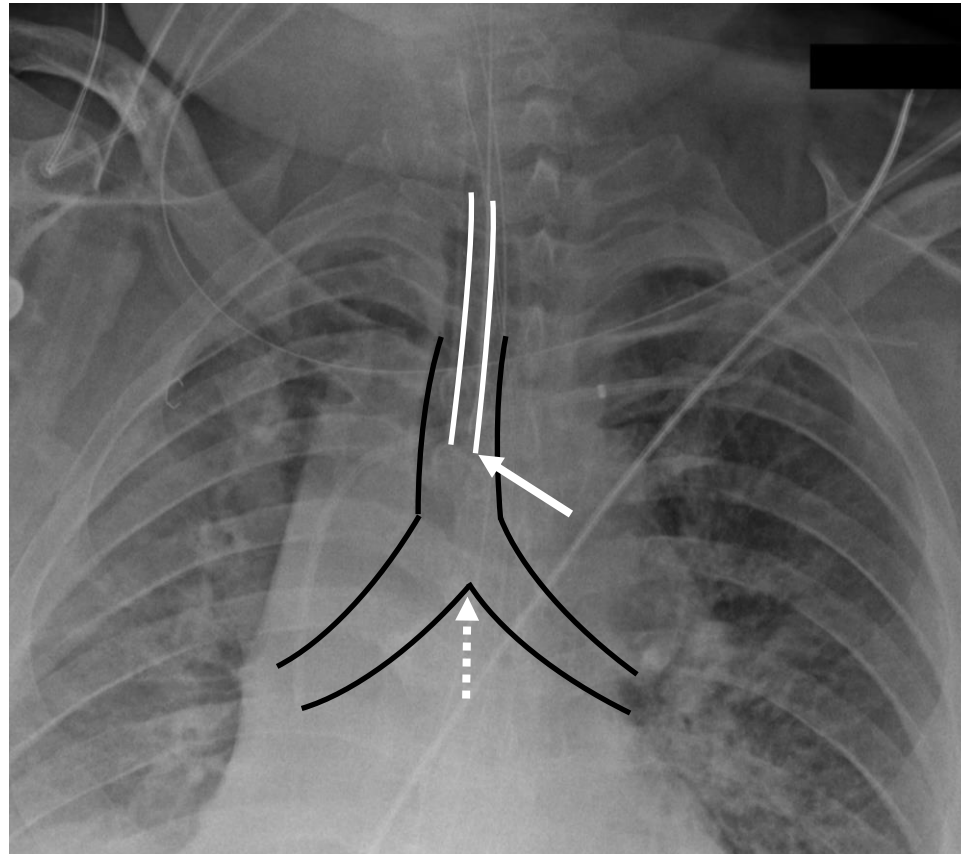


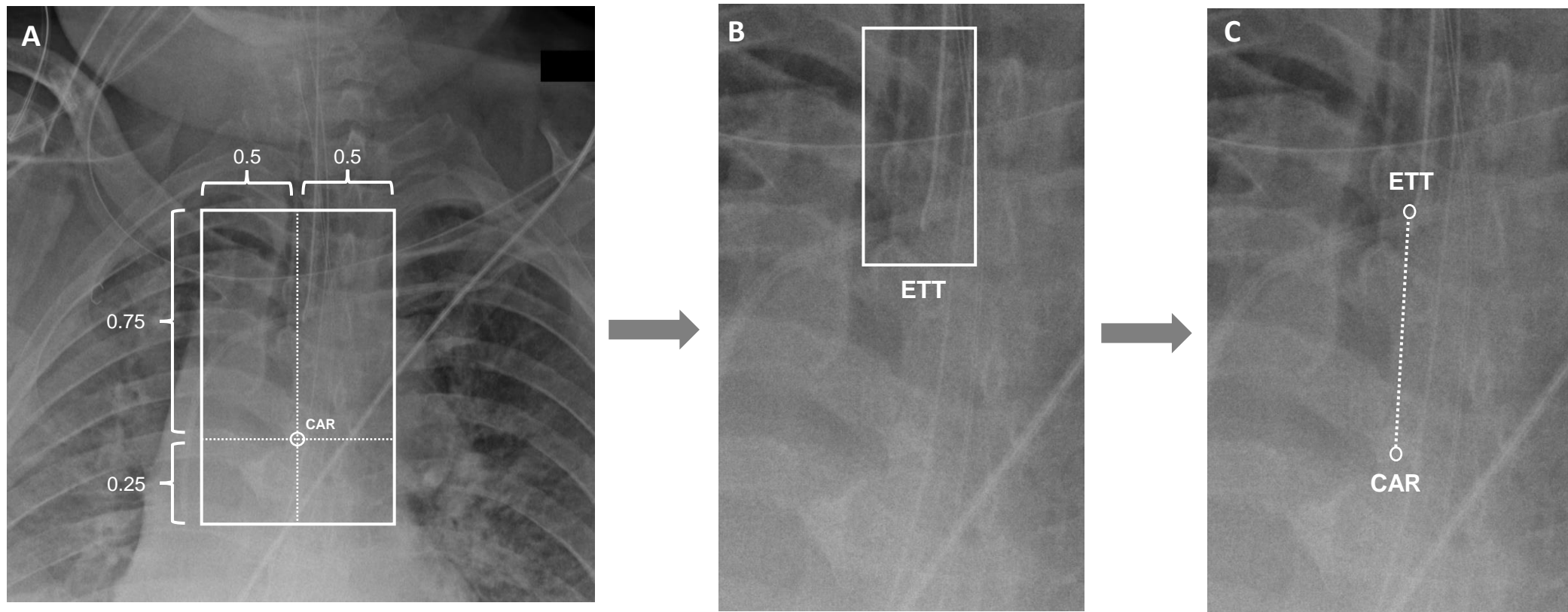
**Fig. 1: Frontal Chest Radiograph**

Single frontal radiograph demonstrating position of an endotracheal tube (ETT; solid white outline) relative to the trachea and proximal right and left bronchi (solid black outline). Optimal positioning is commonly evaluated by characterizing the distal ETT tip (solid white arrow) relative to the carina (dotted white arrow).



**Fig. 2: Serial Multi-Step CNN Architecture**

A total of three different CNN algorithms were created. (A) The first algorithm comprised of a regression loss function network designed to output the estimated  $(y, x)$  coordinate of the carina. From this coordinate position, a  $256 \times 128$  crop of the upper airway was generated. (B) Using this cropped input, a second CNN binary classifier network was used to predict the presence or absence of ETT. (C) Finally, for radiographs containing an ETT, a third CNN regression network was designed to output the estimated  $(y, x)$  coordinates of the carina and the tip of the distal ETT. Abbreviations: CAR (carina), ETT (endotracheal tube).

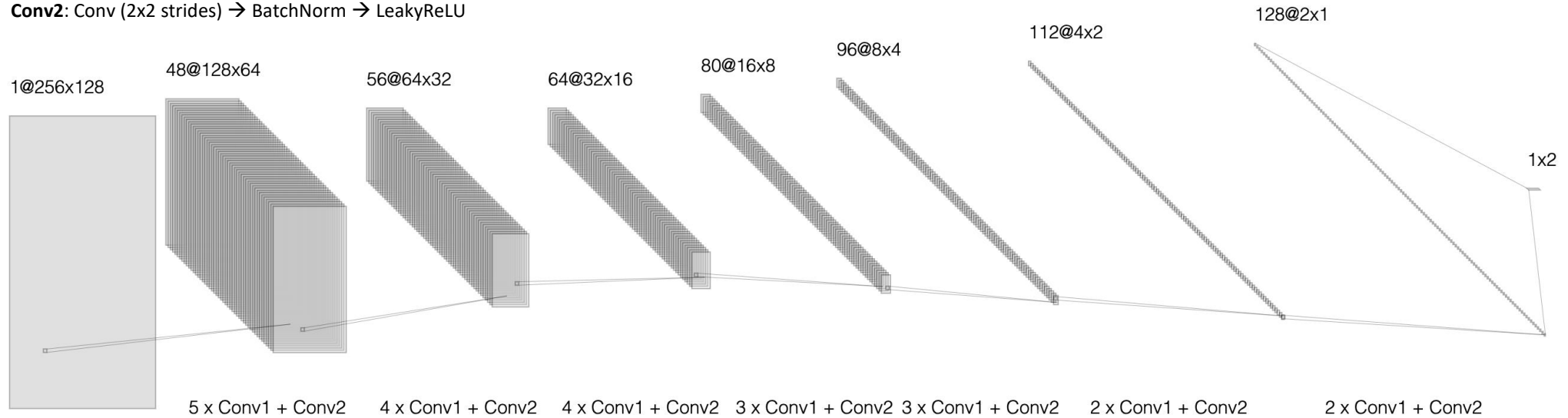


**Fig. 3: Common CNN Backbone**

For ease of model development, a common CNN backbone architecture was used for all three target tasks, with slight modifications in input matrix shape and loss function between each network. The final network topology comprised a total of 30 convolutional blocks, defined as serial application of a 3 x 3 convolutional operation, batch normalization and a LeakyReLU activation function. Subsampling was implemented via convolutional operations with a stride of 2. A total of 7 subsampling operations were performed; with each decrease in feature map size the corresponding channel depth was increased from 48 to 128.

**Conv1:** Conv (no strides) → BatchNorm → LeakyReLU

**Conv2:** Conv (2x2 strides) → BatchNorm → LeakyReLU



**Fig. 4: CNN Model Predictions**

Estimate of carina location by initial first-step CNN algorithm (A), which is subsequently refined by the final CNN algorithm (B). In addition, the final CNN algorithm produces an estimate of the coordinate location of the distal ETT. Manually annotated positions are shown as the white circles, while the CNN derived estimates are shown as the black circles.

