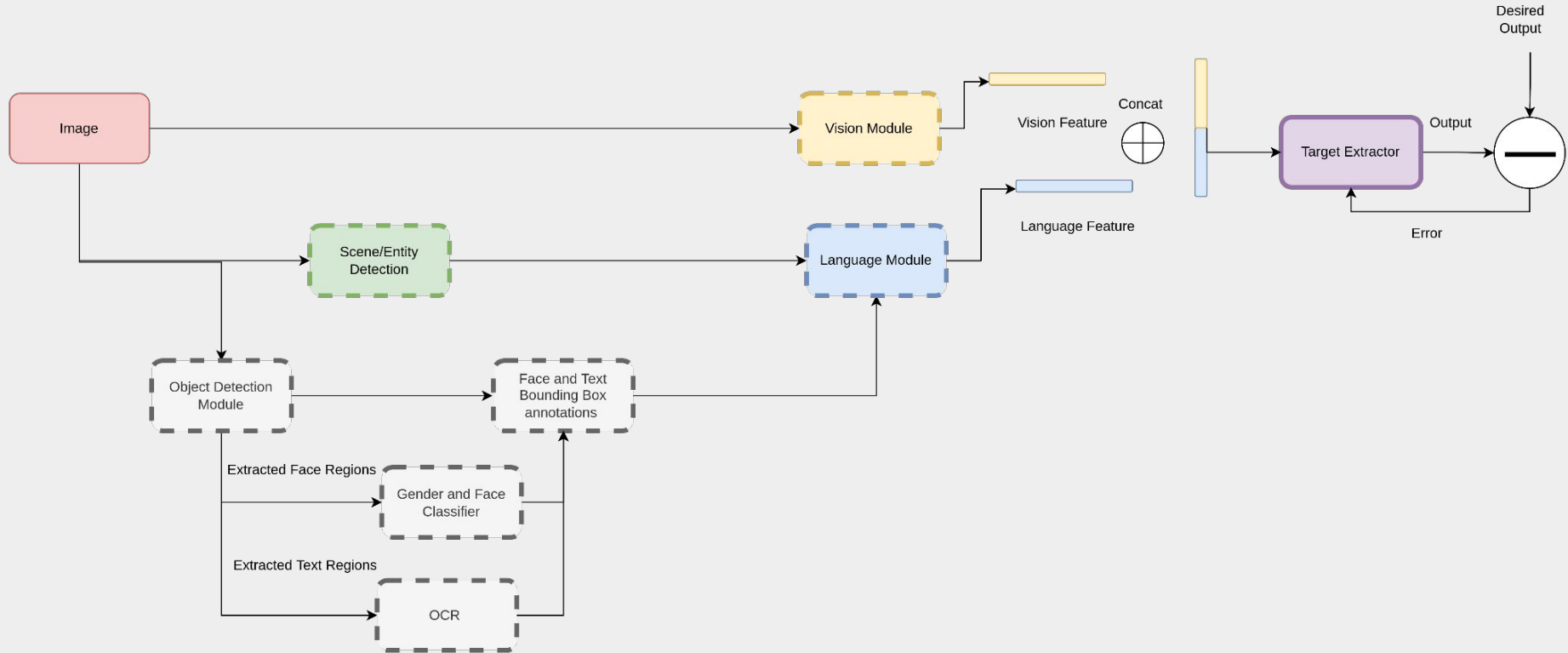


Executive Summary

Context	<ol style="list-style-type: none">1. Memes are the third-most shared type of content on social media, behind only personal news and funny videos, as per a survey of US and UK internet users base.2. Memes can combine harmful captions with innocent images in a malicious manner. Most of the time, the meme's hateful part is linked by subtle wording or external knowledge about the real-world event.3. Memes are highly complex, contextual content which present significant challenges in the content moderation.
Objective	<ol style="list-style-type: none">1. Create an algorithm that identifies multimodal hate speech related to race or ethnicity in internet memes that can help in the content moderation task.
Key Benefits	<ol style="list-style-type: none">1. Reduce the need for human moderation in the pre-moderation stage and reduce the impact on them of viewing harmful content.2. Cheaper and easier to scale as the volume of content uploaded increases as compared to expensive human moderators.

End to End Learning for Hate Speech Detection System

Scenario-1



Scene and Entity Detection

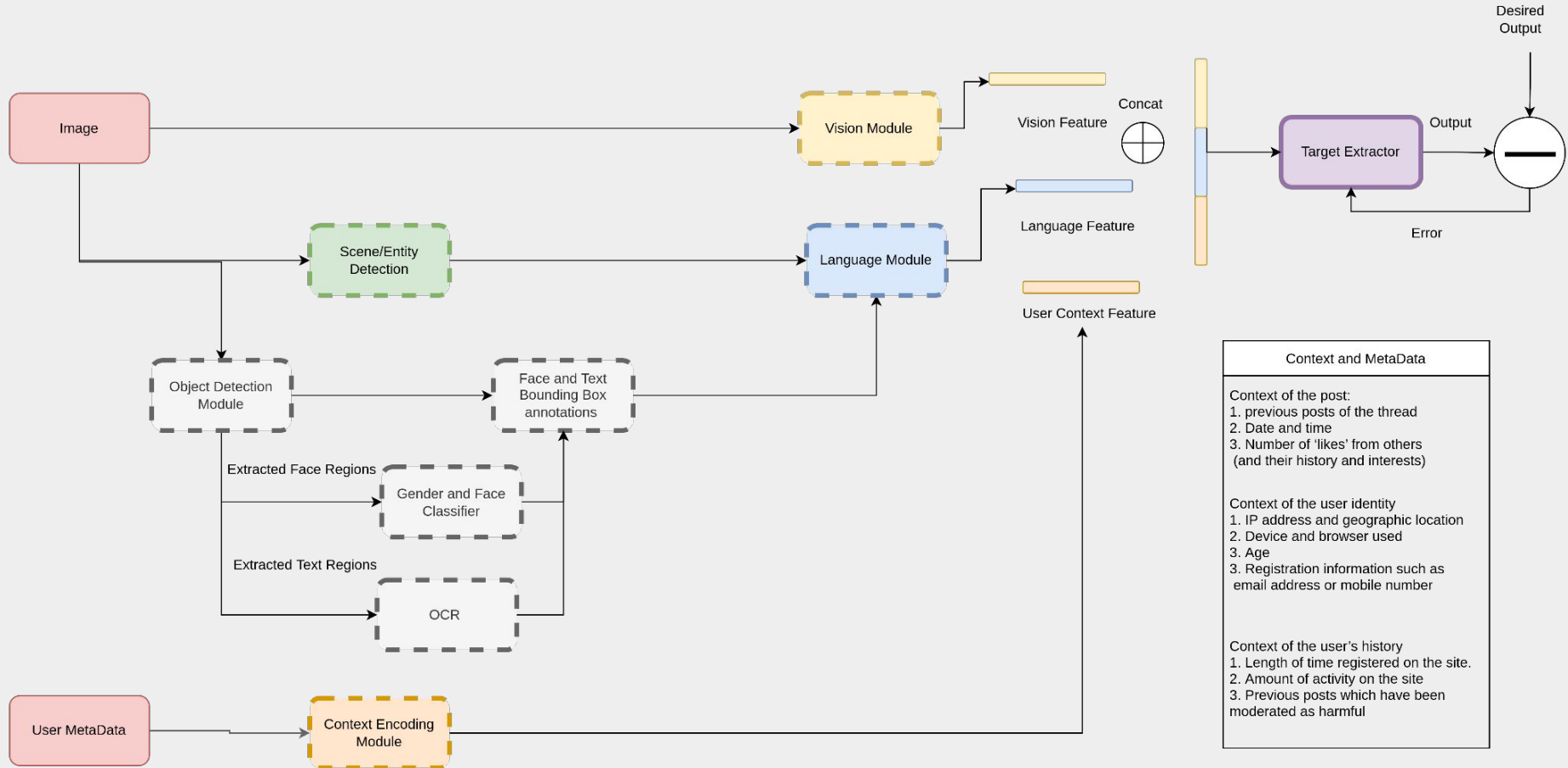


Image credit: [Quinten de Graaf](#) on [Unsplash](#).

Category	Responses
Web entities	<ul style="list-style-type: none">• <i>entityId</i>: /m/02p7_j8, <i>score</i>: 1.3225499, <i>description</i>: Carnival in Rio de Janeiro• <i>entityId</i>: /m/06gmr, <i>score</i>: 1.1684971, <i>description</i>: Rio de Janeiro• <i>entityId</i>: /m/04cx88, <i>score</i>: 1.05945, <i>description</i>: Brazilian Carnival• ...
Full matching images	<ul style="list-style-type: none">• <i>url</i>: https://1000lugaresparair.files.wordpress.com/2017/11/quinten-de-graaf-278848.jpg• <i>url</i>: https://freewalkingtourrotterdam.com/wp-content/uploads/2017/07/quinten-de-graaf-278848.jpg• ...
Partial matching images	<ul style="list-style-type: none">• <i>url</i>: https://www.linnanneito.fi/wp-content/uploads/sambakarnevaali-riossa.jpg• <i>url</i>: https://static.airhelp.com/wp-content/uploads/2019/02/26105557/two-women-in-carnival-costumes.jpg• ...
Pages with matching images	<ul style="list-style-type: none">• <i>url</i>: https://travelnoire.com/best-carnival-celebrations-around-the-world/, <i>pageTitle</i>: Best \u003cb\u003eCarnival\u003c/b\u003e Celebrations Around The World - Travel Noire, <i>fullMatchingImages</i>: [{<i>url</i>: https://travelnoire.com/wp-content/uploads/2019/02/quinten-de-graaf-278848-unsplash.jpg}]• <i>url</i>: https://bespokebrazil.com/rio-carnival-2019/, <i>pageTitle</i>: Visit \u003cb\u003eRio Carnival 2019\u003c/b\u003e with the Brazil Specialists - Bespoke

End to End Learning for Hate Speech Detection System

Scenario-2



Modules to explore

1. Automated way of clustering images for faster data tagging of rare examples.
2. Trying out Visual Linguistic class of pretrained model like VL Bert that can encode the text and image together to better align the visual-linguistic clues during training.