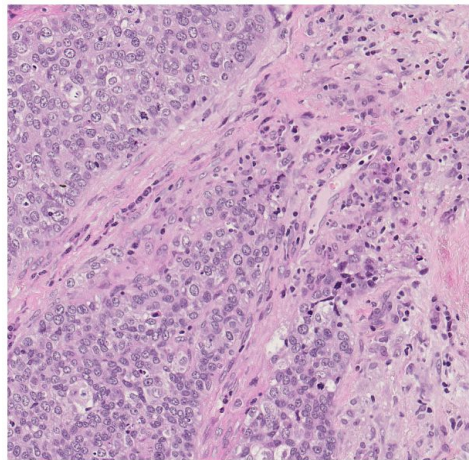


# Förutsäga utfallen av patienter med trippelnegativa bröstcancer med hjälp av maskininlärning

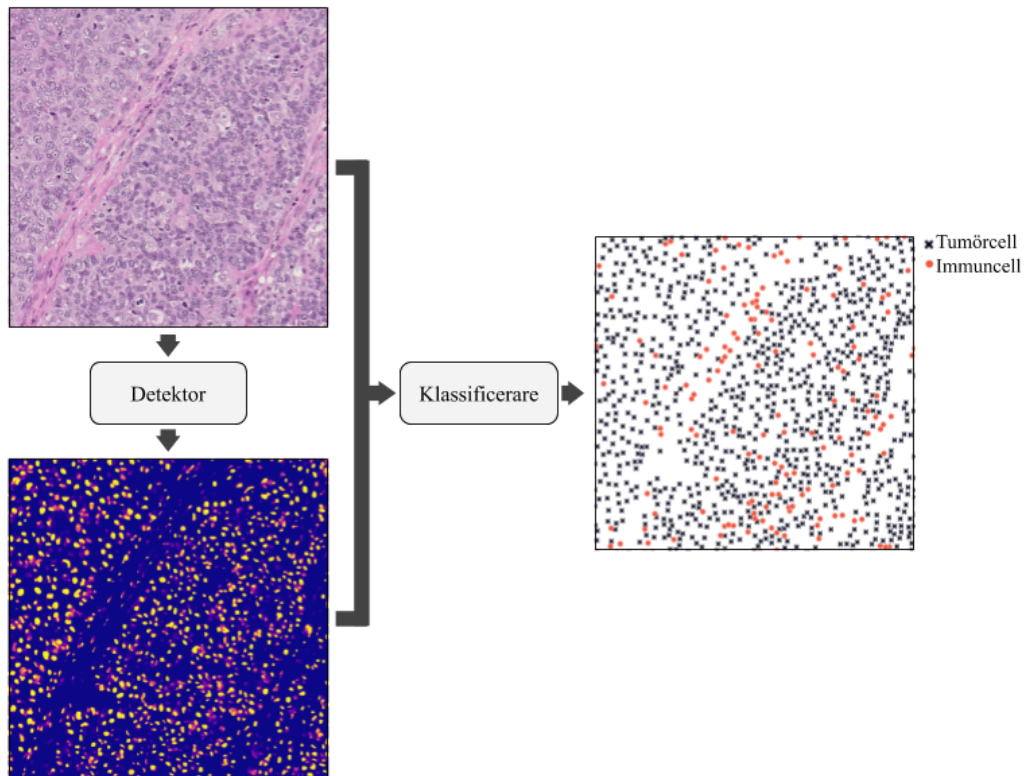
## En populärvetenskaplig sammanfattning

Bröstcancer är den vanligaste cancerformen bland kvinnor. Trippelnegativ bröstcancer är en särskilt aggressiv typ av bröstcancer som utgör 10 % av alla bröstcancerfall. Cancerformen är aggressiv i den mening att den drabbar fler unga och att fjärrmetastaser utvecklas i högre grad jämfört med andra sorters bröstcancer. Den vanligast behandlingen är att tumören opereras bort. Unga och i övrigt friska patienter kan sedan få behandling i form av strålbehandling och kemoterapi för att minska sannolikheten att canceren kommer tillbaka. Dessa behandlingar kommer med en mängd bieffekter och det är därför viktigt att inte behandla fler än nödvändigt. Syftet med det här projektet var att hitta en metod för att förutsäga hur illa det kommer gå för en patient, baserat på vävnadsbilder från den bortopererade tumören. En sådan metod skulle kunna användas för att undvika överbehandling av patienter som inte hade behövt det. Detta är ett exempel på hur en vävnadsbild kan se ut:



Vävnaden har färgats med hematoxylin och eosin, vilket gör det möjligt att urskilja tumörcellerna och immuncellerna. Tumörcellerna är de ljusa stora cellerna i nedre vänstra hörnet medan immuncellerna är de små mörka cellerna som går att hitta i det övre högra hörnet. Immuncellerna utgör en del av kroppens inbyggda försvar mot cancer, och en stor mängd av dessa celler har visat sig indikera en god prognos.

Med detta som utgångspunkt har en metod utformats som bland annat gör det möjligt att räkna antalet immunceller automatiskt. Processen sker i flera steg. En celldetektor avgör var i bilden cellerna finns. När cellernas platser är kända används en cellklassificerare för att avgöra vilka celler som är tumörceller och vilka som är immunceller. En överblick över processen kan ses nedan.



Cellernas mängd och placering används sedan för att träna en modell som kan förutsäga utfall för patienten.

Celldetektorn består av ett neutralt nätverk av typen U-Net. Nätverket tar in en bild på vävnad och avgör vilka pixlar i bilden som tillhör celler och vilka som utgör bakgrunden. Denna process kallas för bildsegmentering. U-Net innehåller tiotusentals parametrar. Värdet på parametrarna bestäms genom att nätverket tränas på färdiga exempel. Ju fler exempel modellen tränas på desto bättre blir resultatet.

Cellklassificeraren är en modell av typen logistisk regression och innehåller endast fyra parametrar. Denna modell tränas också utifrån exempel. Modellen avgör om en cell är en tumör- eller immuncell baserat på dess ljushet, ljushetsvarians och storlek.

När cellernas platser och klass är kända kan dessa användas som bas för en mängd modeller som får tävla mot varandra om vilken som bäst kan förutspå patientutfallen. Modellernas prestanda jämförs även mot modeller som är tränade på faktorer som är uppskattade av experter.

Resultaten visade att antalet immunceller som räknats av maskininlärningsmodellerna är en bättre bas för utfallsprediktionen jämfört med när man låtit experter uppskatta mängden immunceller. Detta indikerar att maskininläring skulle kunna hjälpa läkare och forskare att fatta bättre beslut.