This supplemental table summarizes the used ML approach, the basis of the ground truth assignment, data set size and data variance, as well as the intended design space of the final ML model for each application example discussed in the paper.

Goal	Reference	ML approach	Ground truth	Data set size	Data variance	Design space
Segmentation ferrite vs. pearlite/ bainite/ martensite/tempered martensite in two-phase-steels (LOM and SEM images)	Azimi et al. [1]	DL: Semantic Segmentation	Easy separation of foreground class (ferrite) vs. background classes (pearlite, bainite, martensite, tempered martensite) Differentiation of pearlite/ bainite/ martensite/tempered martensite based on round robin consensus and sample processing history	2205 images of second phase objects (4 different classes)	Medium variance	Robust model application possible with SEM images More data needed to improve the performance with LOM images
Classification of bainitic subclasses in two-phase-steels (SEM images)	Müller et al. [47]	Conventional ML: object-wise feature extraction (image textural features) + support vector machine classification	Expert consensus from round robin tests + correlative characterization (SEM + EBSD)	3903 images of second phase objects (7 different classes)	Low variance	Tailored protocol for sample contrasting and image acquisition to guarantee the data quality needed for this sophisticated classification (i.e., five bainite classes to be distinguished)
Segmentation of lath-shaped bainite in multi-phase steels (LOM and SEM images)	Durmaz et al. [25]	DL: Semantic Segmentation	Correlative characterization (LOM or SEM + EBSD)	754 patches extracted from 51 LOM images 413 patches extracted from 36 SEM images	Low variance	Tailored for one specific analysis pipeline in quality control (i.e., same microstructure type, sample etching, same microscope)
Segmentation of upper bainite, lower bainite, martensite, tempered martensite in quenched and quenched and tempered steels (LOM and SEM images)	Bachmann et al. [30]	DL: Classification with CNN	Correlative characterization (LOM + SEM + EBSD)	6500 patches extracted from SEM images 2200 patches extracted from LOM images	Medium variance	Tailored protocol for sample contrasting and image acquisition to guarantee the data quality needed for this sophisticated classification
Segmentation of prior austenite grains after picric acid based etching	Laub et al. [92]	DL: Semantic Segmentation	Correlative characterization (LOM + EBSD)	8000 patches extracted from 30 LOM images	High variance	Model robust against etching conditions, etching artefacts, microscope settings, grain sizes
Segmentation of prior austenite grains after Nital etching	Bachmann et al. [46]	DL: Semantic Segmentation	Correlative characterization (LOM + SEM + EBSD)	1420 patches extracted from 13 LOM images	Medium variance	Model robust against etching conditions, etching artefacts, microscope settings, grain sizes; Certain basic level of contrasting of PAG must be visible