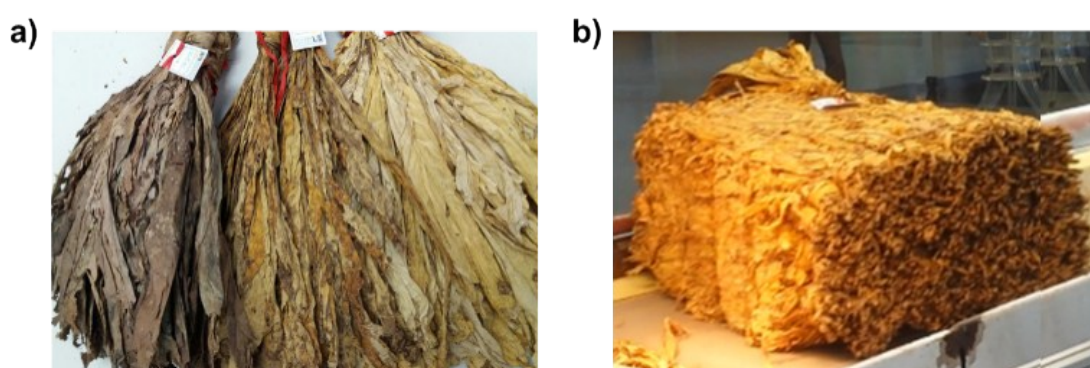


## Electronic Supplementary Information (ESI)

### Fast Inline Tobacco Classification by Near-Infrared Hyperspectral Imaging and Supporting Vector Machines-Discriminant Analysis

Marcelo C. A. Marcelo<sup>a</sup>, Frederico L. F. Soares<sup>a</sup>, Jorge A. Ardila<sup>a</sup>, Jailson C. Dias<sup>a</sup>, Ricardo Pedó<sup>a</sup>, Samuel Kaiser<sup>\*a</sup>, Oscar F. S. Pontes<sup>a</sup>, Carlos E. Pulcinelli<sup>a</sup> and Guilherme P. Sabin<sup>a</sup>



**Figure S1.** a) Bundle of flue-cured Virginia leaves with different colour classes and b) Tobacco bale of flue-cured Virginia at the leaf threshing facility.

**Table S1.** Analyzed flue-cured Virginia dataset stratified by stalk position, leaf colour and leaf quality.

Classification System	Class	Number of Bundles	Number of Bales
Stalk Position	Leaf	906	26374
	Cutters	791	12770
	Lugs	694	23294
	Primings	649	7535
Leaf Colour	Pale Lemon	488	2133
	Lemon	59	3472
	Orange	1726	51317
	Light Mahogany	143	1899

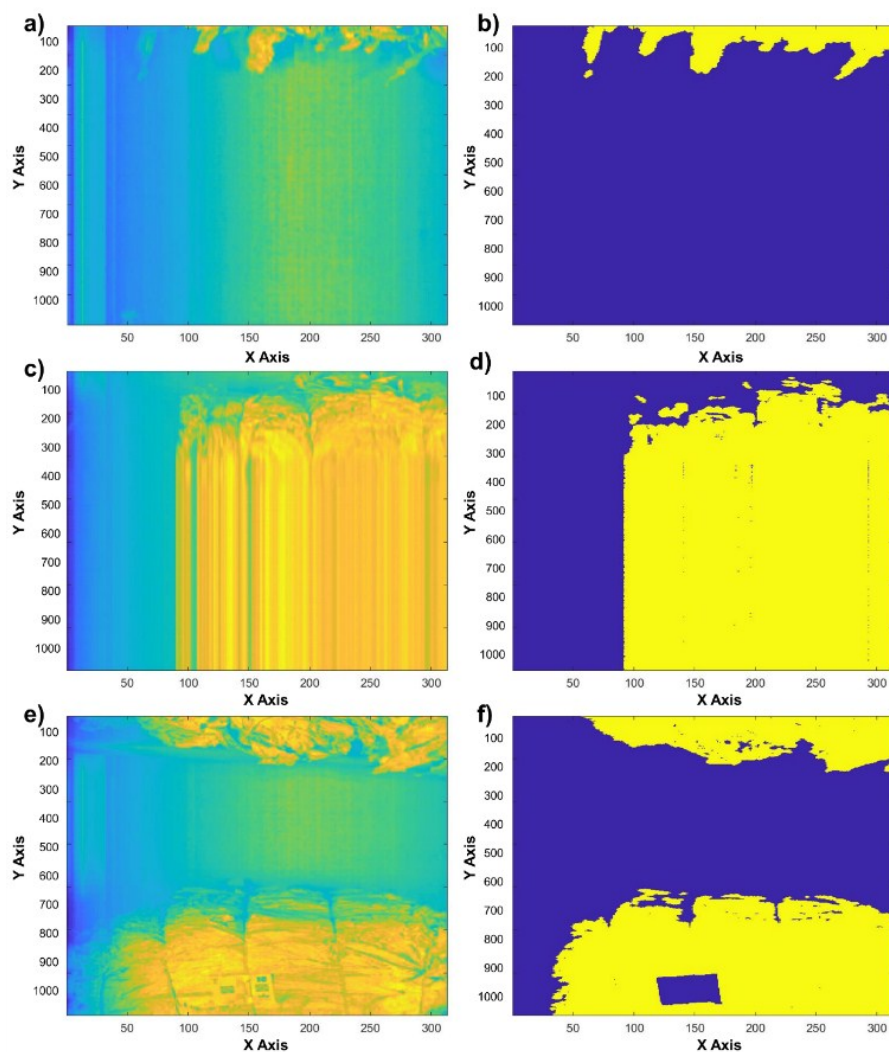
	Mahogany	624	11152
Leaf Quality	High quality, homogeneous and high ripeness	101	103
	High quality and homogeneous	128	411
	High quality, high ripeness	320	2413
	High quality	517	19577
	High quality, low ripeness	64	395
	Medium quality, thick leaf	75	1811
	Medium quality and high ripeness	249	6564
	Medium quality and extreme thick leaf	133	5264
	Medium quality and low ripeness	27	-
	Medium quality	24	738
	Medium-low quality	153	9712
	Medium-low quality and high ripeness	64	4092
	Medium-low quality and low ripeness	122	4009
	Medium-low and heterogeneous	403	10694
	Low quality and ripe defects	269	1413
	Low quality and high ripeness	49	376
	Low quality and overripe	165	2085
	Low quality and unripe	59	163
	Low quality and greenish	54	153
	Low quality, dark, overripe and brittle	64	-

**Table S2.** Analysed air-cured Burley dataset stratified by stalk position, leaf colour and leaf quality.

Classification	Class	Number of	Number of
----------------	-------	-----------	-----------

System		Bundles	Bales
Stalk Position	Tips	12 <sup>a</sup>	134 <sup>a</sup>
	Leaf	279	11522
	Cutters	281	8712
	Lugs	210	8120
	Primings	230	1907
Leaf Colour	Light Mahogany	347	7633
	Mahogany	467	22610
	Dark Mahogany	198	152 <sup>a</sup>
Leaf Quality	High quality, homogeneous and high ripeness	44	201
	High quality, high ripeness	161	872
	High quality	276	14203
	Medium quality, thick leaf	49	3811
	Medium quality and extreme thick leaf	10 <sup>a</sup>	-
	Medium quality and heterogeneous	186	10345
	Low quality and ripeness defects	237	752
	Low quality and unripe	10 <sup>a</sup>	-
	Low quality and greenish	39	211

a: class not included in the model due to low number of samples.



**Figure S2.** Corrected images at 1528.8 nm (a, c and e) and ROI selection (b, d and f) of samples with issues in acquisition: a) and b) Mostly belt conveyor image; c) and d) Sample stopped under the camera; e) and f) Two bales at the same image.