A woman with a tennis racket about to swing at a ball .

woman   tennis racket   swing   ball
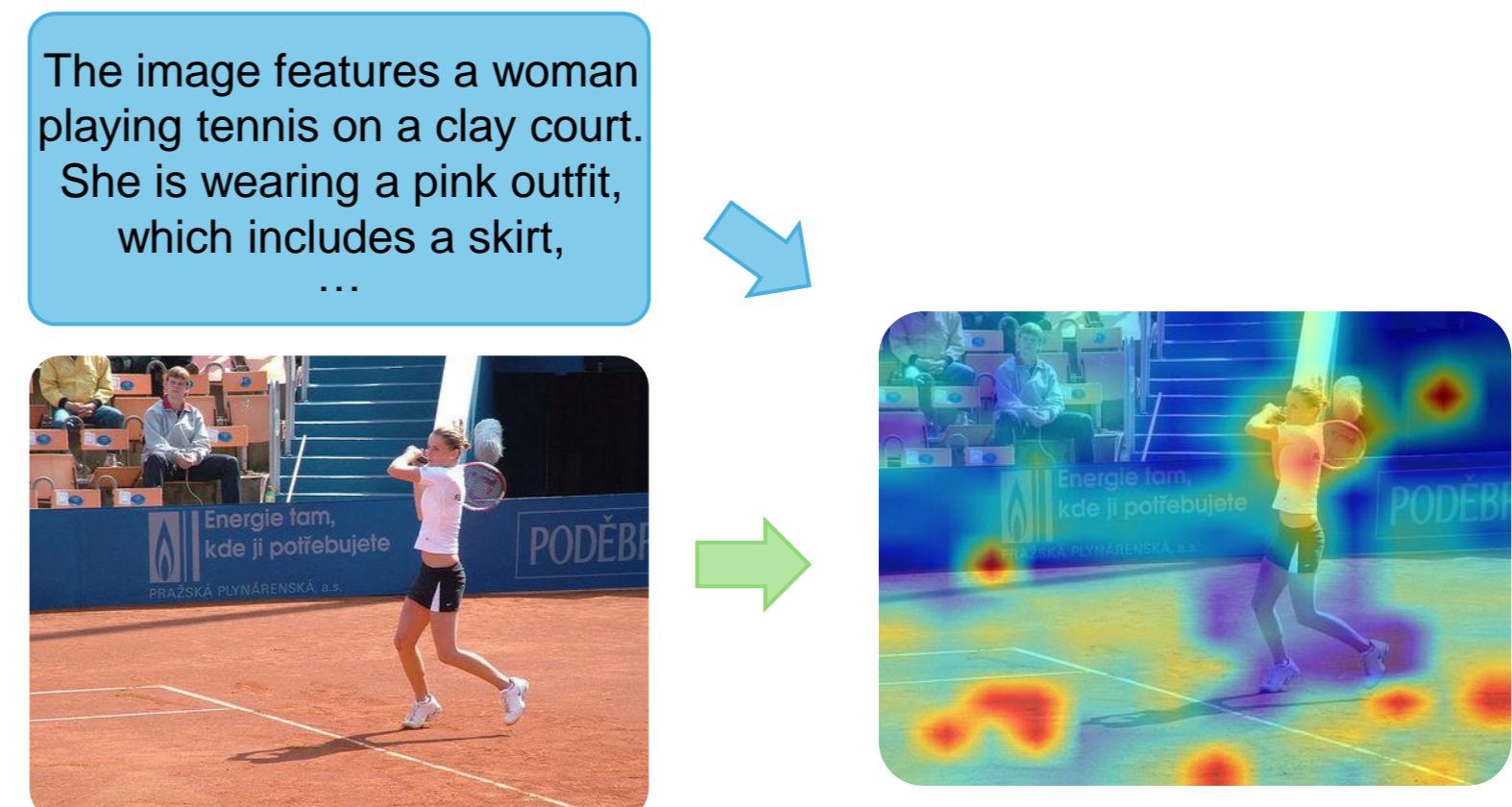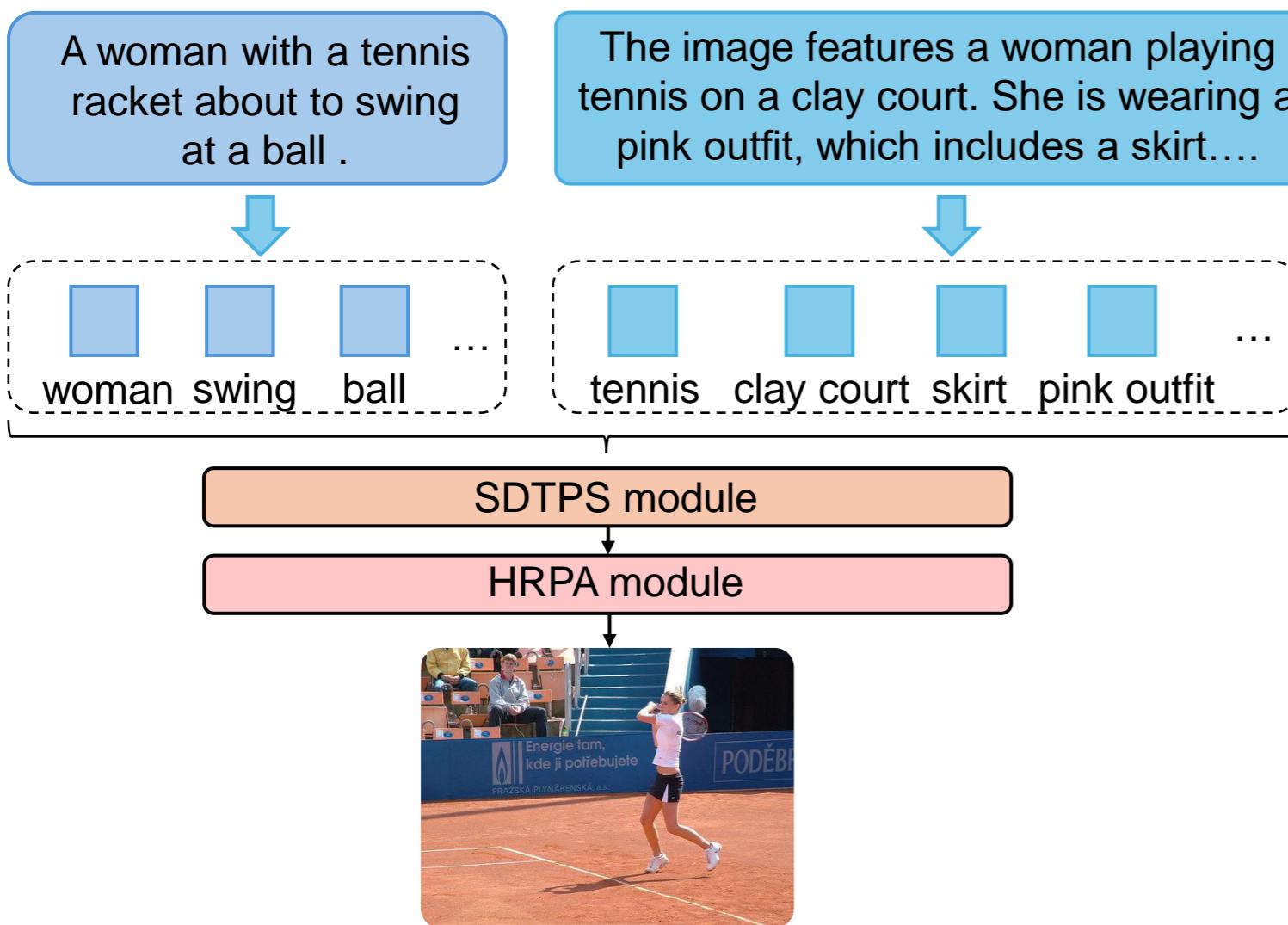
?   Patch Ambiguity

Patch Redundancy

A woman with a tennis racket about to swing at a ball .

(a) The limitation of semantic gap between image and sparse text

A woman with a tennis racket about to swing at a ball .

The image features a woman playing tennis on a clay court. She is wearing a pink outfit, which includes a skirt....

woman   swing   ball   ...

tennis   clay court   skirt   pink outfit   ...

SDTPS module

HRPA module

The image features a woman playing tennis on a clay court. She is wearing a pink outfit, which includes a skirt, ...

A woman with a tennis racket about to swing at a ball .

(b) SEPS framework bridge semantic gap through leveraging dense text