

Determining image usability and femoral head coverage using artificial intelligence: A future avenue in newborn hip screening

Singh Abhinav, Stamper Andrew, Perry Daniel, Hemmadi Sandeep, Voiculescu Irina

Background: Developmental dysplasia of the hip (DDH) represents the largest cause of hip osteoarthritis in young adults. Late diagnosis leads to multiple operations and persistently poor outcomes. The UK screening programme has not improved early DDH detection since 1986. Artificial Intelligence (AI) assisted diagnostics represents an evolving field with the potential to improve accuracy and objectivity of DDH screening.

Methods: This pilot study utilised an anonymised dataset obtained by experienced radiologists during routine NHS practice. The dataset contained 190 2D hip ultrasound images (normal= 71, dysplastic= 66, dislocated= 53) from 100 multi-ethnic babies (aged 4-12 weeks). Two clinical experts provided the anatomical ground truth masks for ilium, femoral head, and labrum. They used literature-derived femoral head coverage (FHC%) ranges to classify screening diagnosis - normal (>50%) or abnormal (<50%). We designed a multiclass convolutional neural network (modified U-Net, 94:47:49 training:validation:testing data split) to produce anatomical masks for classifying clinically usable images, calculating FHC% and providing a screening diagnosis.

Results: Within the test dataset (n=49), ilium, femoral head and labrum were identified in all images. We measured the pixel overlap per structure between ground truth and AI-derived masks. Mean pixel overlap sensitivities were - ilium (88.9% range 73.6-99.3%), femoral head (98.2% range 92.3-100%) and labrum (72.7% range 3.7-100%). Automated FHC% diagnosis agreed with clinical experts in 89.8% (n=44) of the images with a sensitivity of 90.6% and specificity of 88.2%.

Conclusions: Our AI successfully identifies key features and generates masks to decide whether an ultrasound image is clinically usable. The algorithm-generated diagnosis agrees with expert clinicians on 89.8% of images, exceeding the existing state of the art AI performance.

Implications: When translated into the clinical setting, this method will reduce reporting workload, and improve accuracy and objectivity of DDH screening.

Disclosure: None.