

Populärvetenskaplig sammanfattning av examensarbetet:

Ett första steg mot en algoritm för förutsägelse av omoperation av bröstcancer med hjälp av maskininlärning och mammografibilder.

Student: Emma Jönsson

Handledare: Ida Arvidsson, Jennie Karlsson

Examinator: Anders Heyden

Ett första steg mot att förutsäga omoperation av bröstcancer med artificiell intelligens

Bröstcancer är den vanligaste formen av cancer hos kvinnor. Mellan 2008 och 2020 genomgick 9% av de som var diagnostiserade med bröstcancer en omoperation. Genom att använda artificiell intelligens för att förutse om en omoperation kommer behövas kan man förhoppningsvis sänka förekomsten av omoperationer.

Bröstcancer står för 30% av alla cancerfall hos kvinnor. En vanlig behandling av bröstcancer är bröstbevarande operation, där en bit av bröstet kirurgiskt tas bort. Operationer är både dyrt och har en betydande inverkan på kroppen. För en del kvinnor med en bröstcancerdiagnos krävs dessutom en omoperation. Syftet med arbetet har varit att undersöka om det är möjligt att förutsäga om en person utgör en större risk för omoperation genom att använda artificiell intelligens och mammografibilder.

Artificiell intelligens är en datoralgoritm som försöker efterlikna människans förmåga att hantera information och kan användas inom många områden. Inom medicinsk bildanalys har användandet av neurala nätverk ökat, vilket är en underkategori till artificiell intelligens. Neurala nätverk är ett samlingsnamn på olika självlärande algoritmer som försöker efterlikna funktionen i de biologiska neurala nätverken. Ett vanligt sätt att lära ett nätverk att klassificera bilder är genom att visa bilder och berättar vad som visas. Av detta kan nätverket lära sig och dra slutsatser om vad en ej tidigare sedd bild visar.

Arbetet är ett första steg för att kunna klassificera och förutse omoperationer av bröstcancer. Då bilder med information om omoperationer var svåra att få tag i användes istället bilder med benigna och maligna tumörer från två olika

databaser. (1) *The Chinese Mammography Database* med totalt 2142 bilder och (2) *Curated Breast Imaging Subset of Digital Database for Screening Mammography* med totalt 327 bilder. För att klassificera bilderna som benign och malign skapades 2 olika typer av neurala nätverk. Det totala antalet bilder på 2469 är ganska få när det kommer till att träna neurala nätverk. För att skapa fler bilder vändes bilderna i horisontellt plan, samt roterades och zoomades. De två olika nätverken tränades med både data från enbart den kinesiska databasen samt med data från båda databaserna.

Resultaten visade att fler bilder blev rätt klassificerat om enbart en databas användes i träningen istället för att använda båda. Detta beror antagligen på att bilderna i de två olika databaserna ser olika ut, samt att den ena databasen har lite text på bilderna. En annan sak som kunde konstateras var att resultaten i detta arbete var lägre än andra arbeten med samma data. Detta beror antagligen på att detta arbete använde hela mammografibilder medan liknande arbete har segmenterat ut den intressanta delen av bilden.

Den bästa noggrannheten på klassificering som uppnåddes var 62%. Genom att optimera nätverket ytterligare och använda data med rätt information skulle det kunna vara möjligt att förutse omoperationer i framtiden.