# Hybrid Retrieval-Generation Reinforced Agent for Medical Image Report Generation

Yuan Li, Xiaodan Liang, Zhiting Hu, Eric P. Xing — NeurIPS 2018

CS885 Fall 2021 - Reinforcement Learning

Presenter: **Gustavo Sutter**
*(gsutterp@uwaterloo.ca)*

UNIVERSITY OF WATERLOO

# Report generation

- <u>Goal</u>: generate long and topic-coherent stories or reports to describe visual contents
- <u>Challenges</u>:
  - The report must be a long narrative with multiple sentences, having plausible logic and consistent topics
  - There is a presumed content coverage and specific terminology, both depending on the task at hand

# Medical image reports


**Comparison:**

**Indication:** 60-year-old male with seizure, ethanol abuse

**Findings:** The heart size and mediastinal contours appear within normal limits. There is blunting of the right lateral costophrenic sulcus which could be secondary to a small effusion versus scarring. No focal airspace consolidation or pneumothorax. No acute bony abnormalities.

**Impression:** Blunting of the right costophrenic sulcus could be secondary to a pleural effusion versus scarring.



UNIVERSITY OF WATERLOO

# Why not use caption generation models?

- Most of the sentences present in medical reports are normal findings (low variance of language)
  - Retrieval-based systems work well for those sentences
- Abnormal findings are rare and very diverse, but very important
  - Text generation models fail to deal capture the diversity
  - Models based only on text generation tend to focus on generating sentences that look natural, but not necessarily supported by the visual evidences



UNIVERSITY OF WATERLOO

# HRGR-Agent

- Hybrid Retrieval-Generation Reinforced Agent (HRGR-Agent)
- Combines traditional retrieval-based approaches with modern text generation
- Two main modules:
  - Retrieval policy module
  - Generation module
- Modules are jointly trained using RL

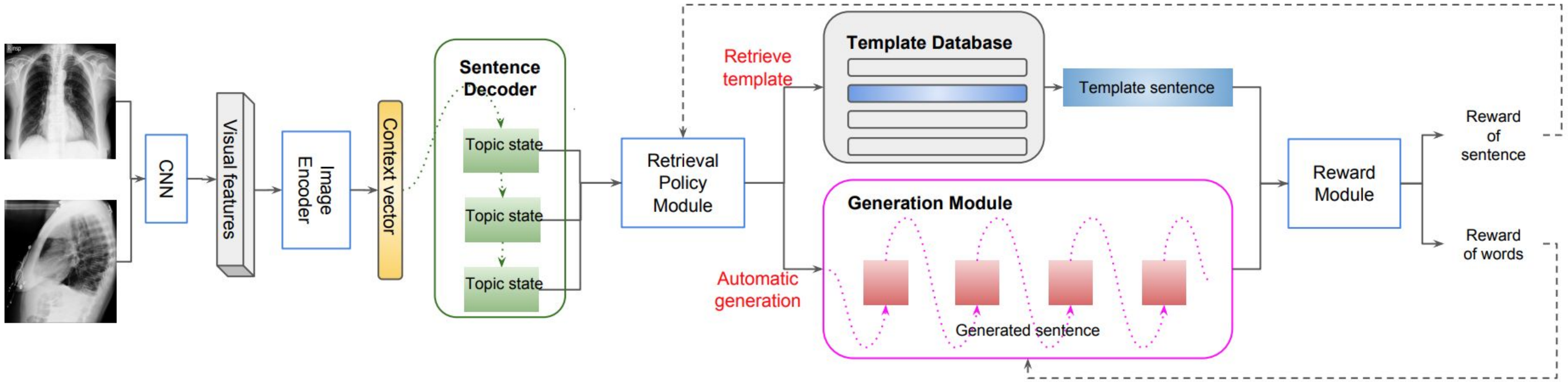
# HRGR-Agent: overview

- Given a set of medical image $\mathbf{I} = \{I_j\}_{j=1}^{K}$ the goal is to generate a sequence ( $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M)$ $\mathbf{y}_i = (y_{i,1}, y_{i,2}, \ldots, y_{i,N})$
- From the extracted features of the images the model generates a topic state for each sentence
- Using the topic state the module can either choose a template sentence or generate a novel sentence from scratch
  - The template databas $\mathbb{T}$ T is extracted from the training corpus
- Supervision comes on both sentence-level and word-level.



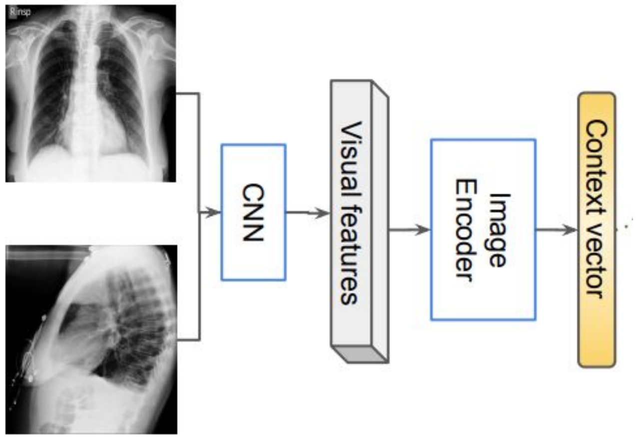
UNIVERSITY OF WATERLOO

# HRGR-Agent: overview



# HRGR-Agent: Image Encoder

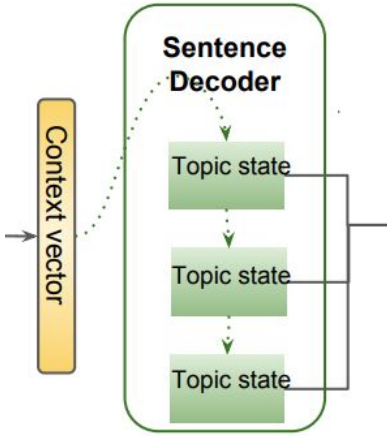
- Given a set of images $\{I_j\}_{j=1}^K$ a CNN is used to extract their features $\{\mathbf{v}_j\}_{j=1}^K$
- The images features are averaged and passed to a fully-connected layer to get the context vector $\mathbf{h}^v$



# HRGR-Agent: Sentence Decoder

- Stacked RNN layers that take the context vector $\mathbf{h}^v$ and generate a sequence of topic states $\mathbf{q}_i$

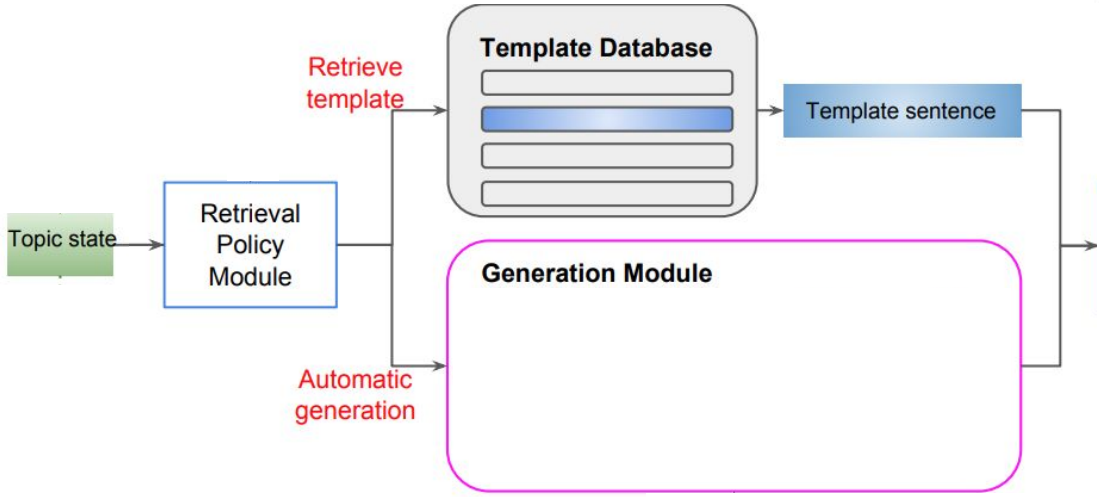
$$\mathbf{c}_i^s = F_{\text{attn}}^s(\mathbf{h}^v, \mathbf{h}_{i-1}^s)$$
$$\mathbf{h}_i^s = F_{\text{RNN}}^s(\mathbf{c}_i^s, \mathbf{h}_{i-1}^s)$$
$$\mathbf{q}_i = \sigma(\mathbf{W}_q \mathbf{h}_i^s + \mathbf{b}_q)$$
$$z_i = \text{Sigmoid}(\mathbf{W}_z \mathbf{h}_i^s + \mathbf{b}_z)$$



# HRGR-Agent: Retrieval Policy Module

- Given each topic state $\mathbf{q}_i$ the retrieval policy module has two options:
  - Retrieve a template sentence from $\mathbb{T}$
  - Activate the generation module to produce a sequence of words

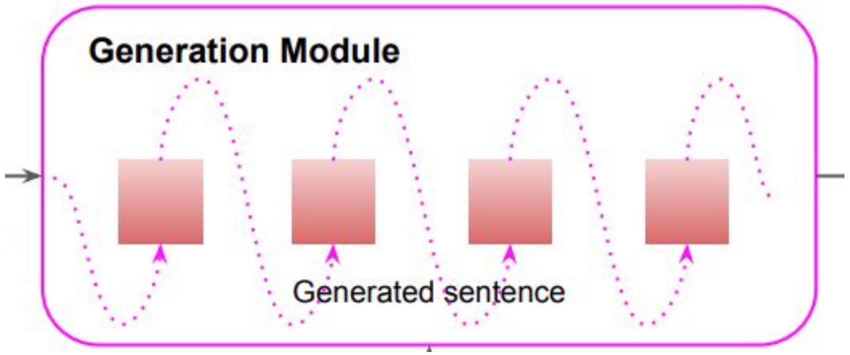
$$
\mathbf{u}_i = \text{Softmax}(\mathbf{W}_u \mathbf{q}_i + \mathbf{b}_u)
$$
$$
m_i = \text{argmax}(\mathbf{u}_i),
$$



# HRGR-Agent: Generation Module

- The generation module consists of usual RNNs with attention, generating new words conditioned on the context vector and the topic state

$$
\begin{aligned}
\mathbf{c}^g_{i,t} &= F^g_{\text{attn}}(\mathbf{h}^v, [\mathbf{e}_{i,t-1}; \mathbf{q}_i], \mathbf{h}^g_{i,t-1}) \\
\mathbf{h}^g_{i,t} &= F^g_{\text{RNN}}([\mathbf{c}^g_{i,t}; \mathbf{e}_{i,t-1}; \mathbf{q}_i], \mathbf{h}^g_{i,t-1}) \\
\mathbf{a}_t &= \text{Softmax}(\mathbf{W}_y \mathbf{h}^g_{i,t} + \mathbf{b}_y) \\
y_t &= \text{argmax}(\mathbf{a}_t) \\
\mathbf{e}_{i,t} &= \mathbf{W}_e \mathbb{O}(y_{i,t}),
\end{aligned}
$$



# HRGR-Agent: Reward Module

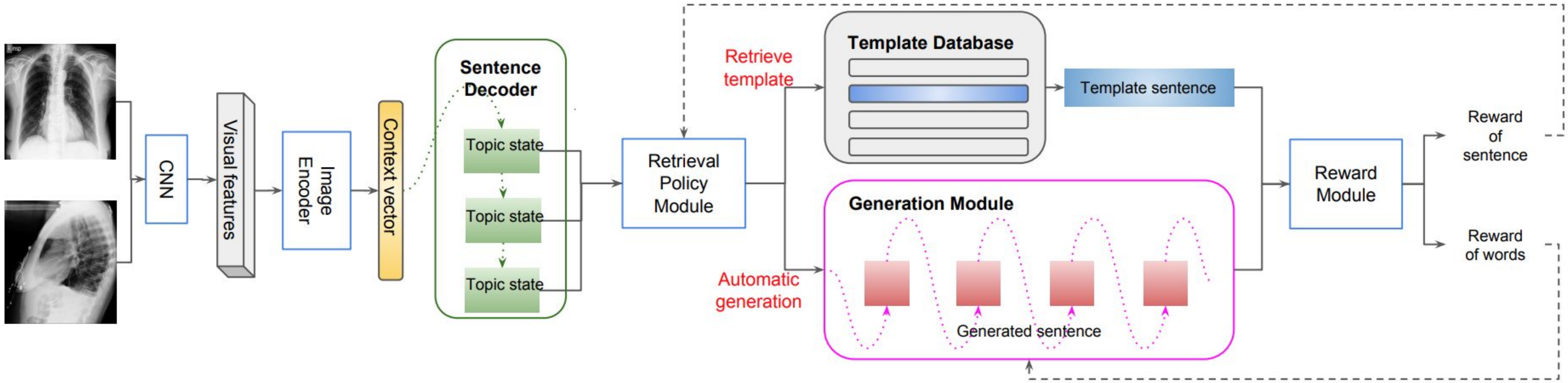
- Rewards are computed using the CIDEr metric
- Sentence-level reward

$$R_{sent}(\mathbf{y}_i) = f(\{\mathbf{y}_k\}_{k=1}^{i}, \mathrm{gt}) - f(\{\mathbf{y}_k\}_{k=1}^{i-1}, \mathrm{gt})$$

- Word-level reward

$$R_{word}(y_t) = f(\{y_k\}_{k=1}^{t}, \mathrm{gt}^s) - f(\{y_k\}_{k=1}^{t-1}, \mathrm{gt}^s)$$

# HRGR-Agent: overview






UNIVERSITY OF WATERLOO

# HRGR-Agent: Hierarchical Reinforcement Learning

- The objective is maximize the reward of the produced sequences given the ground-truth report
- REINFORCE algorithm

$$
\begin{aligned}
\mathcal{L}(\theta) &= -\mathbb{E}_{z,m,y}[R(\mathbf{Y}, \mathbf{Y}^*)] \\
\nabla_\theta \mathcal{L}(\theta) &= -\mathbb{E}_{z,m,y}\left[\nabla_\theta \log p(z,m,y) R(\mathbf{Y}, \mathbf{Y}^*)\right] \\
&= -\mathbb{E}_{z,m,y}\left[\sum_{i=1} \mathbb{1}(z_i < \frac{1}{2}|z_{i-1})\Big(\nabla_{\theta_r}\mathcal{L}(\theta_r) + \mathbb{1}(m_i = 0|m_{i-1})\nabla_{\theta_g}\mathcal{L}(\theta_g)\Big)\right]
\end{aligned}
$$

# HRGR-Agent: Hierarchical Reinforcement Learning

- Policy Upgrade for <u>Retrieval Policy Module</u>
  - Sentence-level

$$R^r(\mathbf{y}_i) = \sum_{j=0}^{\infty} \gamma^j R_{sent}(\mathbf{y}_{i+j})$$

$$\mathcal{L}(\theta_r) = -\mathbb{E}_{m_i}[R^r(m_i, m_i^*)]$$

$$\nabla_{\theta_r}\mathcal{L}(\theta_r) = -\mathbb{E}_{m_i}[\nabla_{\theta_r} \log p(m_i|m_{i-1}) R^r(m_i, m_i^*)]$$

# HRGR-Agent: Hierarchical Reinforcement Learning

- Policy Upgrade for Generation Module
  - Word-level

$$R^g(y_t) = \sum_{j=0}^{\infty} \gamma^j R_{word}(y_{t+j})$$

$$\mathcal{L}(\theta_g) = -\mathbb{E}_{y_t}[R^g(\mathbf{y}_t, \mathbf{y}_t^*)]$$

$$\nabla_{\theta_g} \mathcal{L}(\theta_g) = -\mathbb{E}_{y_t}[\sum_{t=1} \nabla_{\theta_g} \log p(y_t|y_{t-1}) R^g(y_t, y_t^*)]$$

# Experiments: Datasets

- Indiana University Chest X-Ray Collection (IU X-Ray)
  - 7,470 images paired with their corresponding reports
  - 1185 unique tokens
- CX-CHR
  - 35,236 patients (no more than 2 photos for each)
  - 1282 unique tokens
- For each dataset split data into training (70%), validation (20%) and testing (10%)
- Model was trained to predict the findings section of each report

# Experiments: Template Database

- Sentences are selected based on their frequency in the training set
- Candidates that express the same meaning but have little linguistic variation are grouped
  - Only the most frequent sentence of each group will be retrieved by the models
  - Does introduce error in the results, but authors claim that is negligible

# Experiments: Evaluation Metrics

- Automatic metrics
  - CIDEr
  - BLEU
  - ROUGE
- Abnormality detection (select 10 most frequency medical abnormalities)
  - Precision and average false positive (AFP)
- Human evaluation
  - Mechanical Turk surveys to decide which method better matches with the ground-truth

# Experiments: Baselines

- Image captioning models
  - CNN-RNN, LRCN, AdaAtt, Att2in
- Previous work on medical imaging reports
  - CoAtt (uses a different feature extractor)
- Variations of their method
  - Generation — no templates and no RL
  - HRG — no RL
  - Retrieval — same as HRGR-Agent but masking the generated sentences

# Experiments: Results

| Dataset | Model | CIDEr | BLEU-1 | BLEU-2 | BLEU-3 | BLEU-4 | ROUGE |
|---|---|---|---|---|---|---|---|
| **CX-CHR** | CNN-RNN [34] | 1.580 | 0.590 | 0.506 | 0.450 | 0.411 | 0.577 |
| | LRCN [9] | 1.588 | 0.593 | 0.508 | 0.452 | 0.413 | 0.577 |
| | AdaAtt [23] | 1.568 | 0.588 | 0.503 | 0.446 | 0.409 | 0.575 |
| | Att2in [28] | 1.566 | 0.587 | 0.503 | 0.446 | 0.408 | 0.576 |
| | Generation | 0.361 | 0.307 | 0.216 | 0.160 | 0.121 | 0.322 |
| | Retrieval | 2.565 | 0.535 | 0.475 | 0.437 | 0.409 | 0.536 |
| | HRG | 2.800 | 0.629 | 0.547 | 0.497 | 0.463 | 0.588 |
| | HRGR-Agent | **2.895** | **0.673** | **0.587** | **0.530** | **0.486** | **0.612** |
| **IU X-Ray** | CNN-RNN [34] | 0.294 | 0.216 | 0.124 | 0.087 | 0.066 | 0.306 |
| | LRCN [9] | 0.284 | 0.223 | 0.128 | 0.089 | 0.067 | 0.305 |
| | AdaAtt [23] | 0.295 | 0.220 | 0.127 | 0.089 | 0.068 | 0.308 |
| | Att2in [28] | 0.297 | 0.224 | 0.129 | 0.089 | 0.068 | 0.308 |
| | CoAtt* [16] | 0.277 | **0.455** | 0.288 | 0.205 | **0.154** | **0.369** |
| | HRGR-Agent | **0.343** | 0.438 | **0.298** | **0.208** | 0.151 | 0.322 |

Table 2: Automatic evaluation results on CX-CHR (upper part) and IU X-Ray Datasets (lower part). BLEU-n denotes BLEU score uses up to n-grams.

# Experiments: Results

| Dataset | CX-CHR | | | IU X-Ray | | |
|---|---|---|---|---|---|---|
| Models | Retrieval | Generation | HRGR-Agent | CNN-RNN [34] | CoAtt [16] | HRGR-Agent |
| Prec. (%) | 14.13 | 27.50 | **29.19** | 0.00 | 5.01 | **12.14** |
| AFP | 0.133 | 0.064 | **0.059** | 0.000 | **0.019** | 0.043 |
| Hit (%) | – | 23.42 | **52.32** | – | 28.00 | **48.00** |

Table 3: Average precision (Prec.) and average false positive (AFP) of medical abnormality terminology detection, and human evaluation (Hit). The higher Prec. and the lower AFP, the better.

# Conclusion

- The paper introduces a model that bridges tradicional retrieval-based approaches and modern sequence generating methods
- For each sentence a retrieval policy module determines if a template should be retrieved or a novel sentence should be generate from scratch
- The model is trained using RL, defining rewards on word and sentence levels
- HRGR-Agent achieves the state-of-the-art in two medical image report datasets

UNIVERSITY OF WATERLOO

Thank you!