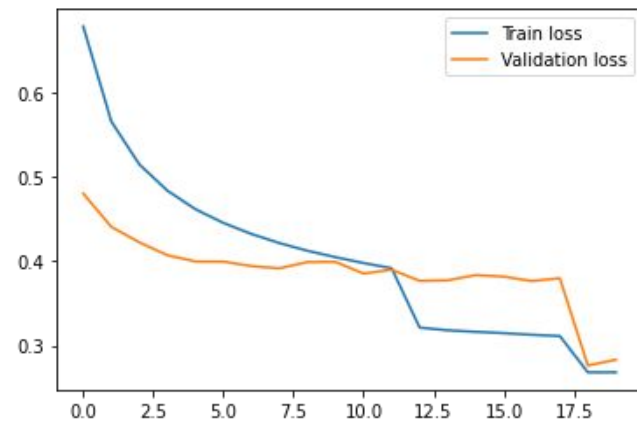
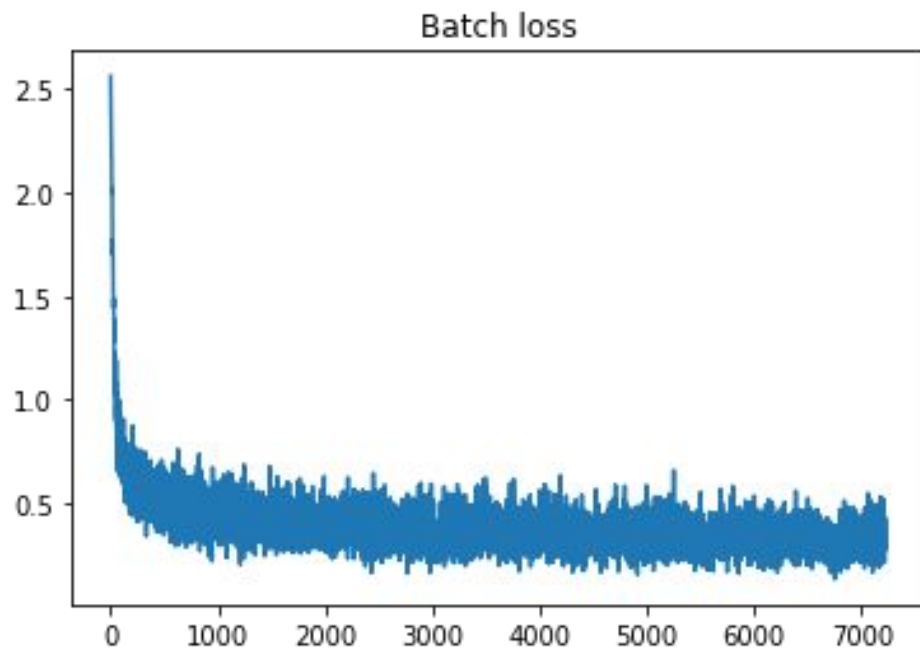
text	image_path	label
late wi lit titer rox set coma cot teeth sa ta...	./train/publication/50729.jpg	7
ate arca ad meter internalization fibroblast g...	./train/publication/8101.jpg	7
research induction resistant y irradiation eff...	./train/publication/88706.jpg	7
tha ae bound compos orc meant proven ate orc m...	./train/publication/45265.jpg	7
attest correlate os te solvent ex te prompt si...	./train/publication/67072.jpg	7

Training Parameters

```
num_epochs = 600
Batch_size = 8*1024
criterion = nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(MyModel.parameters(), lr=0.00005,
weight_decay=0.022)
scheduler = ReduceLROnPlateau(optimizer, 'min', patience=2, factor=0.3)
```

Scheduler reduced learning rate 2 times

Losses and accuracies



Layout_MV2

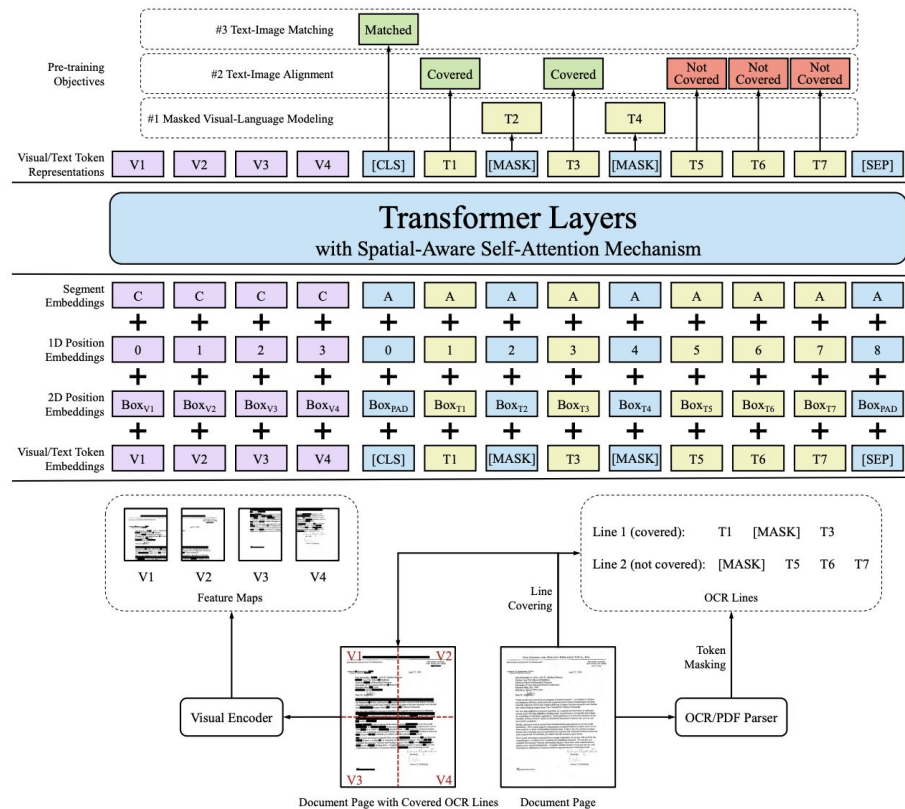


Figure 2: An illustration of the model architecture and pre-training strategies for LayoutLMv2

Training Parameters

```
# Taken from the paper
```

```
optimizer = torch.optim.Adam(LayoutMV2Model.parameters(), lr=0.00002,  
weight_decay=0.01)  
scheduler = ReduceLROnPlateau(optimizer, 'min', patience=2, factor=0.3)
```

Losses and accuracies

No charts for $\frac{1}{3}$ epochs, sorry. Accuracy 80%

Inference Speed

	min	mean	[+/-sd]	median	max
FastText	1391	2705	290.1	2768	2790
VGG16 + Fasttext	1706	3350	291.0	3403	3448
VGG	303	621	67.8	612	1476

Thanks!!!