U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

UNIVERSITY OF TWENTE. | TECHMED CENTRE

LEARNING TO WRITE MEDICAL REPORTS FROM EEG DATA

Automatic report generation

Author: Ana Maria Sousa
Supervised by: Luís Filipe Teixeira
Michel Van Putten
Catarina Lourenço

What is Epilepsy?

Neurological disorder

Over 70 Million people Worldwide

Increased likelihood of unprovoked seizures

Motivation

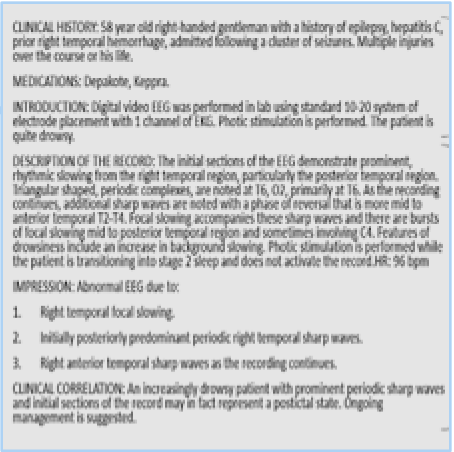
Diagnosis is
difficult/complex

Time consuming

Subjectivity

EEG recording

EEG analysis and
report writing



CLINICAL HISTORY: 58 year old right-handed gentleman with a history of epilepsy, hepatitis C, prior right temporal hemorrhage, admitted following a cluster of seizures. Multiple injuries over the course or his life.

MEDICATIONS: Depakote, Keppra.

INTRODUCTION: Digital video EEG was performed in lab using standard 10-20 system of electrode placement with 1 channel of EKG. Photic stimulation is performed. The patient is quite drowsy.

DESCRIPTION OF THE RECORD: The initial sections of the EEG demonstrate prominent, rhythmic slowing from the right temporal region, particularly the posterior temporal region. Triangular shaped, periodic complexes, are noted at T6, O2, primarily at T6. As the recording continues, additional sharp waves are noted with a phase of reversal that is more mid to anterior temporal T2-T4. Focal slowing accompanies these sharp waves and there are bursts of focal slowing mid to posterior temporal region and sometimes involving C4. Features of drowsiness include an increase in background slowing. Photic stimulation is performed while the patient is transitioning into stage 2 sleep and does not activate the record.HR: 96 bpm

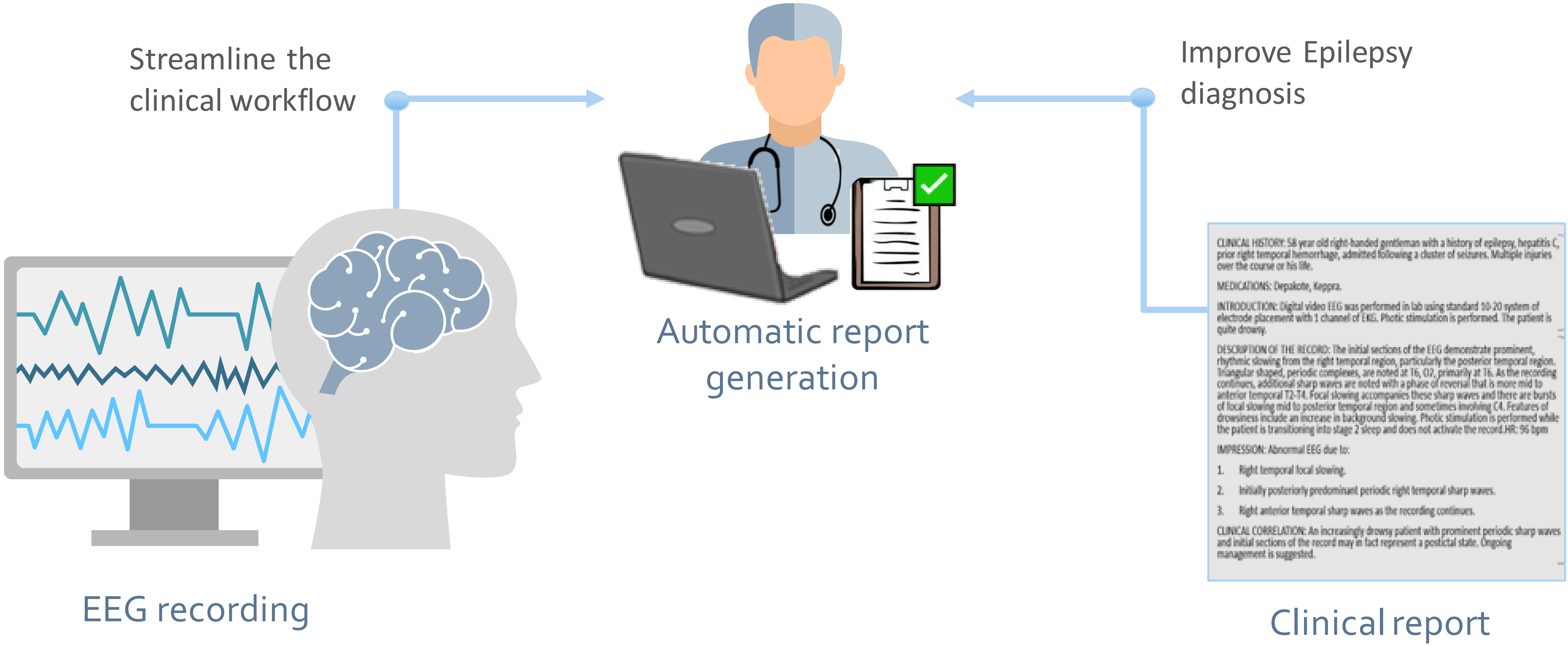
IMPRESSION: Abnormal EEG due to:

1. Right temporal focal slowing.
2. Initially posteriorly predominant periodic right temporal sharp waves.
3. Right anterior temporal sharp waves as the recording continues.

CLINICAL CORRELATION: An increasingly drowsy patient with prominent periodic sharp waves and initial sections of the record may in fact represent a postictal state. Ongoing management is suggested.

Clinical report

What if we could "translate" what the brain is saying?



CLINICAL HISTORY: 58 year old right-handed gentleman with a history of epilepsy, hepatitis C, prior right temporal hemorrhage, admitted following a cluster of seizures. Multiple injuries over the course or his life.

MEDICATIONS: Depakote, Keppra.

INTRODUCTION: Digital video EEG was performed in lab using standard 10-20 system of electrode placement with 1 channel of EKG. Photic stimulation is performed. The patient is quite drowsy.

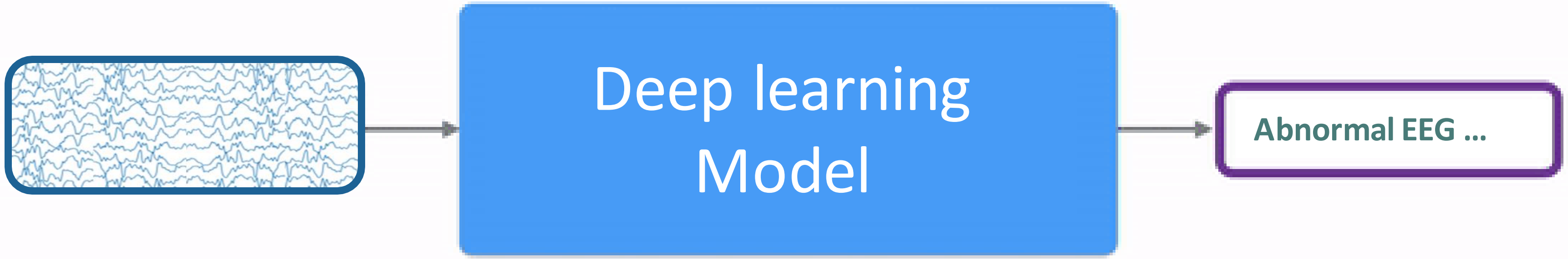
DESCRIPTION OF THE RECORD: The initial sections of the EEG demonstrate prominent, rhythmic slowing from the right temporal region, particularly the posterior temporal region. Triangular shaped, periodic complexes, are noted at T6, O2, primarily at T6. As the recording continues, additional sharp waves are noted with a phase of reversal that is more mid to anterior temporal T2-T4. Focal slowing accompanies these sharp waves and there are bursts of focal slowing mid to posterior temporal region and sometimes involving C4. Features of drowsiness include an increase in background slowing. Photic stimulation is performed while the patient is transitioning into stage 2 sleep and does not activate the record.HR: 96 bpm

IMPRESSION: Abnormal EEG due to:

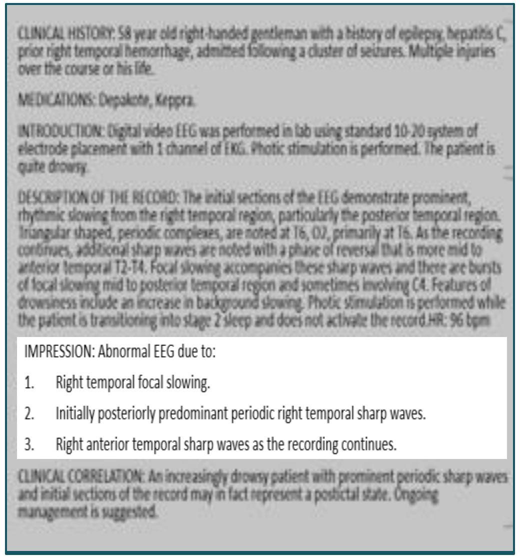
1. Right temporal focal slowing.
2. Initially posteriorly predominant periodic right temporal sharp waves.
3. Right anterior temporal sharp waves as the recording continues.

CLINICAL CORRELATION: An increasingly drowsy patient with prominent periodic sharp waves and initial sections of the record may in fact represent a postictal state. Ongoing management is suggested.

What if we could "translate" what the brain is saying?



Data Preparation


CLINICAL HISTORY: 58 year old right-handed gentleman with a history of epilepsy, hepatitis C, prior right temporal hemorrhage, admitted following a cluster of seizures. Multiple injuries over the course or his life.

MEDICATIONS: Depakote, Keppra.

INTRODUCTION: Digital video EEG was performed in lab using standard 10-20 system of electrode placement with 1 channel of EKG. Photic stimulation is performed. The patient is quite drowsy.

DESCRIPTION OF THE RECORD: The initial sections of the EEG demonstrate prominent, rhythmic slowing from the right temporal region, particularly the posterior temporal region. Triangular shaped, periodic complexes, are noted at T6, O2, primarily at T6. As the recording continues, additional sharp waves are noted with a phase of reversal that is more mid to anterior temporal T2-T4. Focal slowing accompanies these sharp waves and there are bursts of focal slowing mid to posterior temporal region and sometimes involving C4. Features of drowsiness include an increase in background slowing. Photic stimulation is performed while the patient is transitioning into stage 2 sleep and does not activate the record.HR: 96 bpm

IMPRESSION: Abnormal EEG due to:

1. Right temporal focal slowing.
2. Initially posteriorly predominant periodic right temporal sharp waves.
3. Right anterior temporal sharp waves as the recording continues.

CLINICAL CORRELATION: An increasingly drowsy patient with prominent periodic sharp waves and initial sections of the record may in fact represent a postictal state. Ongoing management is suggested.

abnormal eeg due to right temporal focal slowing and predominance of sharp waves, initially in the right posterior temporal region and continuing in the right anterior temporal region.

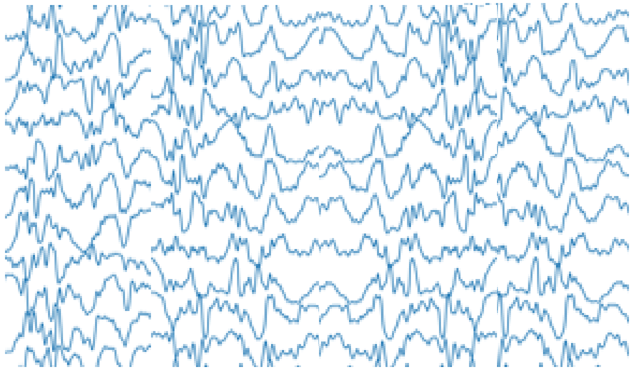
Clean and normalize text

<START>
abnormal
eeg
due
to
right
temporal
focal
....
<END>

Tokenization

Vocabulary creation

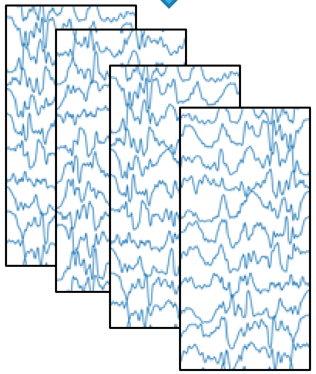
1
5
56
28
15
48
18
23
....
2



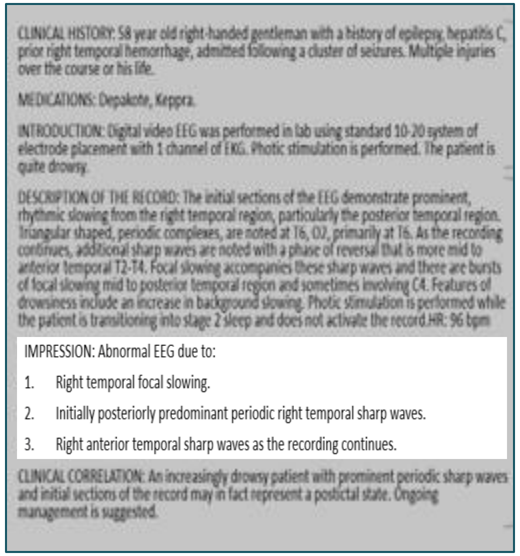
Downsampling

Filtering [0.5 , 35] Hz

Split in 2s epochs



Data Preparation



CLINICAL HISTORY: 58 year old right-handed gentleman with a history of epilepsy, hepatitis C, prior right temporal hemorrhage, admitted following a cluster of seizures. Multiple injuries over the course or his life.

MEDICATIONS: Depakote, Keppra.

INTRODUCTION: Digital video EEG was performed in lab using standard 10-20 system of electrode placement with 1 channel of EKG. Photic stimulation is performed. The patient is quite drowsy.

DESCRIPTION OF THE RECORD: The initial sections of the EEG demonstrate prominent, rhythmic slowing from the right temporal region, particularly the posterior temporal region. Triangular shaped, periodic complexes, are noted at T6, O2, primarily at T6. As the recording continues, additional sharp waves are noted with a phase of reversal that is more mid to anterior temporal T2-T4. Focal slowing accompanies these sharp waves and there are bursts of focal slowing mid to posterior temporal region and sometimes involving C4. Features of drowsiness include an increase in background slowing. Photic stimulation is performed while the patient is transitioning into stage 2 sleep and does not activate the record.HR: 96 bpm

IMPRESSION: Abnormal EEG due to:

1. Right temporal focal slowing.
2. Initially posteriorly predominant periodic right temporal sharp waves.
3. Right anterior temporal sharp waves as the recording continues.

CLINICAL CORRELATION: An increasingly drowsy patient with prominent periodic sharp waves and initial sections of the record may in fact represent a postictal state. Ongoing management is suggested.

abnormal eeg due to right temporal focal slowing and predominance of sharp waves, initially in the right posterior temporal region and continuing in the right anterior temporal region.

Clean and normalize text

<START>
abnormal
eeg
due
to
right
temporal
focal
....
<END>

Tokenization

Vocabulary creation

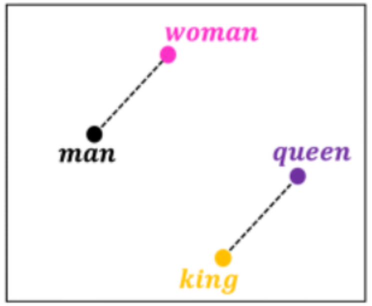
1
5
56
28
15
48
18
23
....
2

One-hot encoding

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 ...
0 0 0 0 **1** 0 0 0 0 0 0 0 0 0 0 ...
0 0 0 0 0 0 0 0 .... 0 0 0 **1** ....
0 0 0 0 0 0 0 0 ... **1** 0 0 0 ...
...
0 0 **1** 0 0 0 0 0 0 0 0 0 0 0 0 ...

Word embedding

-0,2
0,5
0,33
-0,8
0,75
-0,61



Word2Vec
fastText

Dataset

Clinical EEG database of Temple University Hospital (TUH EEG Corpus)

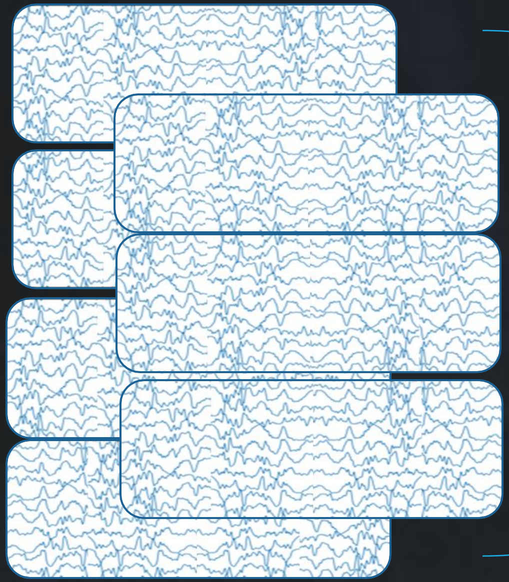
- Patient gender:
  51% Female
  49% Male
- Patient age :
  1 - 90 years old
  average 51.6 years

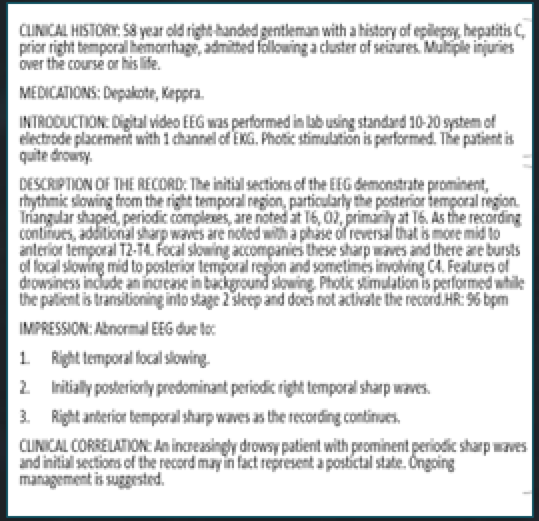
Table 1. TUH EEG Seizure Corpus

Set	Total AR Sessions	Sessions with Impression section	Patients	Epochs (2s)	Times (hours)
Train	687	641	304	904640	251
Test	182	167	36	263307	146
Total	869	808	340	1167947	398

Challenges & Limitations

- EEG signals are extremely variable and affected by noise.





CLINICAL HISTORY: 58 year old right-handed gentleman with a history of epilepsy, hepatitis C, prior right temporal hemorrhage, admitted following a cluster of seizures. Multiple injuries over the course or his life.

MEDICATIONS: Depakote, Keppra.

INTRODUCTION: Digital video EEG was performed in lab using standard 10-20 system of electrode placement with 1 channel of EKG. Photic stimulation is performed. The patient is quite drowsy.

DESCRIPTION OF THE RECORD: The initial sections of the EEG demonstrate prominent, rhythmic slowing from the right temporal region, particularly the posterior temporal region. Triangular shaped, periodic complexes, are noted at T6, O2, primarily at T6. As the recording continues, additional sharp waves are noted with a phase of reversal that is more mid to anterior temporal T2-T4. Focal slowing accompanies these sharp waves and there are bursts of focal slowing mid to posterior temporal region and sometimes involving C4. Features of drowsiness include an increase in background slowing. Photic stimulation is performed while the patient is transitioning into stage 2 sleep and does not activate the record.HR: 96 bpm

IMPRESSION: Abnormal EEG due to:

1. Right temporal focal slowing.
2. Initially posteriorly predominant periodic right temporal sharp waves.
3. Right anterior temporal sharp waves as the recording continues.

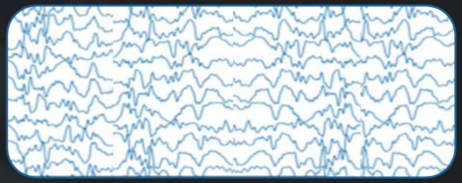
CLINICAL CORRELATION: An increasingly drowsy patient with prominent periodic sharp waves and initial sections of the record may in fact represent a postictal state. Ongoing management is suggested.

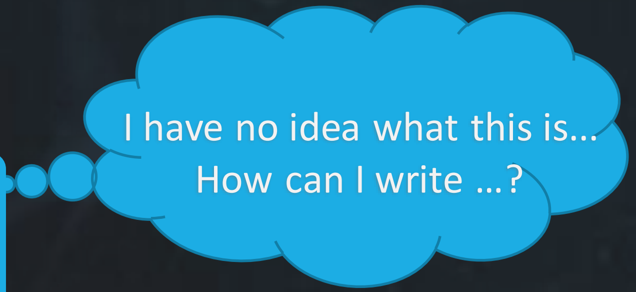
- The content, the **detail** of the report and the **style of writing** are very characteristic of each doctor

- EEG signals are time series with different lengths

Challenges & Limitations

- Deep learning models need a lot of data to learn efficiently





Train: 641 reports (80%)
Test: 167 reports (20%)

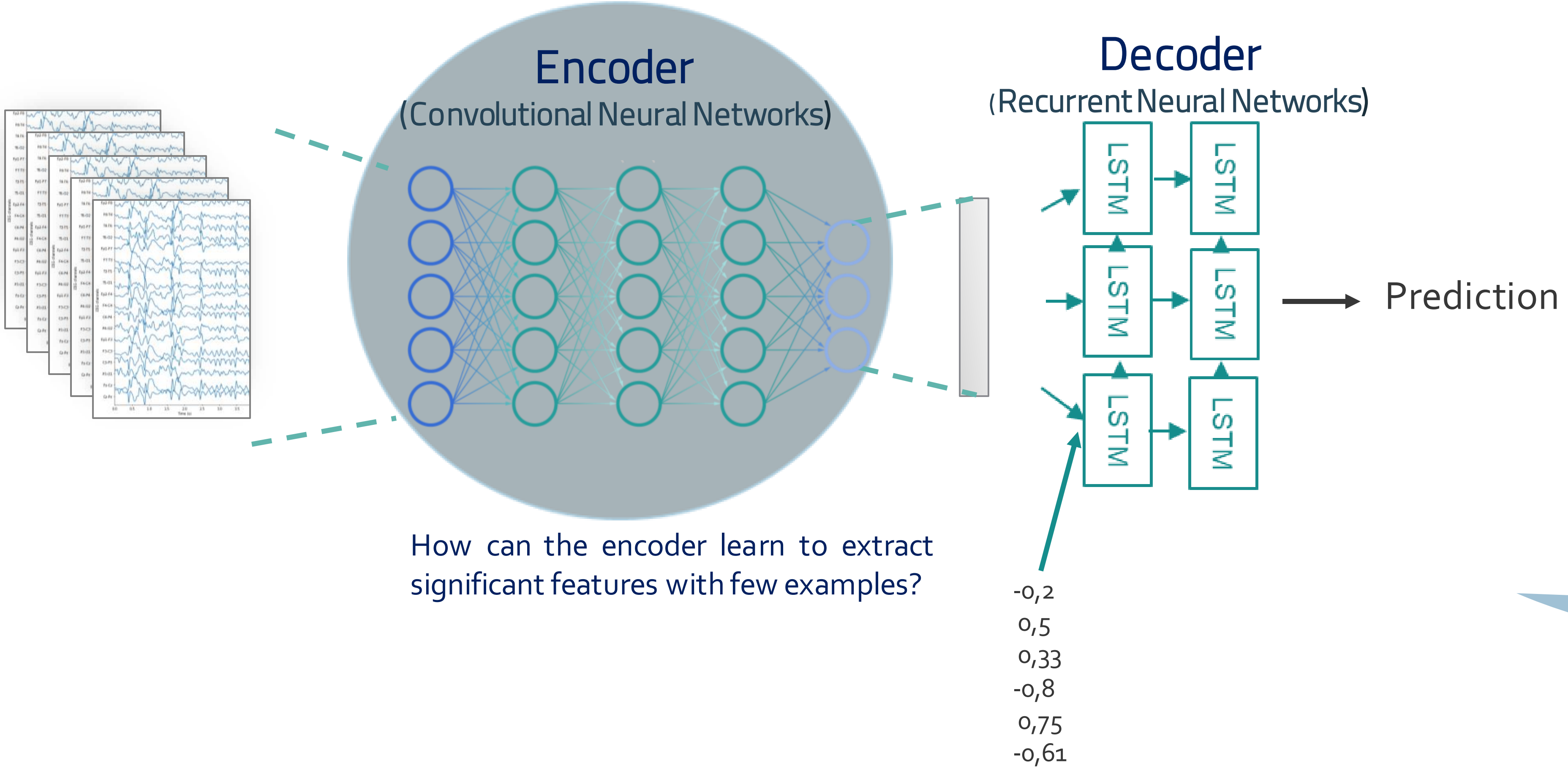
- Lack of annotations of TUH EEG dataset and data imbalance

TUH EEG Seizure Corpus (epileptic/no-epileptic)

Phenotypes:

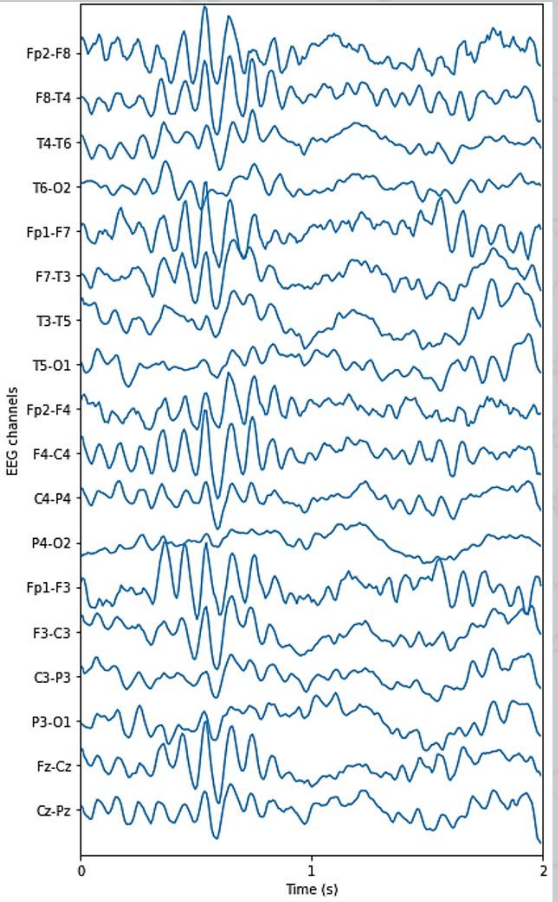
- normality
- generalized slowing
- focal slowing
- epileptiform discharges
- seizures
- abnormal delta waves
- sharp waves
- …

Encoder-Decoder architecture

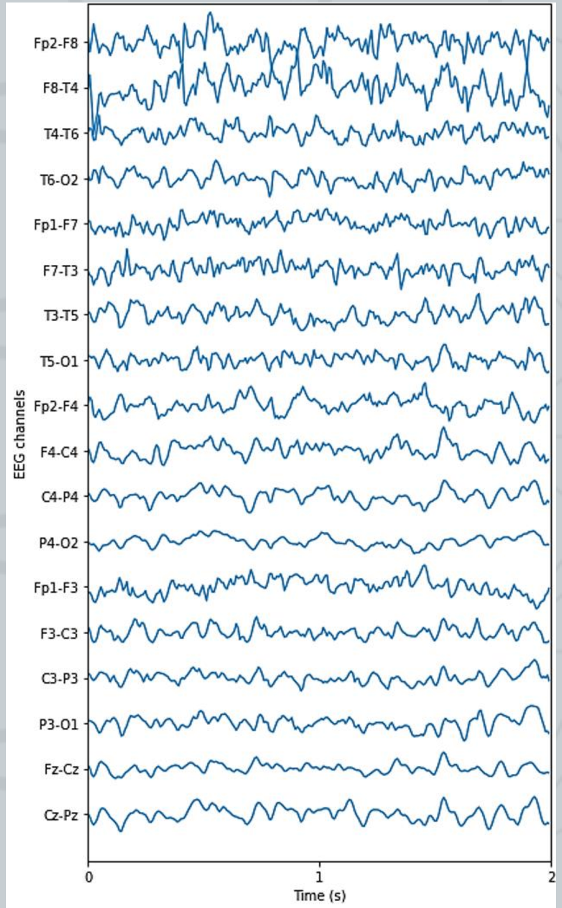


How can the encoder learn to extract significant features with few examples?

Pre-training Encoder

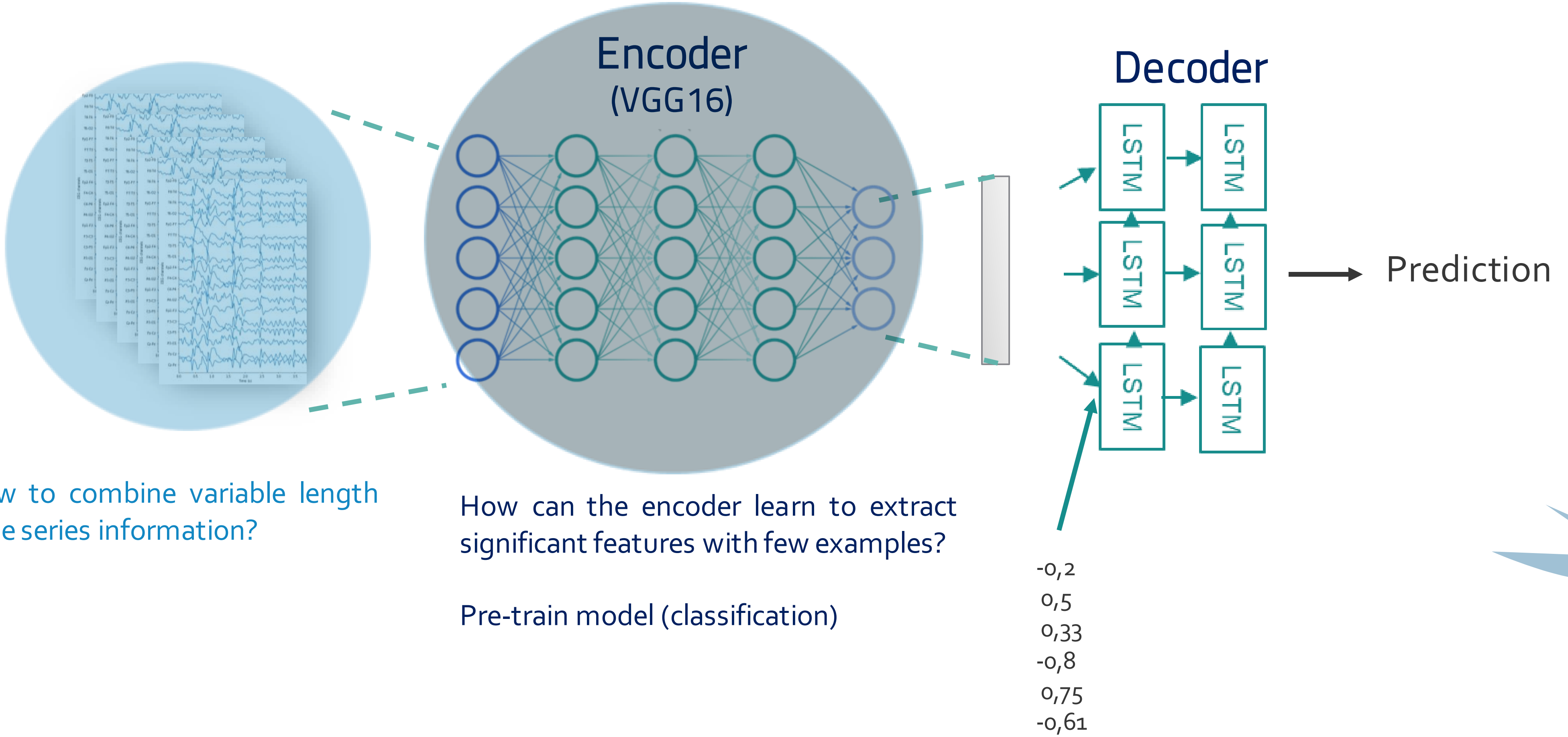


Epileptic event
(Spikes and/or Sharp Waves: SPSW)



No-Epileptic event
(Background: BCKG)

Encoder-Decoder architecture

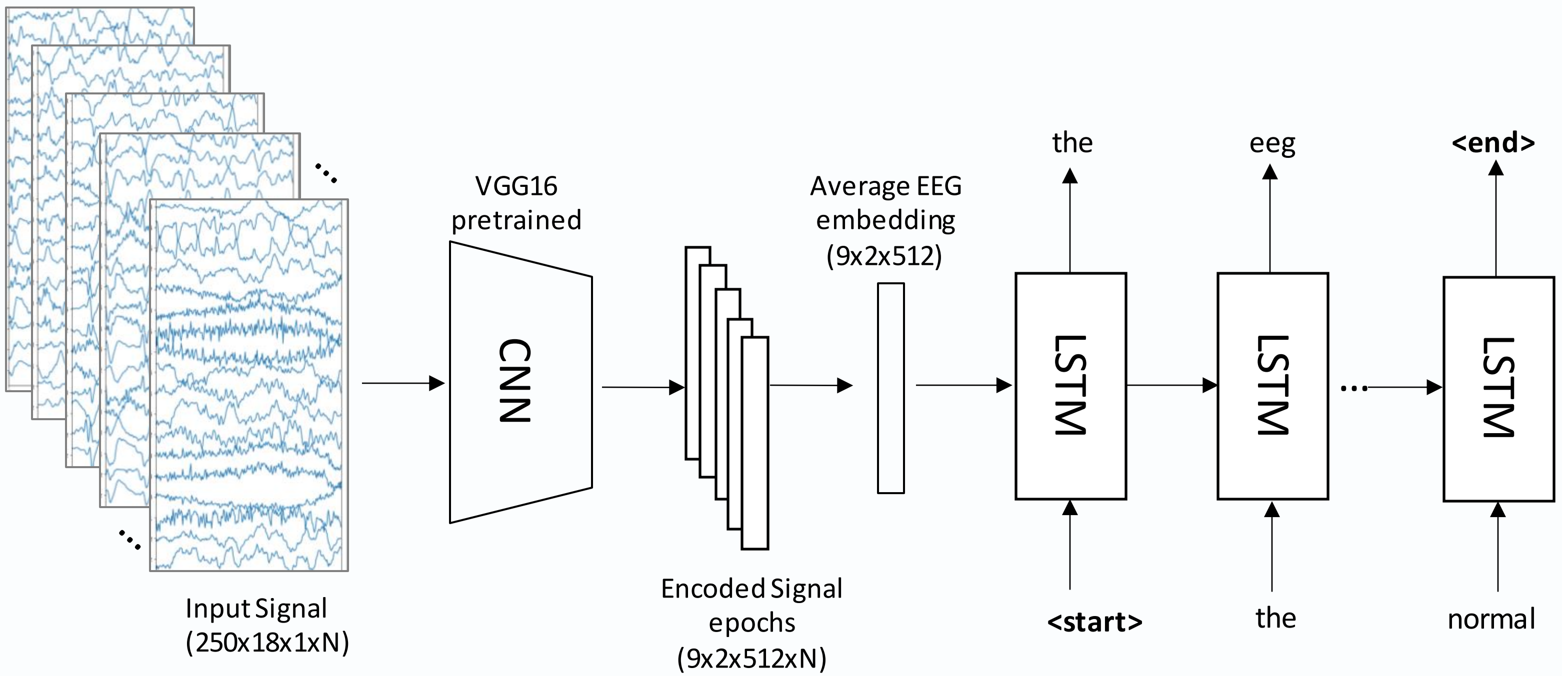


How to combine variable length time series information?

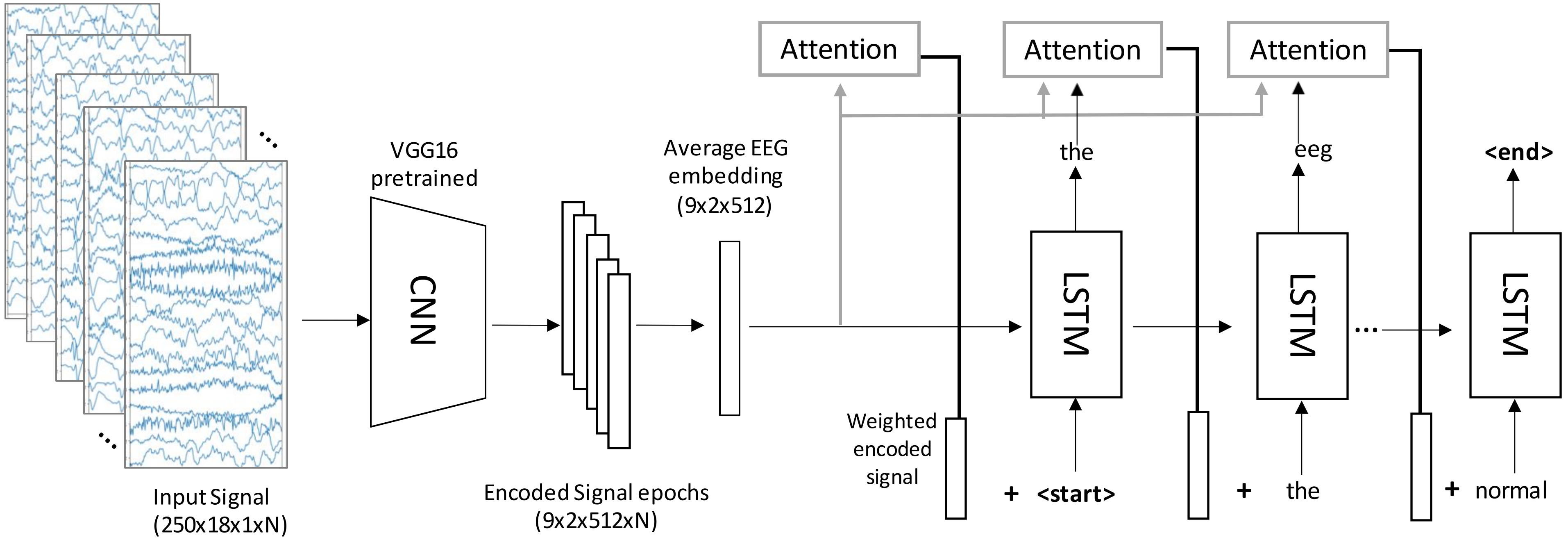
How can the encoder learn to extract significant features with few examples?

Pre-train model (classification)

Average EEG Embedding (CNN-LSTM)

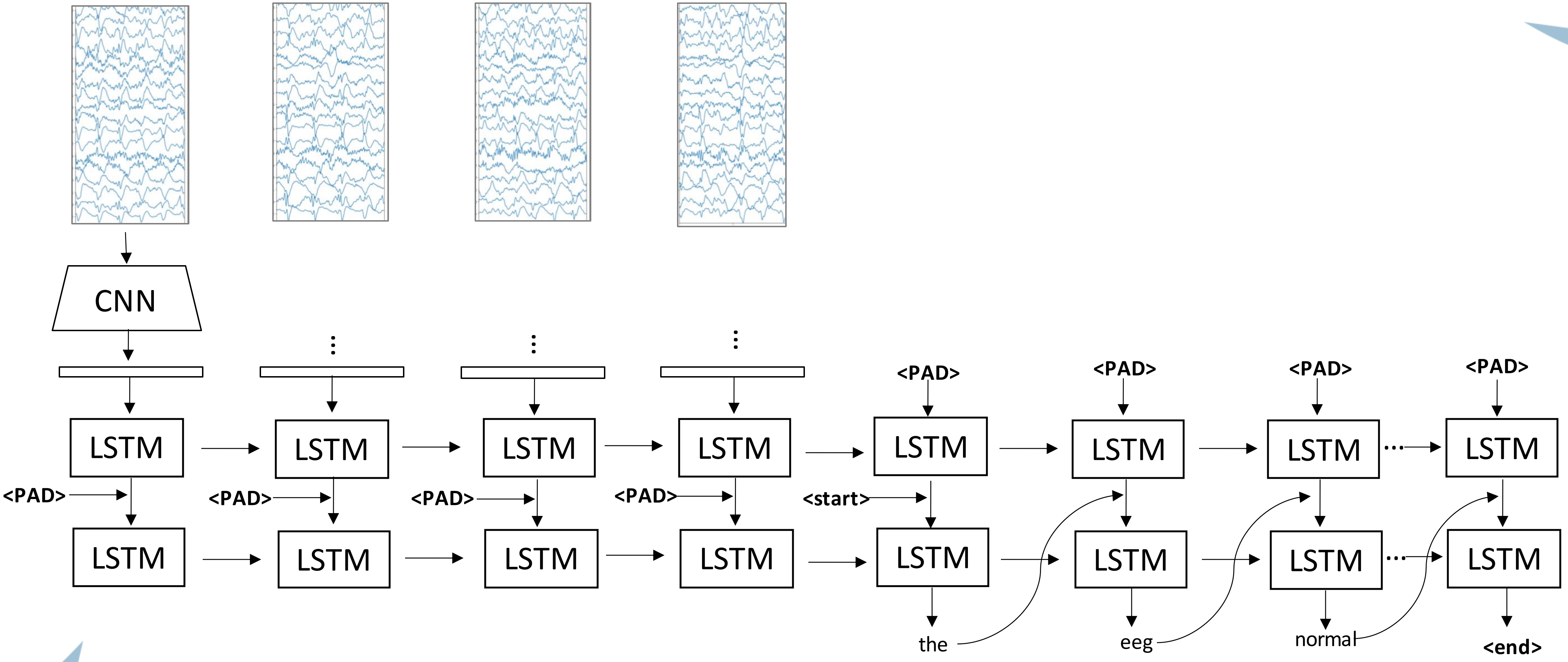


Average EEG Embedding (CNN-Att-LSTM)

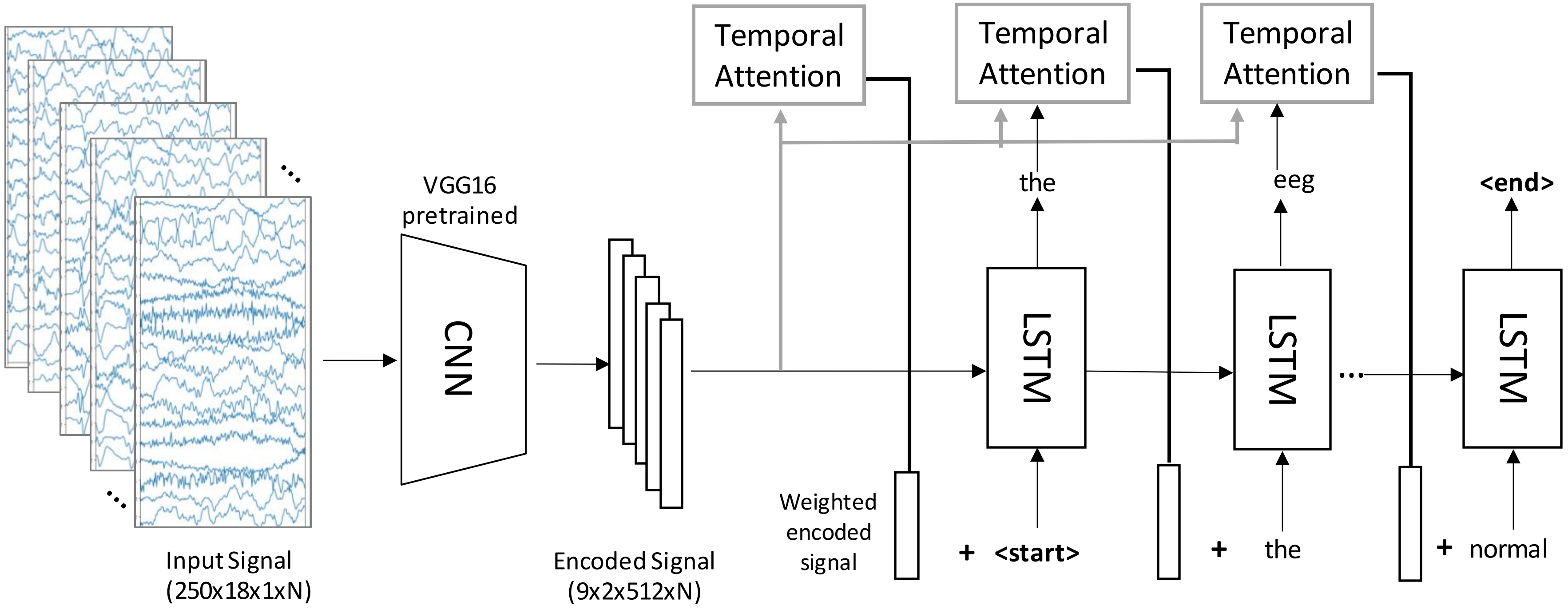


How about exploring the EEG time component?

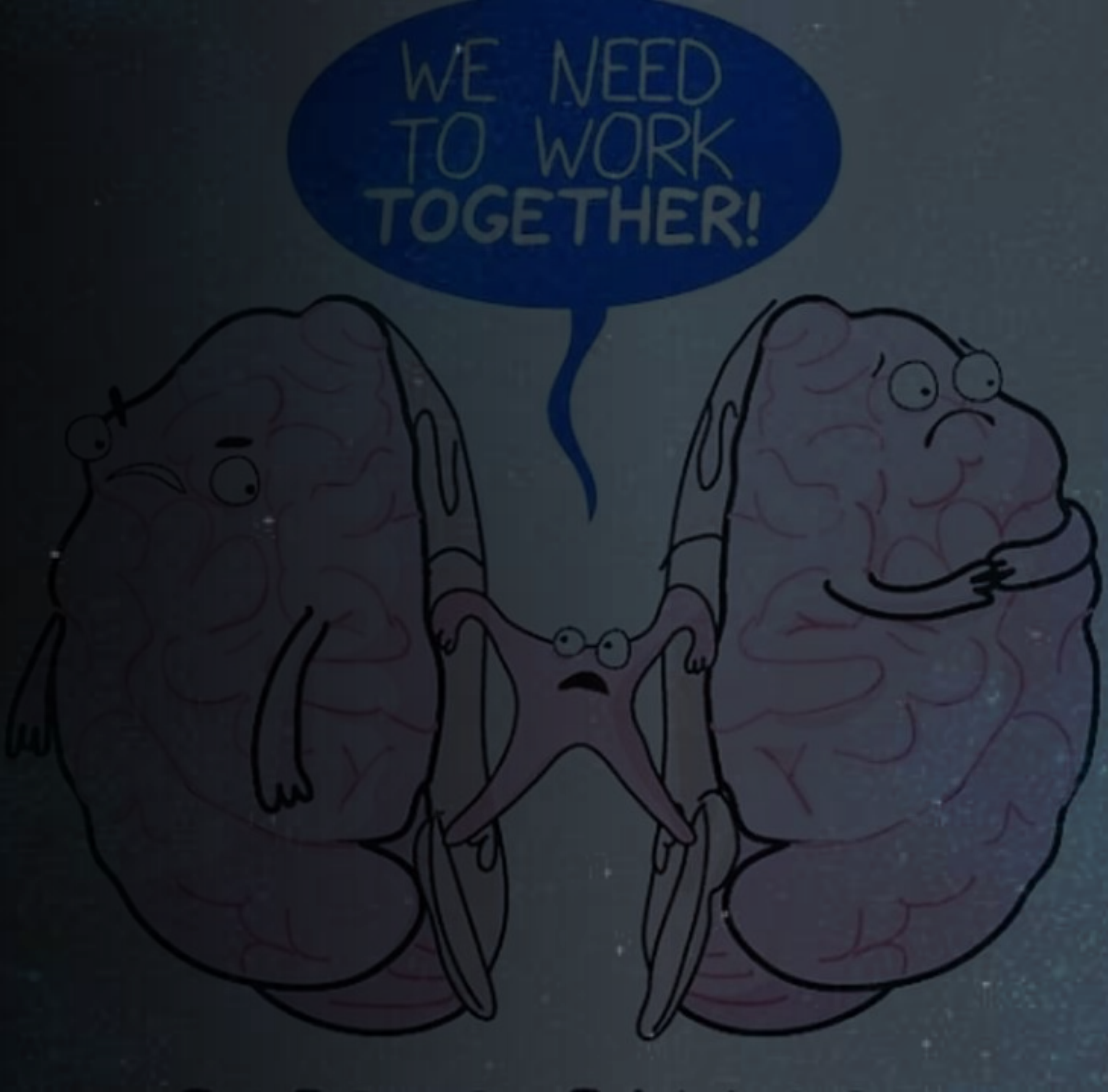
Sequence to Sequence (Seq2seq)



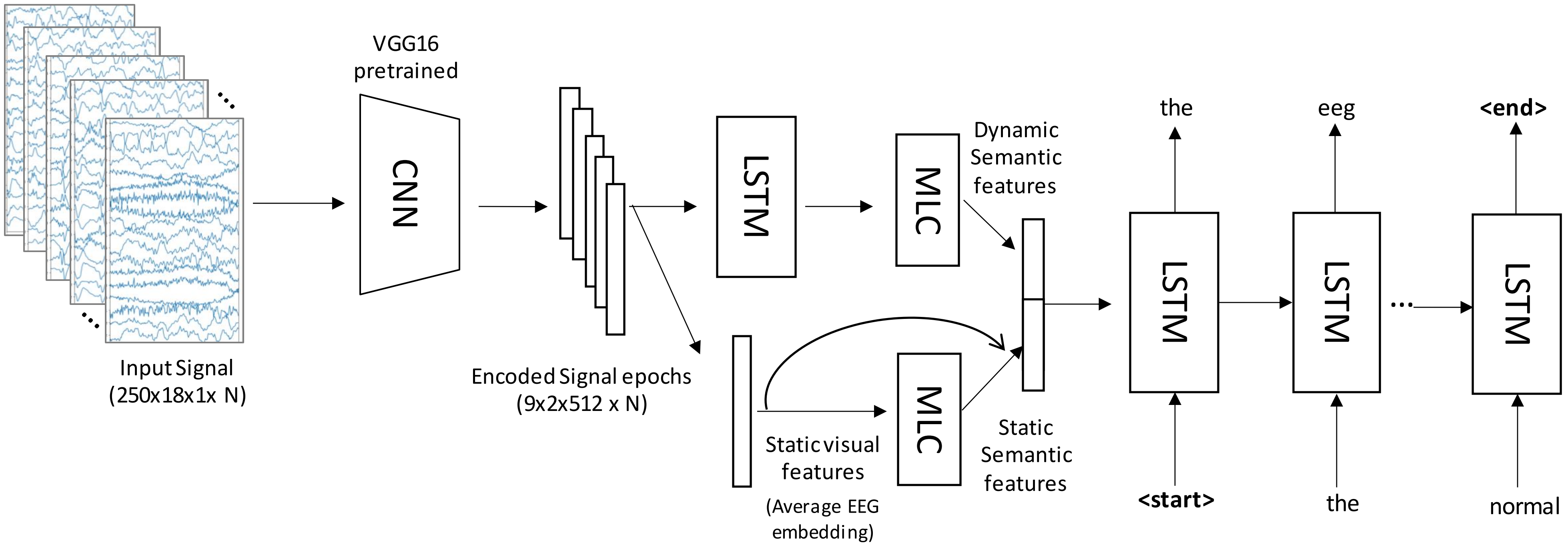
Temporal Attention Mechanism (TAM)



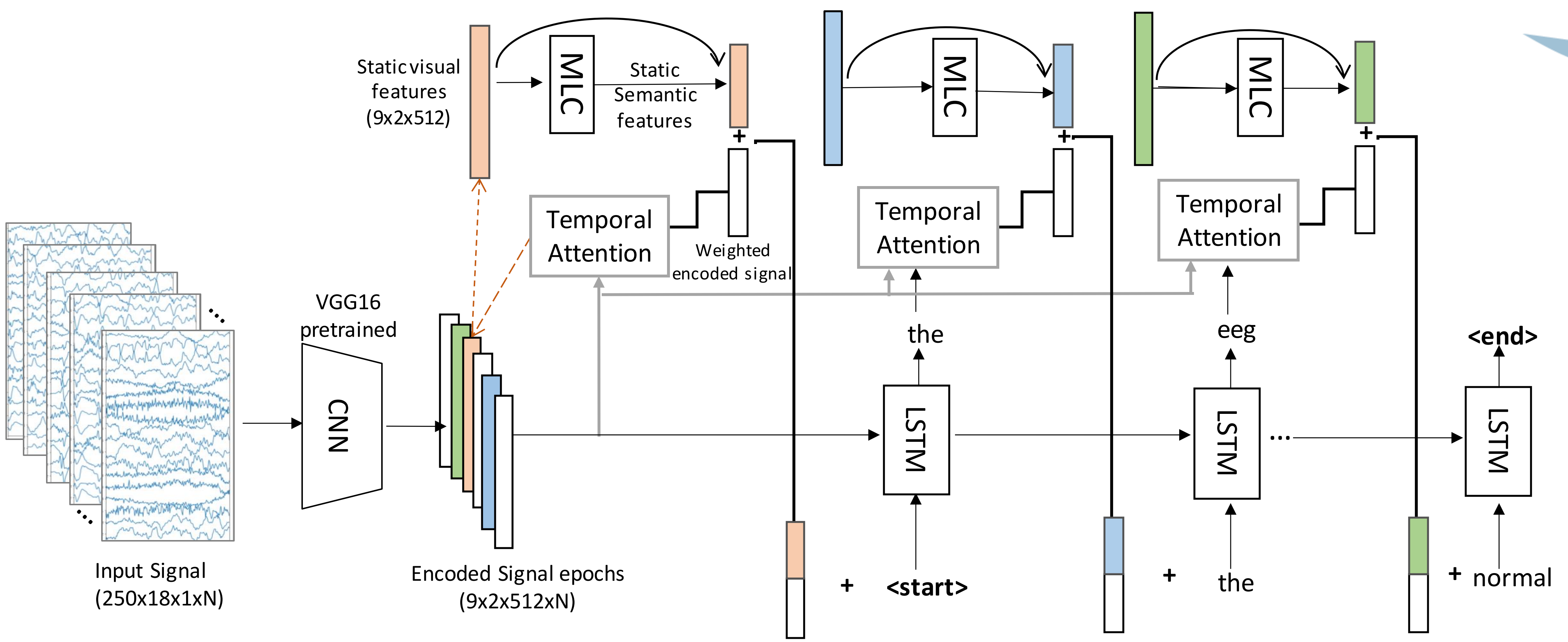
Maybe combine static and dynamic?



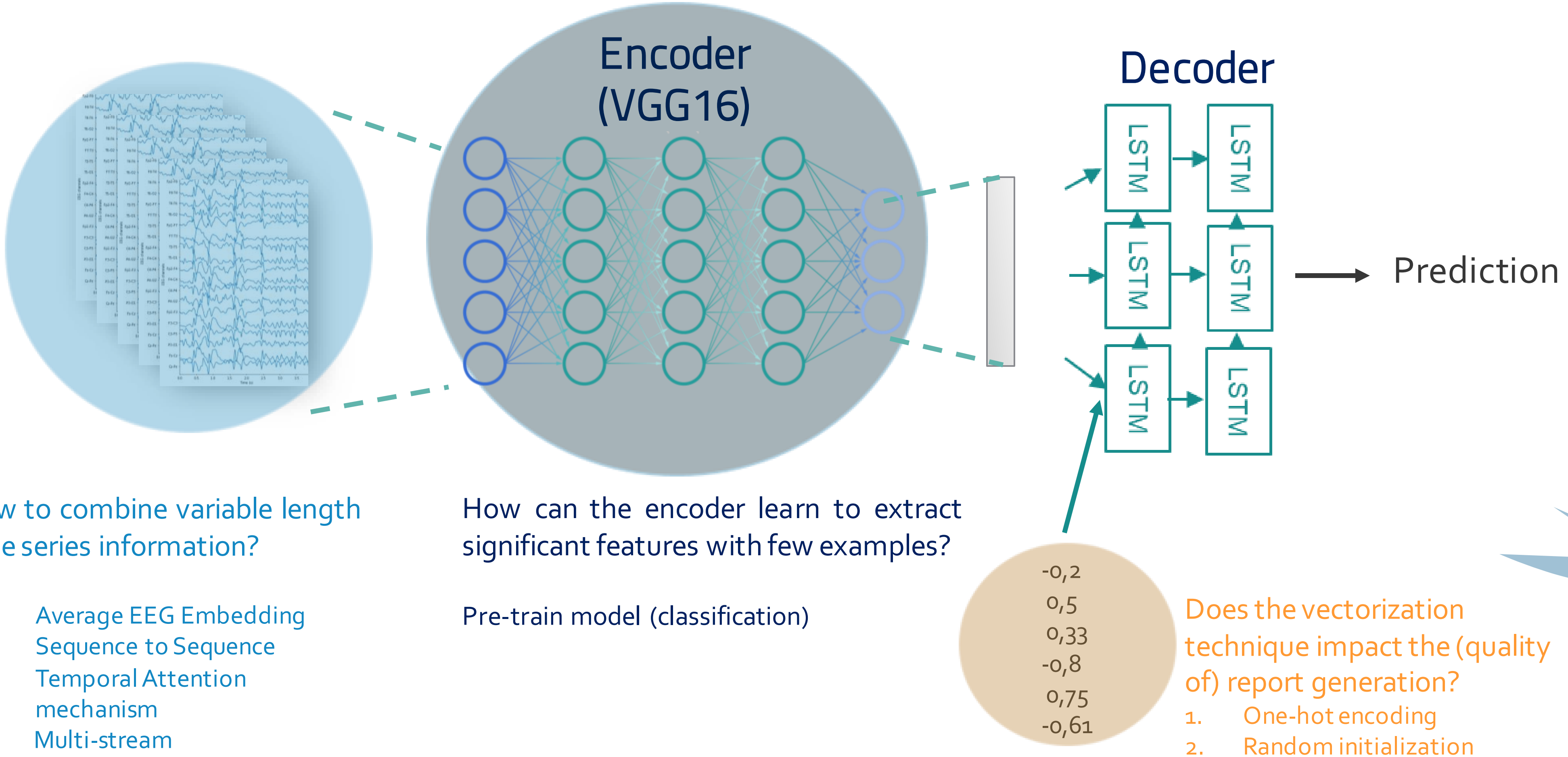
Multi-Stream Models
(Multi-stream-Avg)



Multi-Stream Models (Multi-stream-TAM)



Encoder-Decoder architecture



How to combine variable length time series information?

1. Average EEG Embedding
2. Sequence to Sequence
3. Temporal Attention mechanism
4. Multi-stream

How can the encoder learn to extract significant features with few examples?

Pre-train model (classification)

Does the vectorization technique impact the (quality of) report generation?

1. One-hot encoding
2. Random initialization
3. Word2Vec
4. fastText

Results

Classification Task

Table 2. VGG16 Model performance for binary classification (SPSW and BCKG)

Set	***AUC***	***Sensitivity***	***Specifity***	***Precision***	***F1-score***
Train	0,81	88%	75%	25%	0,39
Test	0,80	82%	77%	28%	0,42

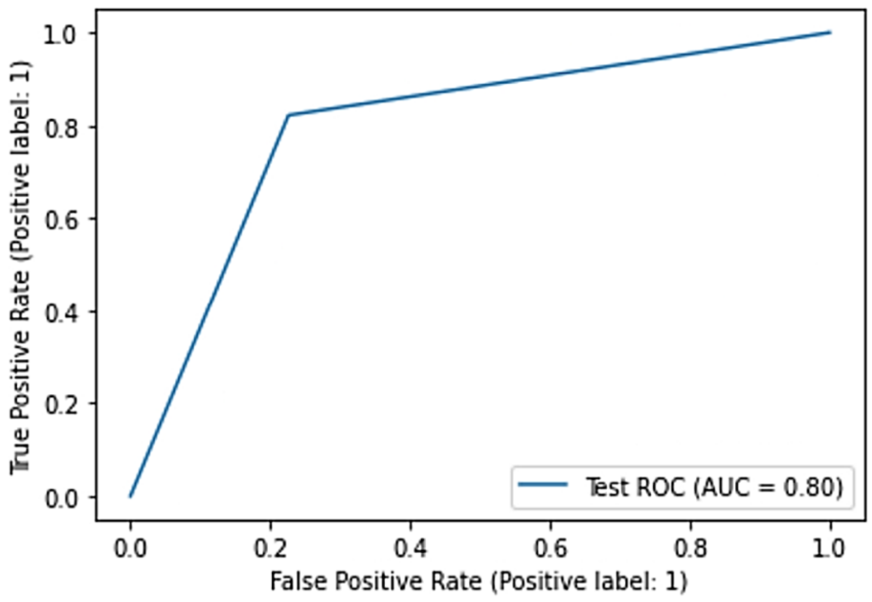


Fig 1. ROC curve on test set

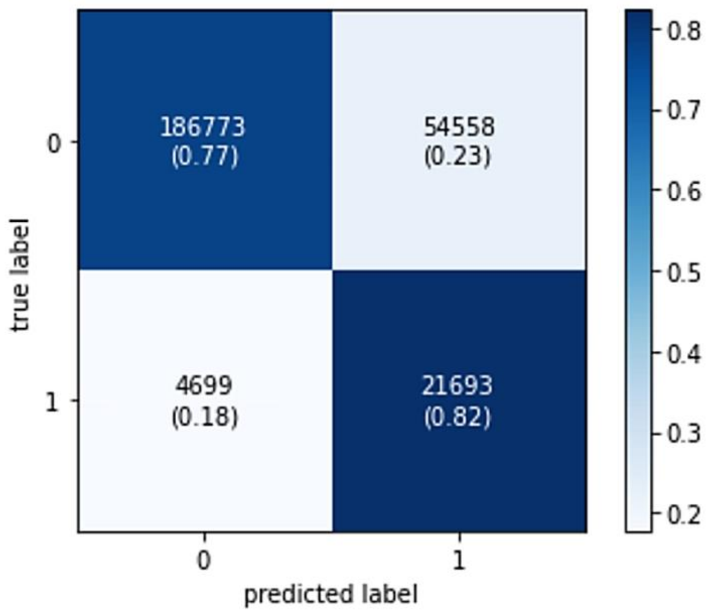


Fig 2. Confusion matrix
(0-negative, 1-positive)

Impact of text representation approaches

Table 3: Results from CNN-LSTM using different word embeddings

Method	***BLEU1***	***BLEU2***	***BLEU3***	***BLEU4***	***METEOR***	***ROUGE_L***	***CIDEr***	***SPICE***
One-Hot encod.	53,5	**39,7**	31,6	25,1	21,4	44,2	15,5	**18,2**
Random init.	54,1	40,6	**32,4**	**25,7**	**21,6**	**44,3**	**22,3**	17,3
Word2Vec	**54,6**	39,1	25,6	14,7	17,6	35,8	14,2	17,4
fastText	49,5	37,1	25,4	15,7	17,8	37,2	14,0	16,8

EEG Captioning
(Architecture Comparison)

Table 4. Model performance of all captioning approaches with random embedding initialization

Method	***BLEU1***	***BLEU2***	***BLEU3***	***BLEU4***	***METEOR***	***ROUGE_L***	***CIDEr***	***SPICE***
CNN-LSTM	54,1	40,6	32,4	25,7	21,6	44,3	**22,3**	17,3
CNN-Att-LSTM	50,5	38,9	31,3	25,0	20,7	42,5	16,9	17,1
Seq2seq	51,2	35,9	26,9	19,9	17,6	40,1	16,4	13,6
TAM	52,2	39,8	32,0	25,7	20,0	44,1	16,5	16,9
Multi-stream-Avg	**56,3**	**43,4**	**35,0**	**28,3**	**23,3**	**46,0**	21,8	**18,4**
Multi-stream-TAM	54,8	41,4	32,9	26,3	23,2	44,7	21,5	19,3

EEG Captioning
(Architecture Comparison)

Table 5. Model performance of average embedding-based models

Method	***BLEU1***	***BLEU2***	***BLEU3***	***BLEU4***	***METEOR***	***ROUGE_L***	***CIDEr***	***SPICE***
CNN-LSTM	**54,1**	**40,6**	**32,4**	**25,7**	**21,6**	**44,3**	**22,3**	**17,3**
CNN-Att-LSTM	50,5	38,9	31,3	25,0	20,7	42,5	16,9	17,1

Table 6. Qualitative model evaluation of average embedding-based models

Methods	**Generated report**
CNN-LSTM	this eeg is marked by the occurrence of sharp waves and **generalized slowing**.
CNN-Att-LSTM	this eeg is marked by the occurrence of several **seizures** , in addition to the **focal slowing** in the right posterior quadrant.
Ground truth: abnormal eeg due to the arising of intermittent left occipital focal **seizures**, in addition to **background slowing** and focal left hemispheric **slowing**	

EEG Captioning
(Architecture Comparison)

Table 4. Model performance of all captioning approaches with random embedding initialization

Method	***BLEU1***	***BLEU2***	***BLEU3***	***BLEU4***	***METEOR***	***ROUGE_L***	***CIDEr***	***SPICE***
CNN-LSTM	**54,1**	**40,6**	**32,4**	**25,7**	**21,6**	**44,3**	**22,3**	**17,3**
CNN-Att-LSTM	50,5	38,9	31,3	25,0	20,7	42,5	16,9	17,1
Seq2seq	51,2	35,9	26,9	19,9	17,6	40,1	16,4	13,6
TAM	52,2	39,8	32,0	**25,7**	20,0	44,1	16,5	16,9
Multi-stream-Avg	**56,3**	**43,4**	**35,0**	**28,3**	**23,3**	**46,0**	21,8	**18,4**
Multi-stream-TAM	54,8	41,4	32,9	26,3	23,2	44,7	21,5	19,3

EEG Captioning
(Architecture Comparison)

Table 4. Model performance of all captioning approaches with random embedding initialization

Method	***BLEU1***	***BLEU2***	***BLEU3***	***BLEU4***	***METEOR***	***ROUGE_L***	***CIDEr***	***SPICE***
CNN-LSTM	54,1	40,6	32,4	25,7	21,6	44,3	**22,3**	17,3
CNN-Att-LSTM	50,5	38,9	31,3	25,0	20,7	42,5	16,9	17,1
Seq2seq	51,2	35,9	26,9	19,9	17,6	40,1	16,4	13,6
TAM	52,2	39,8	32,0	25,7	20,0	44,1	16,5	16,9
Multi-stream-Avg	**56,3**	**43,4**	**35,0**	**28,3**	**23,3**	**46,0**	21,8	18,4
Multi-stream-TAM	54,8	41,4	32,9	26,3	23,2	44,7	21,5	**19,3**

Could models effectively characterize the eeg re
What about the localization?

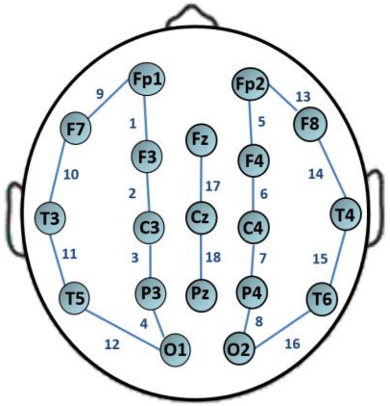


Table 7. Qualitative model evaluation of all captioning approaches

Methods	Generated report
CNN-LSTM	this eeg is marked by the occurrence of sharp waves and generalized slowing.
CNN-Att-LSTM	this eeg is marked by the occurrence of several seizures , in addition to the **focal slowing** in the <u>**right posterior**</u> quadrant.
Seq2Seq	this eeg is remarkable for **focal** <u>**right hemispheric**</u> **slowing** , mild to moderate diffuse slowing and focal voltage attenuation.
TAM	abnormal eeg due to the arising of multiple frequent seizures and moderate diffuse slowing.
Multi-stream-Avg	this eeg is marked by the occurrence of seizure , in addition to continuous **focal slowing** and moderate diffuse background slowing.
Multi-stream-TAM	this eeg is marked by the occurrence of seizure , in addition to **focal slowing** in the <u>**left anterior temporal**</u> region with diffuse background slowing .
Ground truth: abnormal eeg due to the arising of intermittent left occipital focal seizures, in addition to background slowing and **focal** <u>left hemispheric</u> **slowing**	

Could models effectively characterize the EEG recording?

Table 8. Qualitative evaluation of model performance, generating clinical report from a normal EEG

Methods	Generated report
CNN-LSTM	this eeg is remarkable for focal delta activity in the right hemisphere and marked background slowing.
CNN-Att-LSTM	this eeg is remarkable for focal slowing in the right temporal region.
Seq2Seq	abnormal eeg because of multiple left frontal electrographic seizures , in addition to focal slowing in the left hemisphere.
TAM	this eeg is marked by the occurrence of seizure , in addition to the focal slowing from the right hemisphere .
Multi-stream-Avg	this eeg is marked by the occurrence of seizure , in addition to continuous focal slowing and moderate diffuse background slowing.
Multi-stream-TAM	this eeg is marked by the occurrence of seizure , in addition to focal slowing in the left anterior temporal region with diffuse background slowing .
Ground truth: eeg within the normal limits.	

Conclusions

It is possible and there is a **great potential** in generating clinical reports from EEGs, but EEG captioning models are **not yet ready for** implementation in **clinical practice**

- The main limitations are related to the dataset
- Compromise between the amount of data and the complexity of the models
- Feature extraction and summarization in global embeddings are one of the major challenges
- Models are restricted regarding the diversity of phenotypes
- Models generate incomplete reports and have difficulty in identifying locations
- Model performance is still far from optimal

Conclusions → Future work

- Feature extraction and summarization in global embeddings are one of the major challenges

→ Multi-stream-Avg seem promising but need to be further studied and improved :

  - Datasets
  - Assess the efficiency of feature extraction
  - Pre-training the streams
  - Use 3D-CNN instead of 2D-CNN

- Model are restricted regarding the diversity of phenotypes

→ Train in a more comprehensive database

- Model Generate incomplete reports and have difficulty in identifying locations.

→ Dense Captioning approach

- Model performance is still far from optimal

→ Training with larger datasets;

↘ Transformer-based architectures:

  - Multiple stack transform layers
  - ClinicalBERT or BioBERT

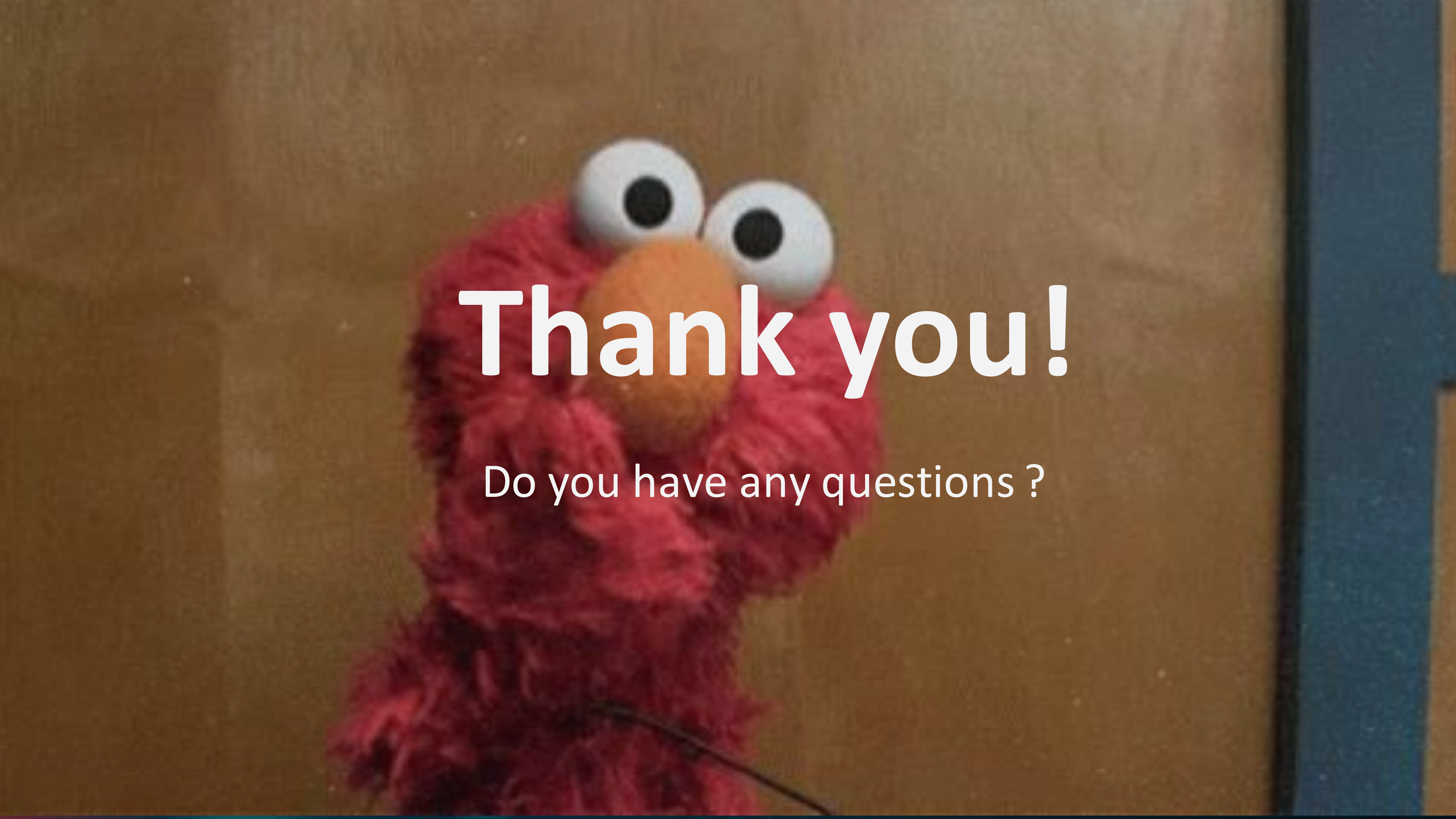
↘ GANs

Future work

- Multi-stream-Avg seem promissing but need to be further studied and improved
- Train in a more comprehensive database
- Dense Captioning approach
- Training with larger datasets
- Transformer-based architectures
- GANs

How far are we from the implementation of such a system in the clinic?

- Clinical validation by a group of experts
- Model interpretability (XAI)



Thank you!

Do you have any questions ?