NII_UIT at TRECVID 2021: Video Summarization Task

Khang Dinh Tran[1,2] , Nhat Pham Le Quang[1,2] , Tien Do[1,2] , Tien-Dung Mai[1,2] , An Pham Nguyen Truong[1,2] , **Duy-Dinh Le**[1,2] , Shin'ichi Satoh[3]

[1] *University of Information Technology (UIT), Ho Chi Minh City, Vietnam*
[2] *Vietnam National University, Ho Chi Minh City (VNU-HCM), Vietnam*
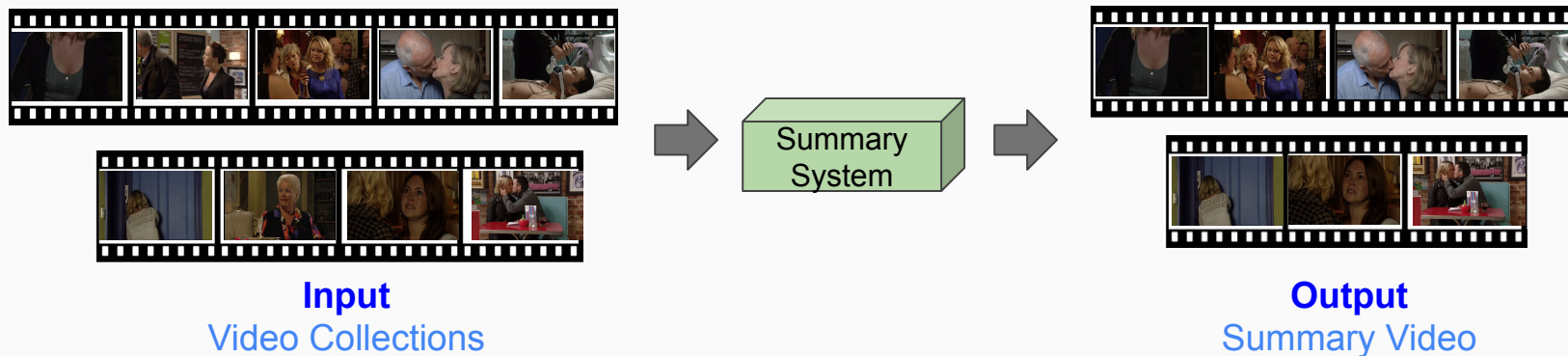[3] *National Institute of Informatics (NII), Tokyo, Japan*

TRECVID Workshop 2021

Overview

- VSUM Task Introduction
- Challenges
- Our Approach
- Experimental Results
- Conclusion & Future Work

VSUM Task Introduction

Introduction - Video Summarization

Input
Video Collections

Summary System

Output
Summary Video

- reduce the **size** → create a short **skim** of the original video
- concentrate the **high amount of high value information** → keep the most **important content**
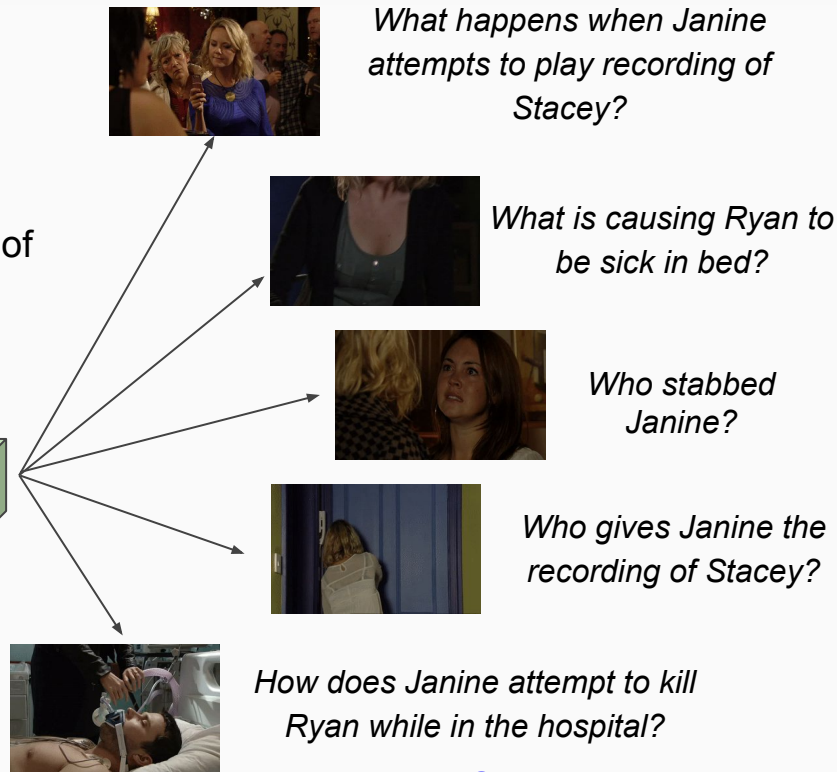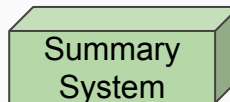
Important content

- ***Major life events*** of a given character:
  - the birth of a child (rather than a short illness),
  - a divorce (rather than an argument with a loved one)
- How they ***intertwine*** with the major life events of ***other specified characters***

Input

Summary System

What happens when Janine attempts to play recording of Stacey?

What is causing Ryan to be sick in bed?

Who stabbed Janine?

Who gives Janine the recording of Stacey?

How does Janine attempt to kill Ryan while in the hospital?

Output

Jack (2021)

Q1 - WHAT happens when <u>police</u> break in the door of <u>Jack and Tanya's home</u>?

Q2 - WHERE are Max and Jack during the violent <u>confrontation</u> between them when a <u>gun</u> is drawn?

Q3 - WHO does Jack offer to pay in order to withdraw their statement to the <u>police</u>?

Q4 - WHY is Jack a <u>suspect</u> in the <u>hit and run</u> on Max?

Q5 - WHAT does Jack <u>reveal</u> to Tanya about his dodgy past?
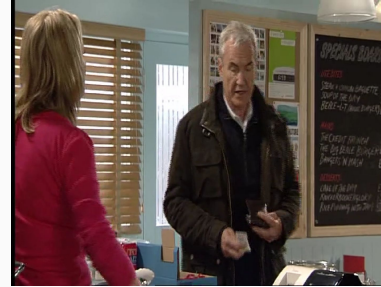
Max (2021)

Q1 - WHAT were the cause of Max's serious <u>injuries</u> which left him in <u>hospital</u>?

Q2 - WHAT is/was the <u>relationship </u>between Max and Tanya?

Q3 - WHAT kind of <u>weapon </u>does Max <u>obtain</u> from Phil?

Q4 - WHERE are Max and Jack during the violent confrontation between them when a <u>gun</u> is drawn?

Q5 - WHO is responsible, or who does Max believe is responsible, for the serious <u>injuries</u> which left him in <u>hospital</u>?

Tanya (2021)

Q1 - WHAT does Tanya <u>reveal</u> to the <u>police</u> while being <u>interviewed</u> at the <u>station</u>?

Q2 - WHAT is/was the <u>relationship</u> between Max and Tanya?

Q3 - WHAT does Jack <u>reveal</u> to Tanya about his dodgy past?

Q4 - WHAT does Tanya <u>discover </u>in <u>the sink</u> and on Jack's <u>clothes</u>?

Q5 - WHAT big move were Tanya and Jack planning for the future?

Archie (2021)

Q1 - WHAT happens when Phil <u>throws</u> Archie in to a <u>pit</u>?

Q2 - WHAT happens after Danielle <u>reveals</u> to Archie that Ronnie is her mother?

Q3 - WHERE do Peggy and Archie get <u>married</u>?

Q4 - WHAT happens when Archie <u>arrives</u> at the <u>pub</u> after Peggy invited him?

Q5 - WHAT happens when Archie is <u>kidnapped</u>?

Peggy (2021)

Q1 - WHO does Peggy ask to <u>kill</u> Archie?
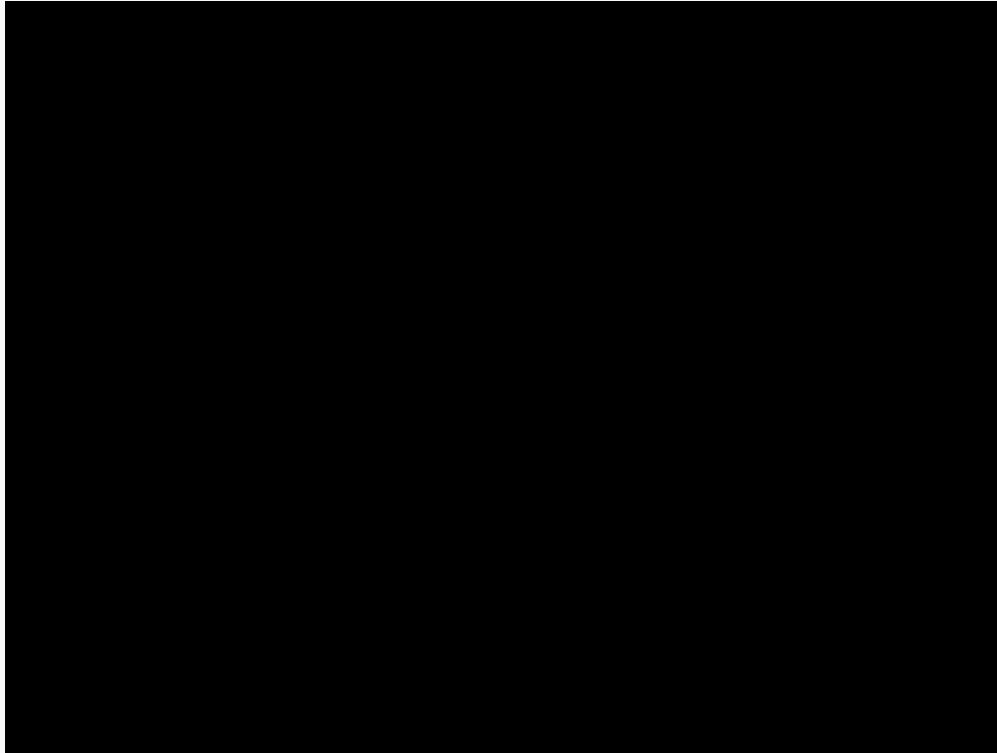
Q2 - WHERE do Peggy and Archie get <u>married</u>?

Q3 - Show one of the challenges WHICH Peggy faces in her election run.

Q4 - WHAT does Peggy overhear Archie saying, which causes their marriage to be over?

Q5 - WHAT is Janine doing to <u>irritate</u> or <u>anger</u> Peggy?

What does Tanya reveal to the police while being interviewed at the station?
What is/was the relationship between Max and Tanya?
What does Jack reveal to Tanya about his dodgy past?
What does Tanya discover in the sink and on Jack's clothes?
What big move were Tanya and Jack planning for the future?

Challenges

Q4 - Who stabbed Janine?

Stacey stabbed Janine

Challenge 1 - Visual Information

- **Visual information** does not always help
  - Different person → Janine
  - Hard to recognize → poison drink
  - There are cases that face recognition failed to recognize the target character and there is no appearance of the target character
- From the guideline → Hit means
  - An example of the video clips which would answer those questions.
  - Note that the *answer does not have to be specifically stated in the videos, just that they can be said to answer those questions.*

Q1 - What is causing Ryan to be sick in bed?

Janine gave Ryan a poison drink

Challenge 2 - Transcript

- Text from transcript only is not enough → Person (face) recognition is needed
  - many short shots → short sentences → difficult to match
  - several people might share same keywords such as giving birth (Stacey, Tanya)

Tanya's storylines have included opening her own salon, giving birth to her third child, discovering that her husband Max Branning (Jake Wood) had an affair with Stacey Slater (Lacey Turner) before divorcing and then burying him alive with the help from her new partner Sean Slater (Robert Kazinsky), marrying Greg Jessop (Stefan Booth), starting an affair with Max on her wedding day to Greg, dealing

Stacey was involved in another issue-led story exploring postpartum psychosis in 2016, following the birth of her son. Turner went on maternity leave in 2019 and prior to her exit, Stacey became involved in a custody battle with Kush Kazemi (Davood Ghadami) and a feud with her cousin Kat Slater (Jessie Wallace). After attacking Phil Mitchell (Steve McFadden), Stacey departs and is off-screen for thirteen

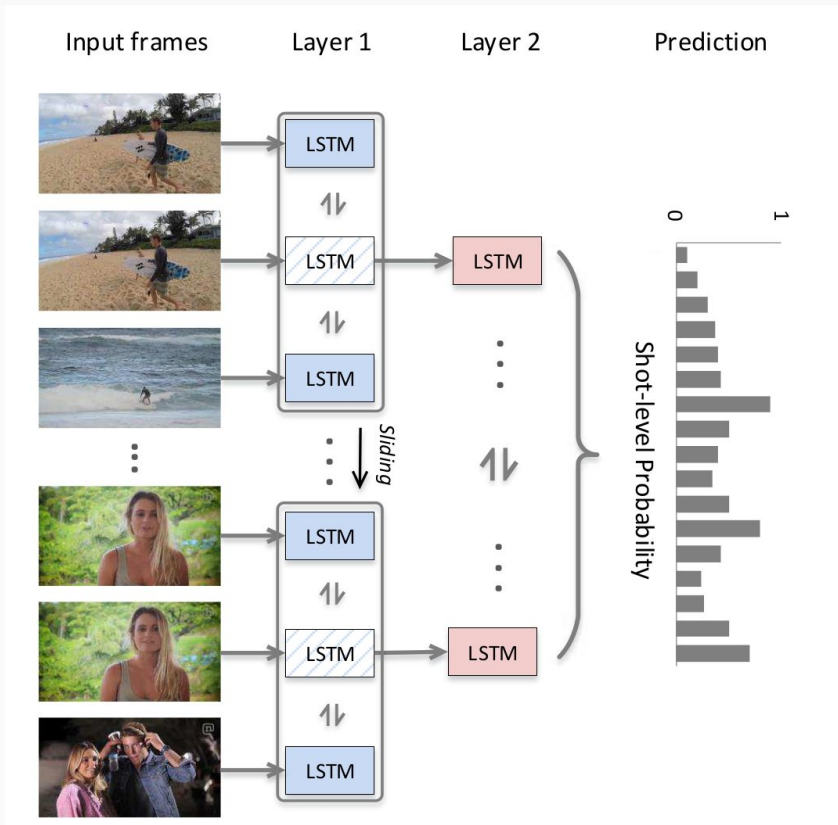- Where are <u>Max</u> and <u>Jack</u> during the violent confrontation between them when a gun is drawn?
- What does <u>Jack</u> reveal to <u>Tanya</u> about his dodgy past?
- What is/was the relationship between <u>Max</u> and <u>Tanya</u>?

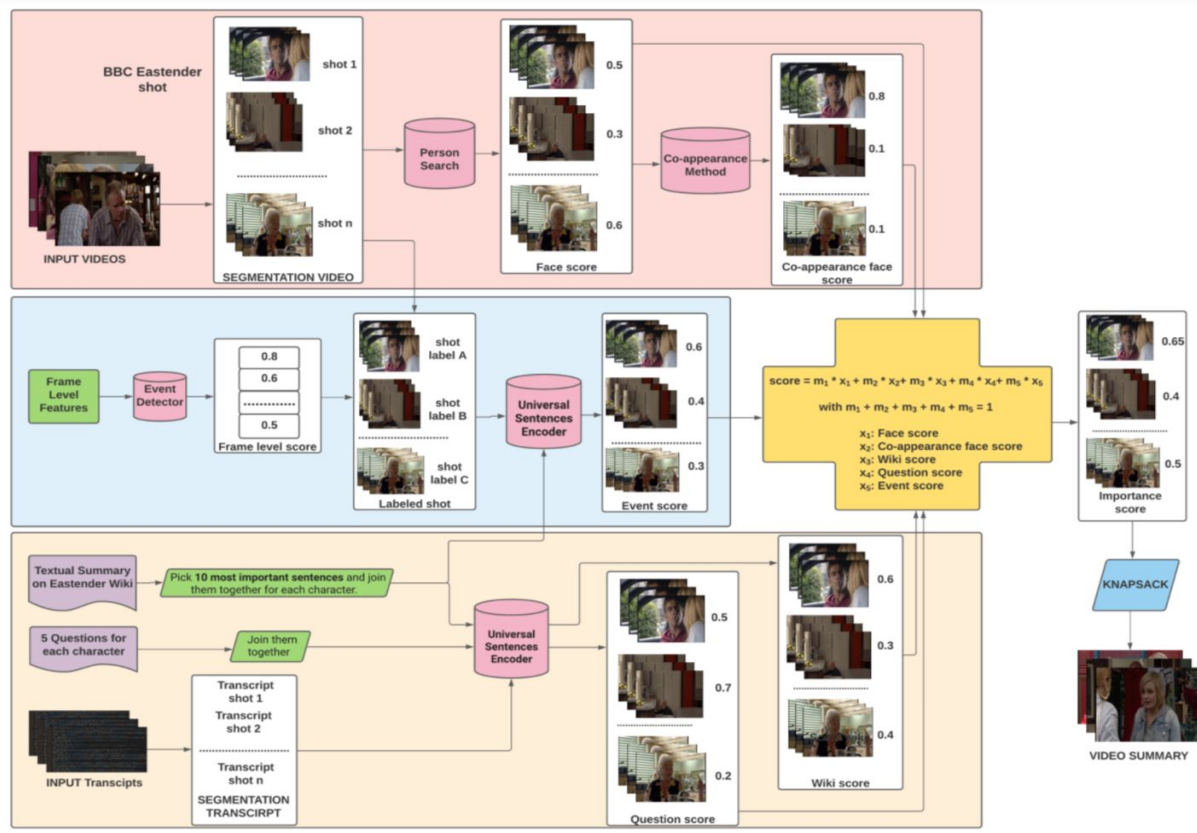|					
Max	What were the cause of Max's serious injuries which left him in hospital?	**What is/was the relationship between Max and Tanya?**	What kind of weapon does Max obtain from Phil?	**Where are Max and Jack during the violent confrontation between them when a gun is drawn?**	Who is responsible, or who does Max believe is responsible, for the serious injuries which left him in hospital?
Jack	What happens when police break in the door of Jack and Tanya's home?	**Where are Max and Jack during the violent confrontation between them when a gun is drawn?**	Who does Jack offer to pay in order to withdraw their statement to the police?	Why is Jack a suspect in the hit and run on Max?	**What does Jack reveal to Tanya about his dodgy past?**
Tanya	What does Tanya reveal to the police while being interviewed at the station?	**What is/was the relationship between Max and Tanya?**	**What does Jack reveal to Tanya about his dodgy past?**	What does Tanya discover in the sink and on Jack's clothes?	What big move were Tanya and Jack planning for the future?

- SumMe and TVSum datasets based supervised video summarization techniques → used for un-edited videos, short length (~ 2 mins) → main problems to solve are redundancy and visual interestingness
- No connection (available datasets) to important life events
- Building annotated datasets for supervised learning is expensive
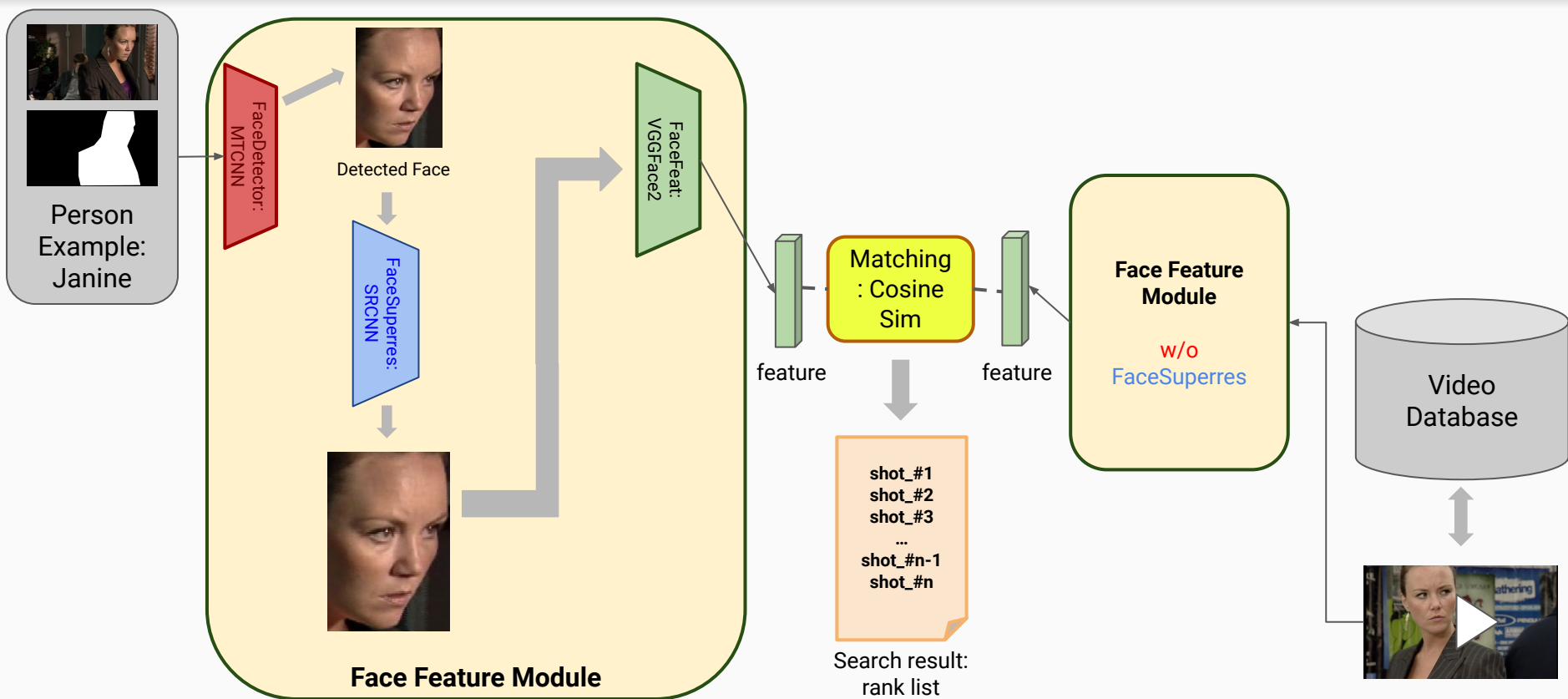
Our Approach

Framework Overview

Problems - Proposed Methods

- Face Search
  - Problem: How to find shots that the target people appear?
    - → *What does <u>Ryan</u> do when <u>Janine</u> is lying in the hospital?*
  - Solution:
    - Face Search from NII_UIT Instance Search Framework
    - Face Co-appearance Score
- Text Search
  - Problem: How to generate text queries
  - Solution:
    - Keywords → generate sentences → match with Wiki text → pick top 10 sentences → match with transcript
    - Image captioning
    - Questions (subtask)
- Fusion → finding appropriate weights

Face Search

Face Co-Appearance Score

- The major life events of a character often involve with one or several other characters. We used the Face Co-appearance Score to model the co-appearance of the person of interest and another character.

- The Face Co-appearance Score of the character A and B is calculated according to the following formula:

$$score_{appearanceAB} = score_{faceA} * score_{faceB}$$

Face Co-Appearance Score

- For example, the group includes three characters A, B and C, the appearance score of each character is calculated as follows:

$$score_{appearanceA} = \max\left(score_{appearanceAB}, score_{appearanceAC}\right)$$

$$score_{appearanceB} = \max\left(score_{appearanceAB}, score_{appearanceBC}\right)$$

$$score_{appearanceC} = \max\left(score_{appearanceAC}, score_{appearanceBC}\right)$$

- The group includes two characters A and B, the appearance score of each character is calculated as follows:

$$score_{appearanceA} = score_{appearanceB} = score_{appearanceAB}$$

Text Query Generation

- Selected keywords → Generate sentences
  - NLTK Context Free Grammar Generation: Keywords → Sentences
- Find the best matching between the generated sentences and sentences describing in Wikipedia
  - Universal Sentences Encoder → representation

- PropN = { 'Archie' | 'Peggy'|'Tanya'|'Max' | 'Jack'|}
- Verb = {'kill' | 'stab' | 'trap'|'reveal'| 'admits to'| 'confronts'|'give'| 'tell'|'attempt to kill' | 'confesses to'|}
- Adj = {'sick'|'angry' | 'frightened'}
- NP = {'hospital' | 'cell' | 'bed' |'restroom'}

Life Event - Tanya@Wikipedia

- From Wikipedia - https://en.wikipedia.org/wiki/Tanya_Branning
  - Tanya's storylines have included **opening her own salon**, **giving birth** to her third child, discovering that her husband Max Branning (Jake Wood) had an affair with Stacey Slater (Lacey Turner) before **divorcing** and then burying him alive with the help from her new partner Sean Slater (Robert Kazinsky), **marrying** Greg Jessop (Stefan Booth), **starting an affair** with Max on her **wedding** day to Greg, dealing with her eldest daughter, Lauren Branning's (Madeline Duggan/Jacqueline Jossa) alcoholism, **battling cervical cancer**, discovering that Max has a secret wife, Kirsty Branning (Kierston Wareing) on their second wedding day, witnessing her daughters Lauren and Abi Branning (Lorna Fitzgerald) fall from the roof of The Queen Victoria public house and subsequently dealing with Abi's **death**, which leads to her **mental breakdown**.

Example of Generated Text Query for Tanya

1. Tanya engaged in an affair with **married** man, Max Branning.
2. Max insults Tanya, and Tanya throws his brother Derek Branning out.
3. Tanya **admits** to the **police** that Tanya tried to **kill** Max; however, Lauren was the real culprit. (TANYA_Q1)
4. Tanya locks Lauren in the house but Tanya escapes through the upstairs window.
5. When Derek discovers that Tanya scared Tanya daughter Alice Branning away, Tanya threatens to **kill** Tanya if Tanya does it again.
6. Tanya **confronts** Cora, who refuses to talk about it, but eventually tells Tanya that Ava died not long after Tanya birth and it hurt so much Tanya never told anyone, not even the father.
7. Tanya goes to the **hospital** where Tanya **tells** Tanya doctor that wants no more treatment as Tanya does not have the strength to fight, despite the fact Tanya could die.
8. Tanya is diagnosed with cervical cancer and **tells** only Lauren and forbids Tanya from **telling** anyone else.
9. Lucy Beale **confronts** Max, who says Tanya was involved in criminal dealings causing someone to get hurt.
10. Derek in response taunts Tanya about what Tanya and Max did in Manchester.

Experimental Results

Result@Main Task

- Fusion weights
  - **NII_UIT_1**: 1.0*MatchingUsingFace
  - **NII_UIT_2:** 0.7*MatchingUsingFace + 0.3*MatchingUsingWiki
  - **NII_UIT_3:** 0.5*MatchingUsingFace + 0.5*MatchingUsingWiki
  - **NII_UIT_4:** 0.2*MatchingUsingFace + 0.7*MatchingUsingWiki + 0.1*MatchingUsingEventDetection

Result@Main Task

- Best run: NII_UIT_3 → balanced weights between MatchingUsingFace and MatchingUsingWiki
- NII_UIT_2 vs NII_UIT_4 → high MatchingUsingWiki weight is better

Main Task	Archie	Peggy	Jack	Max	Tanya	Sum	Ave
NII_UIT_1	0	0	0	0	0	0	0.00%
NII_UIT_2	1	0	0	0	1	2	8.00%
NII_UIT_3	1	1	1	1	2	6	24.00%
NII_UIT_4	0	0	1	1	2	4	16.00%
Solved	20.00%	20.00%	20.00%	20.00%	40.00%		24.00%

Result@Sub Task

- Fusion weights
  - **NII_UIT_1:** 0.25*MatchingUsingFace + 0.25*MatchingUsingFaceCoAppearance + 0.25*MatchingUsingWiki + 0.25*MatchingUsingQuestions
  - **NII_UIT_2:** 0.7*MatchingUsingFace + 0.3*MatchingUsingQuestions
  - **NII_UIT_3:** 0.5*MatchingUsingFace + 0.25*MatchingUsingWiki + 0.25*MatchingUsingQuestions
  - **NII_UIT_4:** 0.5*MatchingUsingFace + 0.1*MatchingUsingFaceCoAppearance + 0.2*MatchingUsingWiki + 0.2*MatchingUsingQuestions

Result@Sub Task

- Best run: NII_UIT_4 → **Answer ALL questions for Tanya**
- MatchingUsingQuestions is better than MatchingUsingWiki
- MatchingUsingFaceCoAppearance helps improvement for Max and Tanya

Sub Task	Archie	Peggy	Jack	Max	Tanya	Sum	Ave
NII_UIT_1	0	0	1	0	1	2	**8.00%**
NII_UIT_2	1	0	0	0	2	3	**12.00%**
NII_UIT_3	2	0	1	0	2	5	**20.00%**
NII_UIT_4	3	1	1	1	5	11	**44.00%**
Solved	**60.00%**	**20.00%**	**20.00%**	**20.00%**	**100.00%**		**44.00%**

Answered Questions - Best Run

Main Task (Best Run)	Archie	Peggy	Jack	Max	Tanya	Sum	Ave
ADAPT_1	3	0	1	1	1	6	24.00%
EURECOM_4	3	1	2	1	2	9	36.00%
NII_UIT_3	1	1	1	1	2	6	24.00%

Sub Task (Best Run)	Archie	Peggy	Jack	Max	Tanya	Sum	Ave
ADAPT_4	0	0	1	0	1	2	8.00%
EURECOM_4	0	2	2	1	3	8	32.00%
NII_UIT_4	3	1	1	1	5	11	44.00%

Answered Questions - Max / Unique

- Max: 16/25 questions can be answered
- Unique answered questions
  - **ADAPT: 2** (Peggy-Q2 = Archie-Q3 → Where do Peggy and Archie get married?)
  - **EURECOM: 2** (Jack-Q5 → What does Jack reveal to Tanya about his dodgy past?, Max-Q3 → What kind of weapon does Max obtain from Phil?)
  - **NII_UIT: 1** (Archie-Q4 → What happens when Archie arrives at the pub after Peggy invited him?)

Team (Max Answered Questions)	Archie	Peggy	Jack	Max	Tanya	Sum	Ave
ADAPT	4	3	1	1	4	13	**52.00%**
EURECOM	3	2	2	2	4	13	**52.00%**
NII_UIT	4	2	1	1	5	13	**52.00%**
Max Answered Questions	**4**	**3**	**2**	**2**	**5**	**16**	**64.00%**

Conclusion & Future Work

Conclusion

- A simple baseline is proposed for VSUM Task
  - Fusion of Face Co-Appearance and Text Matching
  - The best run of the sub task solves 44% of questions
- Evaluation should be automatic in finding hits → fine annotations are needed

Future Work

- Image Matching
  - Given a list of keywords such as wedding, search for images using Google Image Search
  - Use these images to retrieve relevant shots
- Better transcript matching
- Better fusion methods