Model and Hypotheses

```
library(conflicted)

library(kableExtra)
```

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'
```

```
library(glue)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
```

```
conflict_prefer("filter", "dplyr")
```

```
## [conflicted] Will prefer dplyr::filter over any other package
```

```
models <- tribble(
  ~Detectors, ~Stages, ~mAP, ~Vanishing, ~Mislabeling, ~Untargeted,
  "YOLOv3", 1, 33.7, "Object", "Class", "Class, Box, Object",
  "SSD", 1, 29.5, "Class", "Class", "Class, Box",
  "RetinaNet", 1, 36.5, "Class", "Class", "Class, Box",
  "Faster R-CNN", 2, 37.4, "RPN: Object; Det: Class", "Det: Class", "RPN: Object, Box; Det: Class, Box"
  "Cascade R-CNN", 2, 40.3, "RPN 1: Object; RPNs 2, 3 + Det: Class", "RPNs 2, 3: Class; Det: Class", "R
)

models
```

```
## # A tibble: 5 x 6
##   Detectors     Stages   mAP Vanishing                          Misla~1 Untar~2
##   <chr>          <dbl> <dbl> <chr>                              <chr>   <chr>
## 1 YOLOv3             1  33.7 Object                             Class   Class,~
## 2 SSD                1  29.5 Class                              Class   Class,~
## 3 RetinaNet          1  36.5 Class                              Class   Class,~
## 4 Faster R-CNN       2  37.4 RPN: Object; Det: Class           Det: C~ RPN: O~
## 5 Cascade R-CNN      2  40.3 RPN 1: Object; RPNs 2, 3 + Det: Cl~ RPNs 2~ RPN 1:~
## # ... with abbreviated variable names 1: Mislabeling, 2: Untargeted
```

```
note_alpha <- function(txt, num, double_escape = FALSE) {
  glue("{txt}{footnote_marker_alphabet(num, double_escape = double_escape)}")
}

add_linebreak <- function(data) {
  data |> mutate(across(everything(), linebreak))
}
```

```
break_models <- function(col) {
  str_replace_all(col, coll(c("; " = ";\n", " R-CNN" = "\nR-CNN")))
}

models <- models |>
  rename('{note_alpha("Stages", 1)}' := Stages, '{note_alpha("COCO mAP", 2)}' := mAP, '{note_alpha("Unta
  mutate(across(everything(), break_models)) |>
  add_linebreak()

models
```

```
## # A tibble: 5 x 6
##   Detectors                          Stages\\te~1 COCO ~2 Vanis~3 Misla~4 Untar~5
##   <chr>                              <chr>        <chr>   <chr>   <chr>   <chr>
## 1 "YOLOv3"                           1            33.7    "Objec~ "Class" "Class~
## 2 "SSD"                              1            29.5    "Class" "Class" "Class~
## 3 "RetinaNet"                        1            36.5    "Class" "Class" "Class~
## 4 "\\makecell[l]{Faster\\\\R-CNN}"   2            37.4    "\\mak~ "Det: ~ "\\mak~
## 5 "\\makecell[l]{Cascade\\\\R-CNN}" 2            40.3    "\\mak~ "\\mak~ "\\mak~
## # ... with abbreviated variable names 1: `Stages\\textsuperscript{a}`,
## #   2: `COCO mAP\\textsuperscript{b}`, 3: Vanishing, 4: Mislabeling,
## #   5: `Untargeted\\textsuperscript{d}`
```

```
attack_header <- c(3, 3)
names(attack_header) <- c(" ", note_alpha("Attack Losses", 3, double_escape = TRUE))

models |>
  kbl(
    booktabs = TRUE,
    escape = FALSE,
    caption = "Detection models and attack losses. Full details are given in Appendix \\ref{app:mod_los
  ) |>
  kable_styling(
    position = "center",
    latex_options = "striped"
  ) |>
  add_header_above(c(" " = 3, "Targeted" = 2, " " = 1)) |>
  add_header_above(attack_header, escape = FALSE) |>
  footnote(
    alphabet = c(
      "In general, 1-stage detectors are quicker whereas 2-stage detectors are more accurate, though th
      "COCO mean Average Precision (mAP) is the primary metric on the COCO challenge.",
      "The training losses in detectors typically include the box regression loss (Box), the class loss
      "Untargeted attack targets all training losses in a model, i.e.\ the backpropagation loss."
    ),
    threeparttable = TRUE
  )
```

```
sub_results <- function(col) {
  str_replace_all(col, coll(c(">" = "$>$")))
}

results <- tribble(
  ~hp, ~sig,
  "1-stage > 2-stage models (YOLOv3, SSD, RetinaNet > Faster R-CNN, Cascade R-CNN)", "All except Retinal
```

Table 1: Detection models and attack losses. Full details are given in Appendix ??.

			Attack Losses[c]		
			Targeted		
Detectors	Stages[a]	COCO mAP[b]	Vanishing	Mislabeling	Untargeted[d]
YOLOv3	1	33.7	Object	Class	Class, Box, Object
SSD	1	29.5	Class	Class	Class, Box
RetinaNet	1	36.5	Class	Class	Class, Box
Faster R-CNN	2	37.4	RPN: Object; Det: Class	Det: Class	RPN: Object, Box; Det: Class, Box
Cascade R-CNN	2	40.3	RPN 1: Object; RPNs 2, 3 + Det: Class	RPNs 2, 3: Class; Det: Class	RPN 1: Object, Box; RPNs 2, 3 + Det: Class, Box

[a] In general, 1-stage detectors are quicker whereas 2-stage detectors are more accurate, though the 1-stage RetinaNet aims to be both quick and accurate. In a 2-stage detector, the input image passes through a Region Proposal Network (RPN) stage and a detection (Det) stage.

[b] COCO mean Average Precision (mAP) is the primary metric on the COCO challenge.

[c] The training losses in detectors typically include the box regression loss (Box), the class loss on the 80 COCO labels and/or the background class (Class), and the objectness loss on categorizing an image region as background or object (Object).

[d] Untargeted attack targets all training losses in a model, i.e. the backpropagation loss.

```
  "Targeted > Untargeted attack", "Only YOLOv3, SSD",
  "Vanishing > Mislabeling attack", "All except YOLOv3",
  "Larger attack iterations", "All",
  "Less confident targets", "All",
  "Larger perturb boxes", "All except mislabeling attack on Faster R-CNN",
  "Shorter perturb-target distance", "All",
  "Less accurate target COCO class", "Mixed",
  "More probable intended class (mislabeling attack only)", "None except RetinaNet",
  glue("Lower target {note_alpha('IOU', 2)} (untargeted attack only)"), "All"
)

results |>
  rename("Hypotheses (higher success for)" := hp, '{note_alpha("Accepted (across attacks and models)",
  mutate(across(everything(), sub_results)) |>
  add_linebreak() |>
  kbl(
    booktabs = TRUE,
    escape = FALSE,
    caption = "Hypothesis testing in the randomized attack (Sections \\ref{sec:rand_hp} and \\ref{sec:ra
  ) |>
  kable_styling(
    position = "center",
    latex_options = c(
      "striped",
      "hold_position"
    ),
  ) |>
  column_spec(2, width = "1.5in") |>
  column_spec(1, width = "1.25in") |>
```

```
footnote(
  alphabet = c(
    "$p < .05$ for Wald z-test on logistic estimate",
    "intersection-over-union"
  ),
  escape = FALSE
)
```

Table 2: Hypothesis testing in the randomized attack (Sections **??** and **??**)

Hypotheses (higher success for)	Accepted (across attacks and models)[a]
1-stage > 2-stage models (YOLOv3, SSD, RetinaNet > Faster R-CNN, Cascade R-CNN)	All except RetinaNet (YOLOv3, SSD > RetinaNet, Faster R-CNN, Cascade R-CNN)
Targeted > Untargeted attack	Only YOLOv3, SSD
Vanishing > Mislabeling attack	All except YOLOv3
Larger attack iterations	All
Less confident targets	All
Larger perturb boxes	All except mislabeling attack on Faster R-CNN
Shorter perturb-target distance	All
Less accurate target COCO class	Mixed
More probable intended class (mislabeling attack only)	None except RetinaNet
Lower target IOU[b] (untargeted attack only)	All

[a] $p < .05$ for Wald z-test on logistic estimate
[b] intersection-over-union