

Merimiinojen tunnistus pohjakartoituskuvista MATINE hanke

Prof. Jukka Heikkonen

Turun yliopisto, Tietotekniikan laitos

Hankkeelle myönnetty MATINE rahoitus: 94994 euroa

Turun yliopiston omarahoitus: 23 748

Hankkeen kesto: 1.2.2023-31.12.2024

Tutkimuksen tavoitteista

Tutkimuksessa kehitetään koneoppimismenetelmiä tunnistamaan pohjakartoituskuvista miinat (sodanaikaiset/harjoitusmiinat) ja erottamaan nämä luotettavasti Itämeren pohjalla makaavista muista kohteista (esim. kivet, kalliot, hylt)

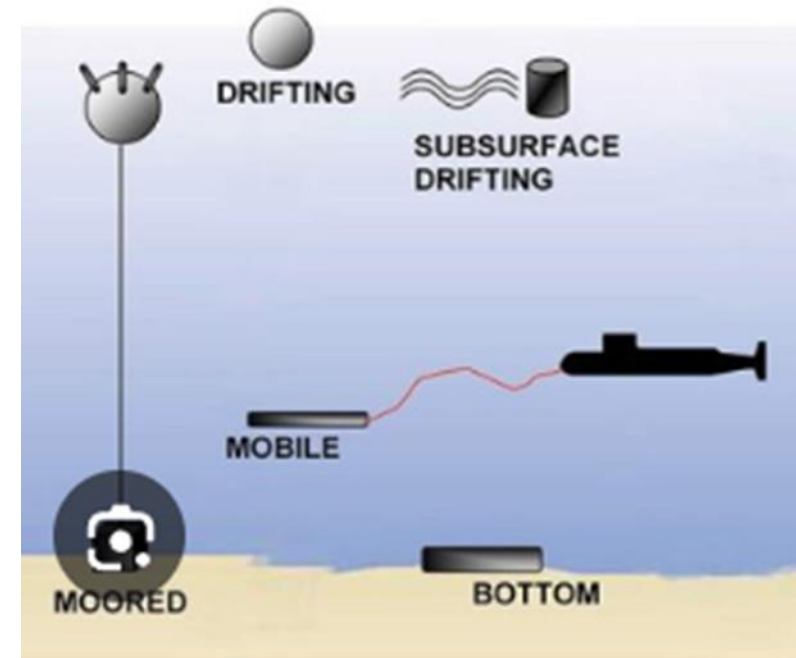
Tutkimushanke tukee Merivoimien väylä- ja meriturvallisuuden kehittämisstrategiaan ja tälle tutkimukselle on selvä tarve **nopeuttamaan** miinanetsintää sekä tekemään se nykyistä paljon **kustannustehokkaammin** ja **luotettavammin**.

Tunnistamismenetelmiä tullaan testataan Merivoimien keräämiin uusiin pohjakartoituskuviin:

- Tuloksia verrataan miinanetsintäoperaattoreiden tunnistamistuloksiin tarkkuuden ja tunnistamisnopeuden suhteen.

Tutkimuksen taustaa

- Sodanaikaisia miinoja Itämeressä arviolta yli 40000
- Iso osa vaarattomia
- Jos ankkuroidaan ja kalastetaan voi syntyä ongelmia
- Itämeren, matalan ja kivikkoinen ja olemassa olemat algoritmit toimivat huonosti:
 - **Varsinkin kivet tunnistetaan usein virheellisesti miinoiksi**
- Pohjamiina on profiililtaan usein sylinterimäinen, kosketusmiinan koho pallon tai ellipsin muotoinen ja kosketusmiinan ankkuri kulmikas.



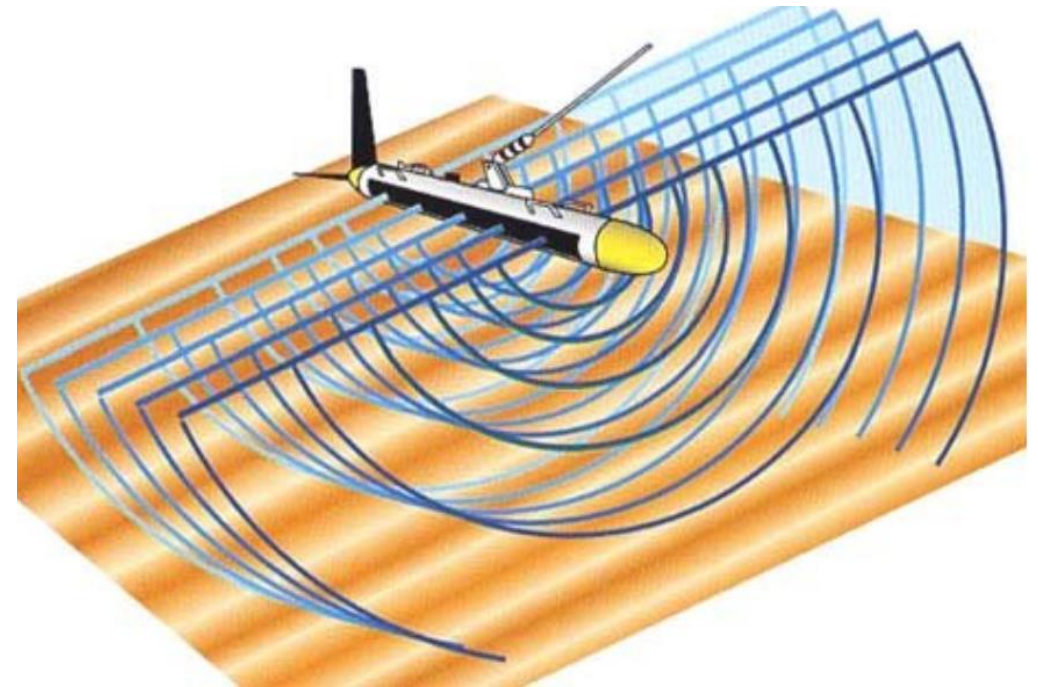
Pohjakartoituskuvat

Kuvantamislaitteistona Klein 5500 viistokaikuluotain:

- Vedettävä
- Laite lähettää ääniaaltoja (5-20kHz) veteen ja rekisteröi niiden paluukaikut.
- Paluukaikujen perusteella laite luo kuvan vedenalaisesta ympäristöstä.
- 5 samanaikaista sädettä -> 100% peitto

Veden akustiset ominaisuudet vaikuttavat mittausten tarkkuuteen

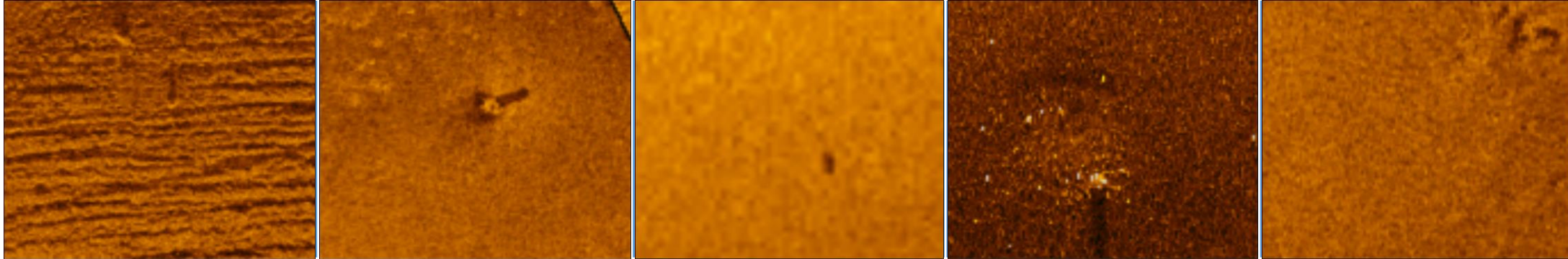
Itämeren mataluus tuottaa häiriöitä, paluukaikuja tulee kivistä ja kallioista.



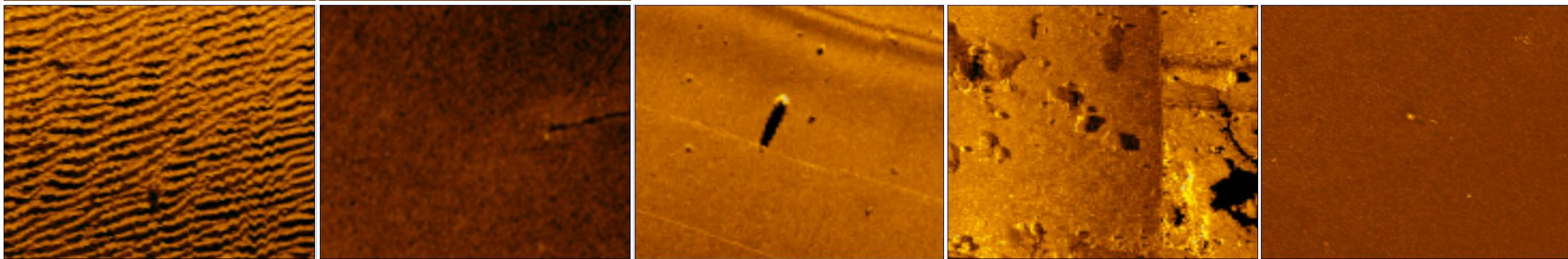
Tutkimuksessa käytettävä aineisto

- Kuvat pdf-formaatissa, kuvien kattama alue vaihtelee 28x28m (skaala 1:150) - 36x36m (1:200) välillä
- Kuvien pikselimäärä noin 2000 x 2000 pikseliä
- Kuvat jaettu neljään eri luokkaan
 1. Miinankaltaiset kohteet (MILCO) - 406 kpl
 2. Kivet - 711 kpl
 3. Hylyt - 84 kpl
 4. Miinat - 205 kpl
- Data annotoitu operaattorien toimesta.
- Kuvat eivät sisällä tietoa kohteiden tarkasta paikasta ja kuvissa suuriakin puutteellisuuksia (esim. vääristymiä ja kaikutietoja puuttuu).
- Lisätietoa käytettävissä (mm. kuvaussyvyys) joidenkin kuvien osalta

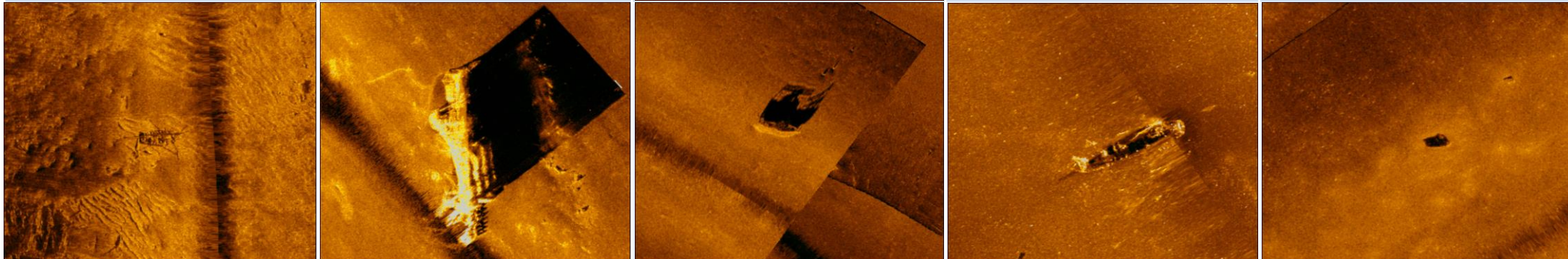
- MILCO



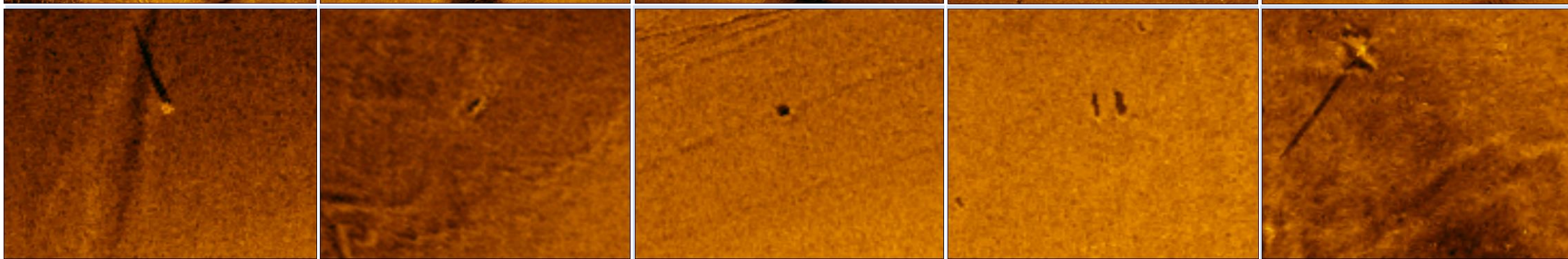
- Kivet



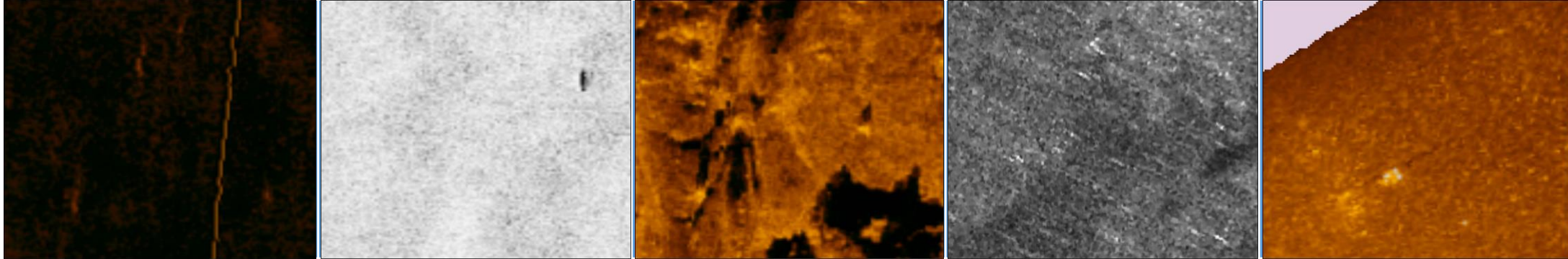
- Hylyt



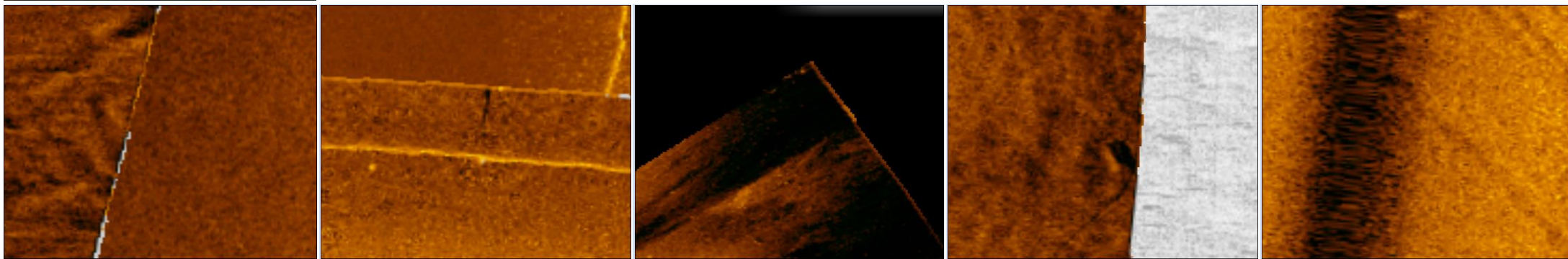
- Miinat



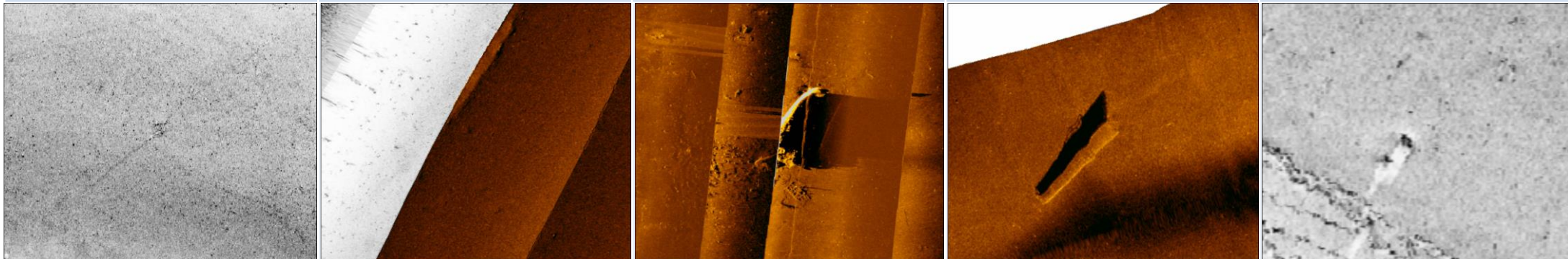
• MILCO



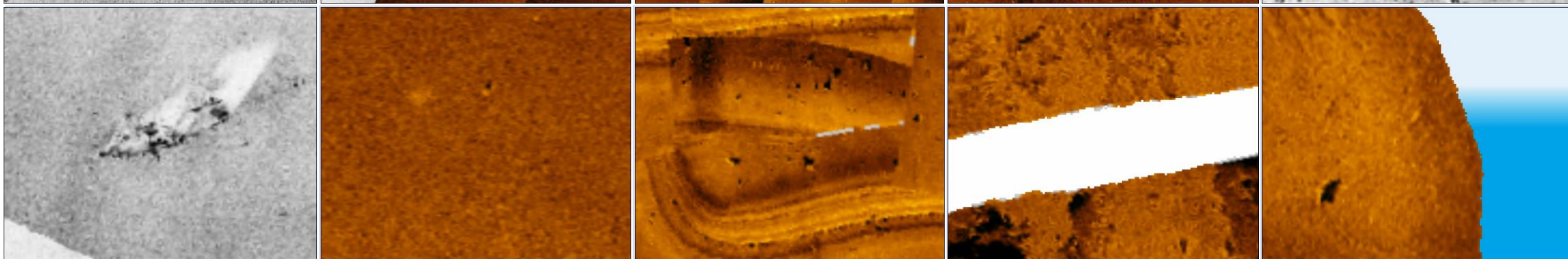
• Kivet



• Hylyt



• Miinat



Operaattorien toiminta

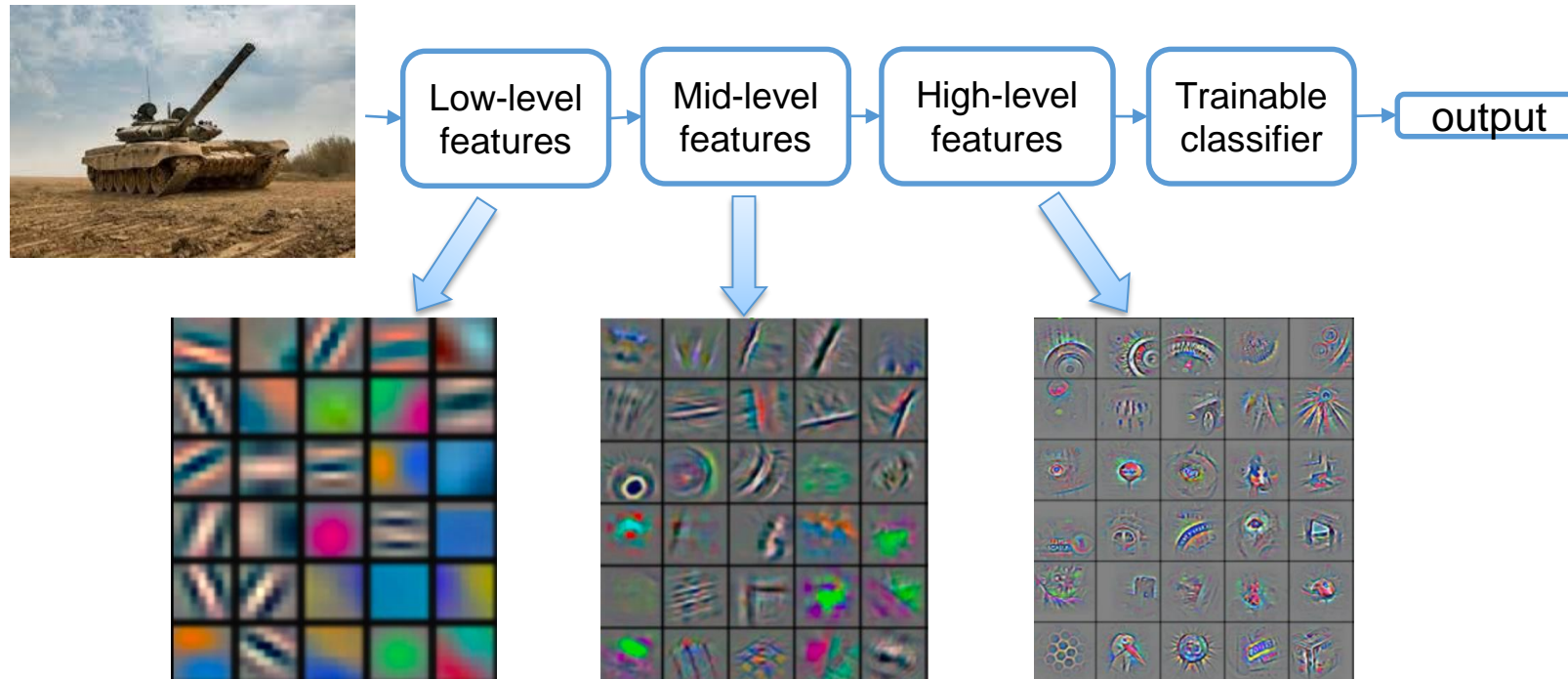
Haastattelun perusteella operaattoreiden kuvien tulkinta perustuu

- Kohteen kokoon
- Kohteen tuottamaan varjoon
- Kohteen muotoon ja rakenteeseen
- Voimakkuuteen (erottuvuuteen)
- Ympäristöön suhteuttuna

Työ aikaa vievää ja potentiaalisten miinakohteiden varmentaminen hidasta ja kallista.

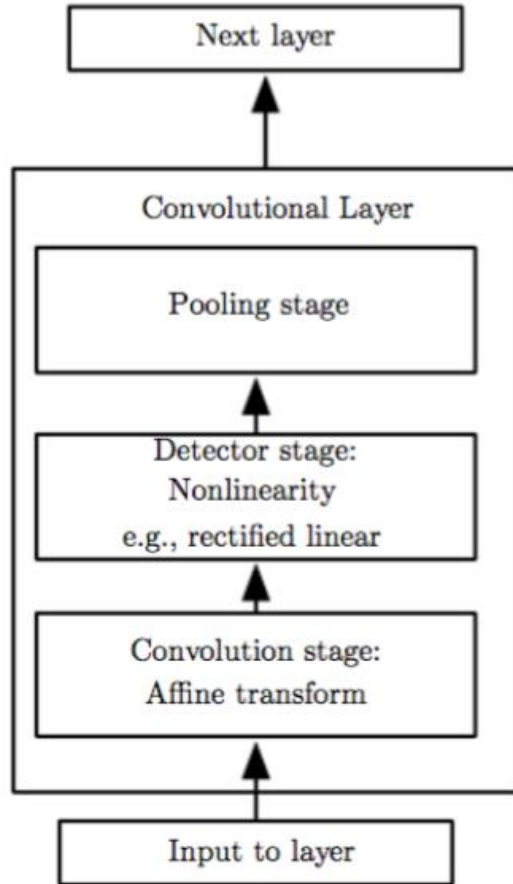
Syväoppimismenetelmistä

- Syväoppimisessä pyritään oppimaan hierarkkisia representaatioita (eli ominaisuuksia) automaattisesti annetusta opetusaineistosta monivaiheisen oppimisprosessin avulla.
- Tekee epälineaarinen pիրremuunnoksen matalamman tason piirteestä korkeamman tason piirteeseen
- Vaatii useinmiten laajan opetusaineiston (kiertotienä ns. siirto-opetus)
- Laskennallisesta haastava reaaliaikasovelluksissa

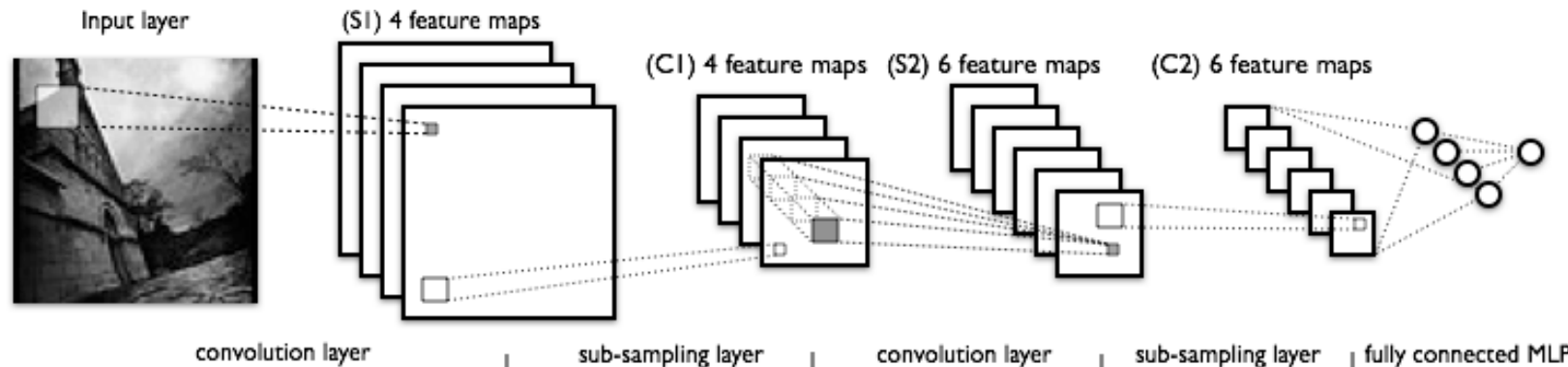


Konvoluutioverkkoista

Complex layer terminology



1. Konvoluutiovaihe
2. Epälineaarisuus eli epälineaarinen muunnos kuten esim. RELU
3. Pooling eli tuotetaan tiivistelmä naapurustosta
 - Max -> maksimiarvo eteenpäin
 - Average -> Keskiarvo eteenpäin
 - Esim. 2x2 naapurusto tekee 4:1 pakkauksen



ViT – Vision Transformers

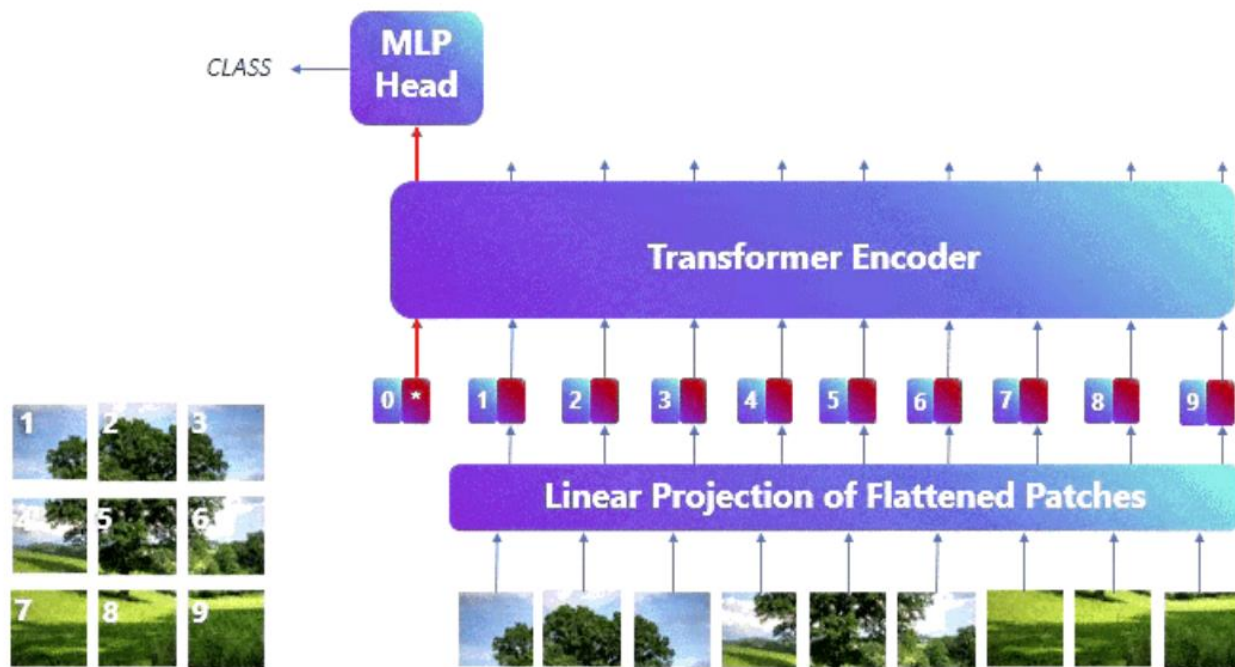
Perustuu Transformer-arkkitehtuuriin, joka on suunniteltu pääasiassa luonnollisen kielen käsittelyyn (NLP)

Koko kuva jaetaan pienempiin osiin (patches), jotka käsitellään sekvenssinä, aivan kuten sana tai lause NLP-mallissa.

Kuvan paikkatiedot lisätään positiivisella koodauksella (sinimuotoinen), jotta malli ymmärtää osien sijainnin kuvassa.

Huomiomekanismi (Self-Attention) sallii mallin painottaa tärkeitä osia kuvasta ja oppia pitkäaikaisia riippuvuuksia.

ViT-mallit saavat parhaan suorituskyvyn, kun käytössä on suuri määrä dataa ja laskentatehoa (esim. miljoonia kuvia).



Menetelmien vertailua

Verrattiin kahta mallia VGG16 ja ViT-B/16

Luokittelu neljään luokkaan: MILCO, kivi, hylky ja miina.

VGG16 erottelee kivet kohtuullisen hyvin, mutta miinojen ja MILCO:iden kanssa on vaikeuksia.

ViT-B/16 on jonkin verran parempi MILCO-luokkien kanssa, mutta sillä on enemmän virheellisiä luokitteluja kivien ja MILCO-luokkien välillä, mikä osoittaa, että se ei ole yhtä luotettava erottelemaan luokkia kokonaisuutena.

VGG16, 4 classes

	MILCO	Rock	Wreck	Mine
MILCO	228	134	3	41
Rock	74	541	11	85
Wreck	0	28	48	8
Mine	36	119	4	46
	MILCO	Rock	Wreck	Mine

Predicted label

ViT-B/16, 4 classes

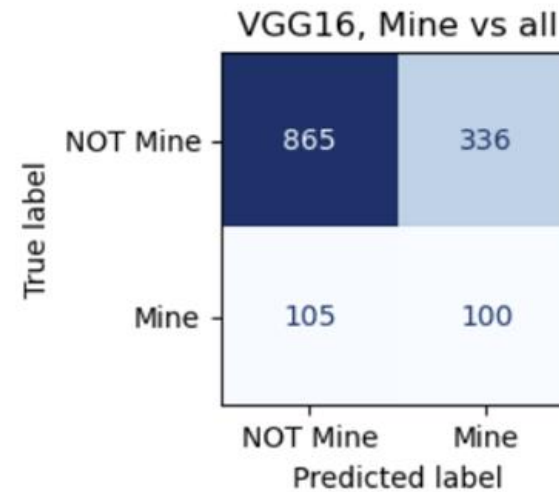
	MILCO	Rock	Wreck	Mine
MILCO	347	48	0	11
Rock	359	314	5	33
Wreck	4	52	22	6
Mine	115	62	1	27
	MILCO	Rock	Wreck	Mine

Predicted label

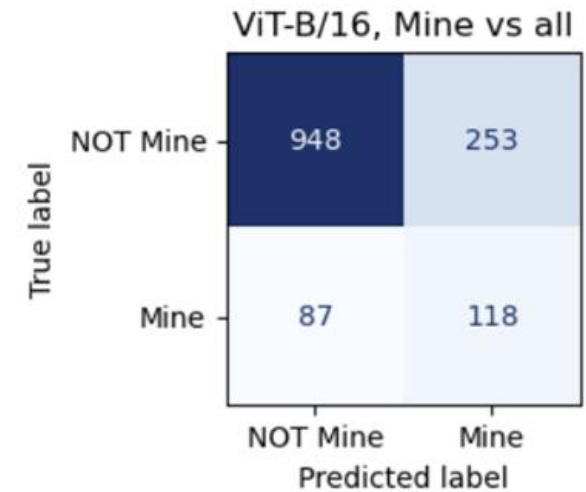
Menetelmien vertailua

Luokittelu kahteen luokkaan
miina vs. muut

ViT-B/16 erottelee miinat
hieman paremmin



(c)



(d)

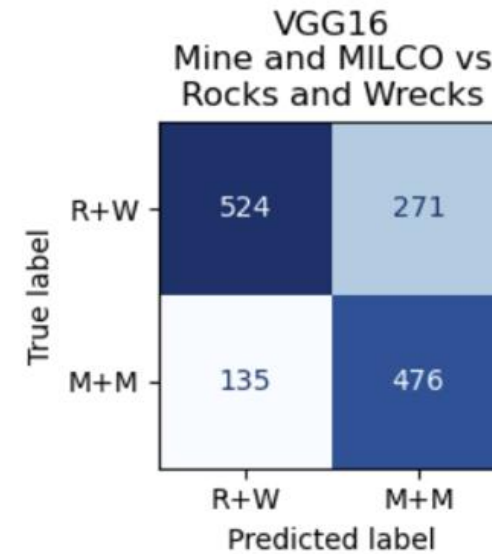
Menetelmien vertailua

miina & MILCO vs. kivi & hylky

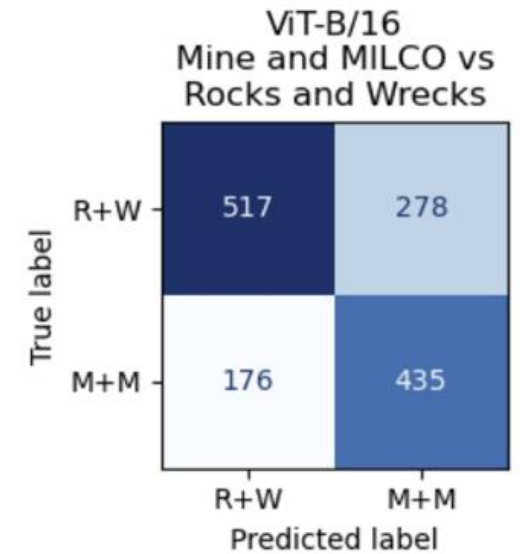
VGG16 on hieman parempi erottamaan miinat/MILCO:t ja kivet/hylkyt toisistaan.

Yhteenvetona voidaan todeta, että ViT-B/16:llä on potentiaalia yleisessä luokkien erottelussa

VGG16:n suorituskyky on hieman tasapainoisempi eri skenaarioissa.



(e)



(f)

Epävarmuuden hyödyntäminen

Monet koneoppimismallit pystyvät ennustamaan todennäköisyyksiä tai todennäköisyysarvoja luokkaan kuulumiselle.

Luokittelijan epävarmuutta voidaan arvioida erilaisilla työkaluilla ja yksi näistä on ns. ROC-käyrä:

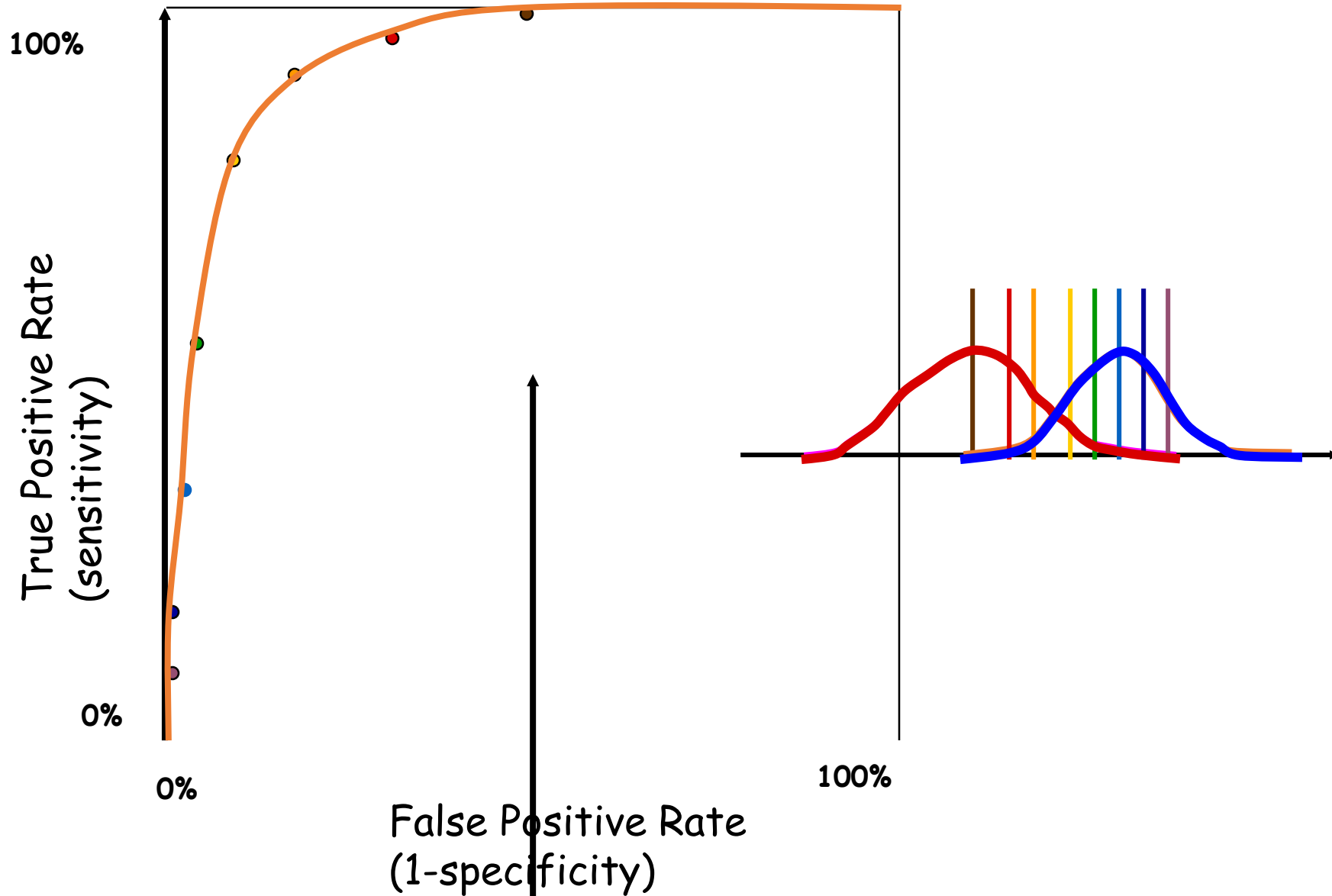
- ROC-käyrä (Receiver Operating Characteristic) havainnollistaa binäärisen luokittelumallin suorituskykyä esittämällä todellisten positiivisten tulosten osuuden ja väärin positiivisten tulosten osuuden eri luokittelukynnyksillä.

ROC analytiikka mahdollistaa asettaa luokittelijalle ”oikean” todennäköisyyskynnysarvon luokkien määrittämiselle

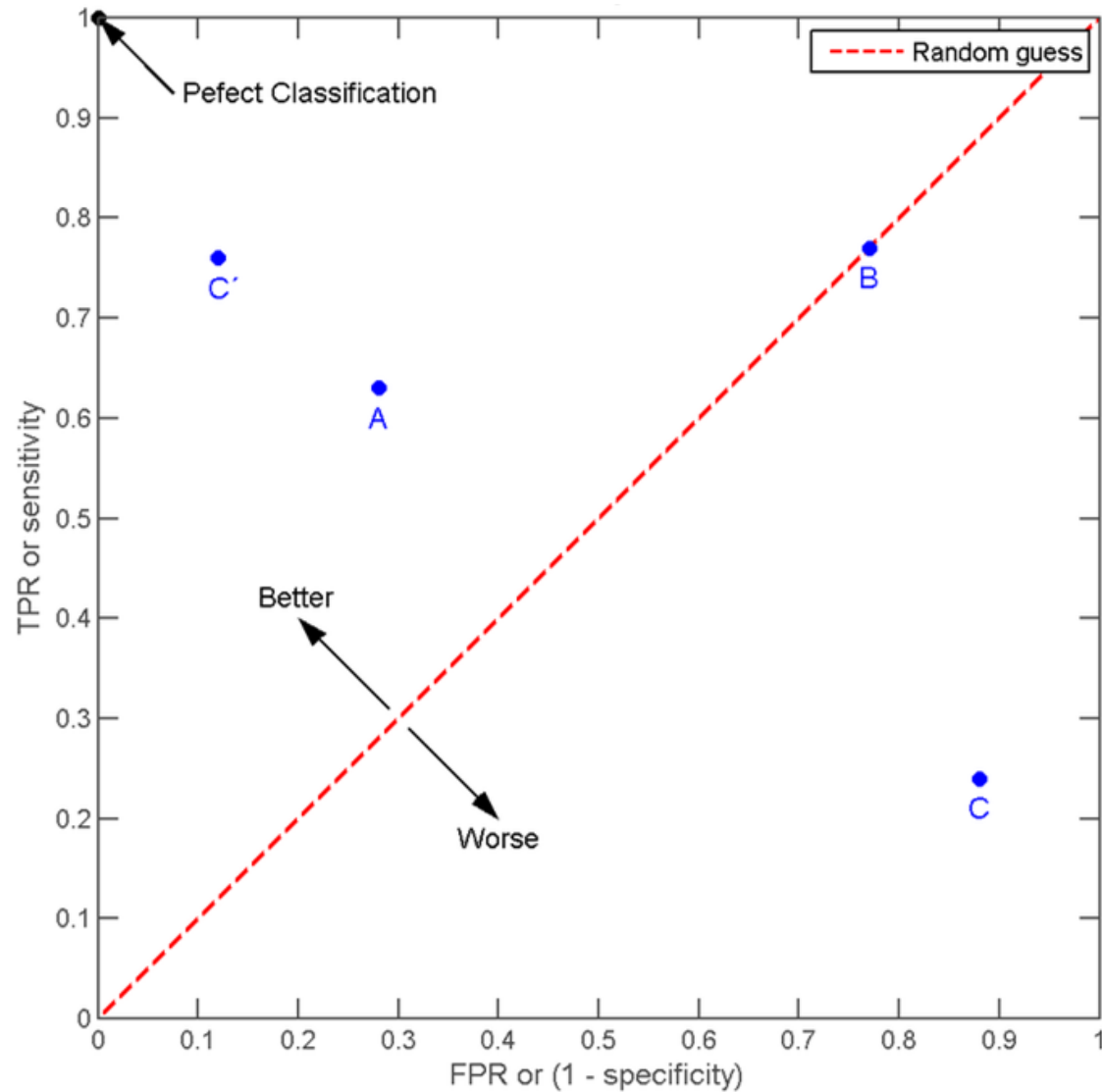
Epävarmuutta voidaan myös hyödyntää ns. Bayesialaisessa päätöksenteossa jossa luokittelutulokseen asetetaan kustannus:

- Minimoidaan esimerkiksi luokitteluvirheestä aiheutuneita kustannuksia:
 - Kustannus kiven luokittelusta miinaksi? (FP)
 - Kustannus miinan luokittelusta kiveksi/muuksi kohteeksi? (FN)

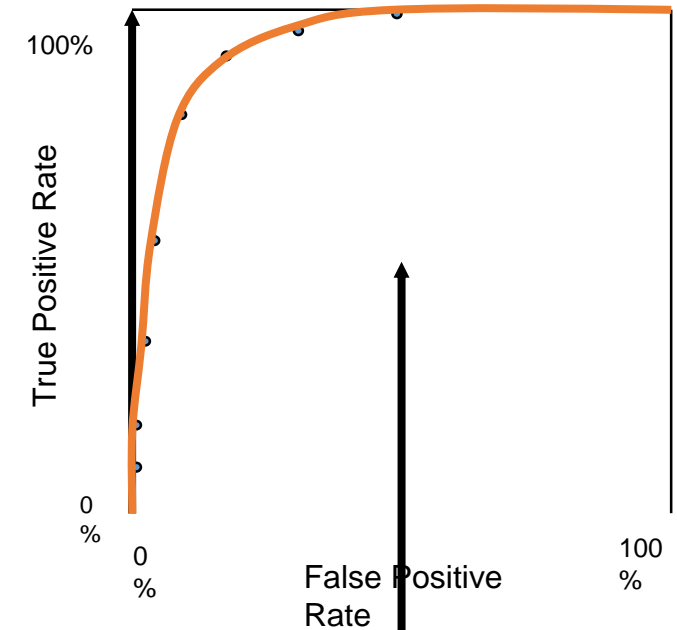
ROC käyrä



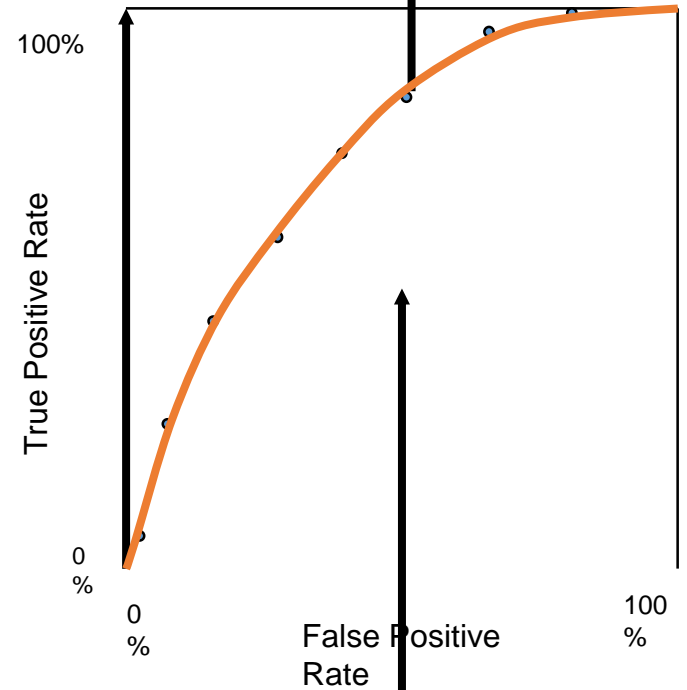
ROC analytiikka



Hyvä malli:



Huono malli:



ROC analytiikka

Verrataan VGG16- ja ViT-B/16-mallien suorituskykyä kahdessa luokittelutilanteessa:

- Miina vs. muut”
- Miina ja MILCO vs. kivet ja hylt.

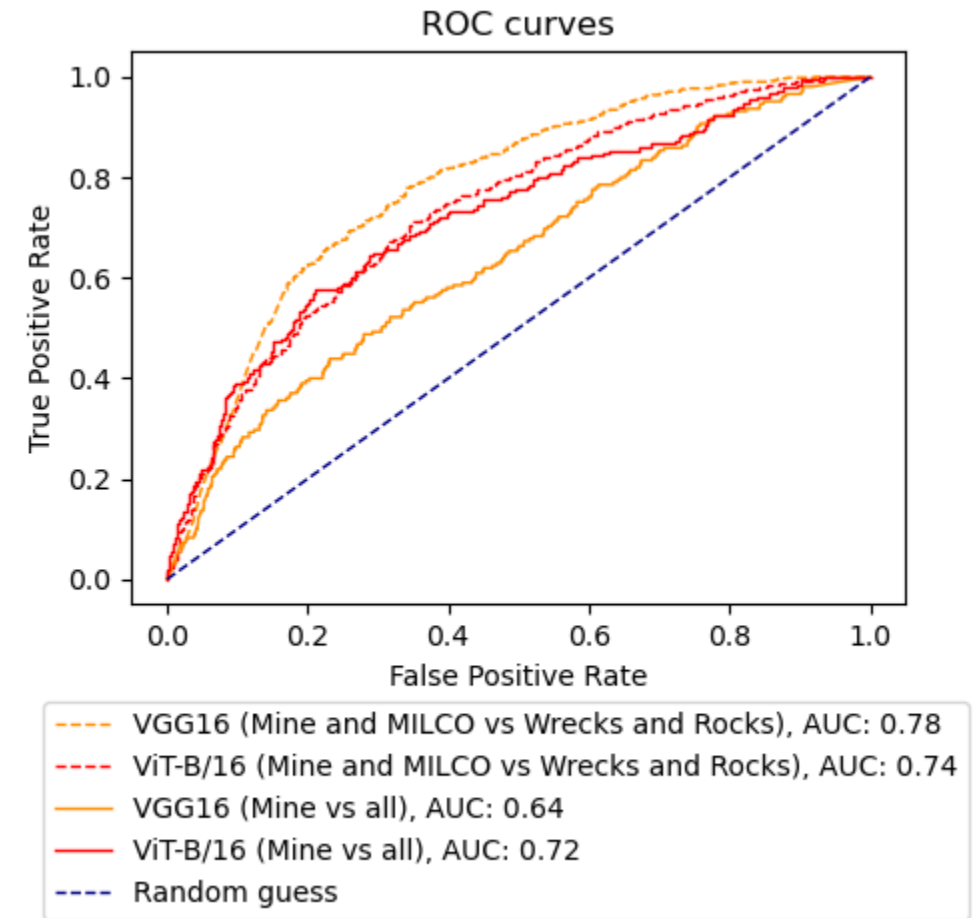
Miina vs. muut:

- ViT-B/16 suoriutuu paremmin AUC-arvolla 0,72
- VGG16:n AUC-arvo 0,64.

Miina ja MILCO vs. kivet ja hylt:

- VGG16 saavuttaa paremman suorituskyvyn AUC-arvolla 0,78
- ViT-B/16:n AUC-arvoon 0,74.

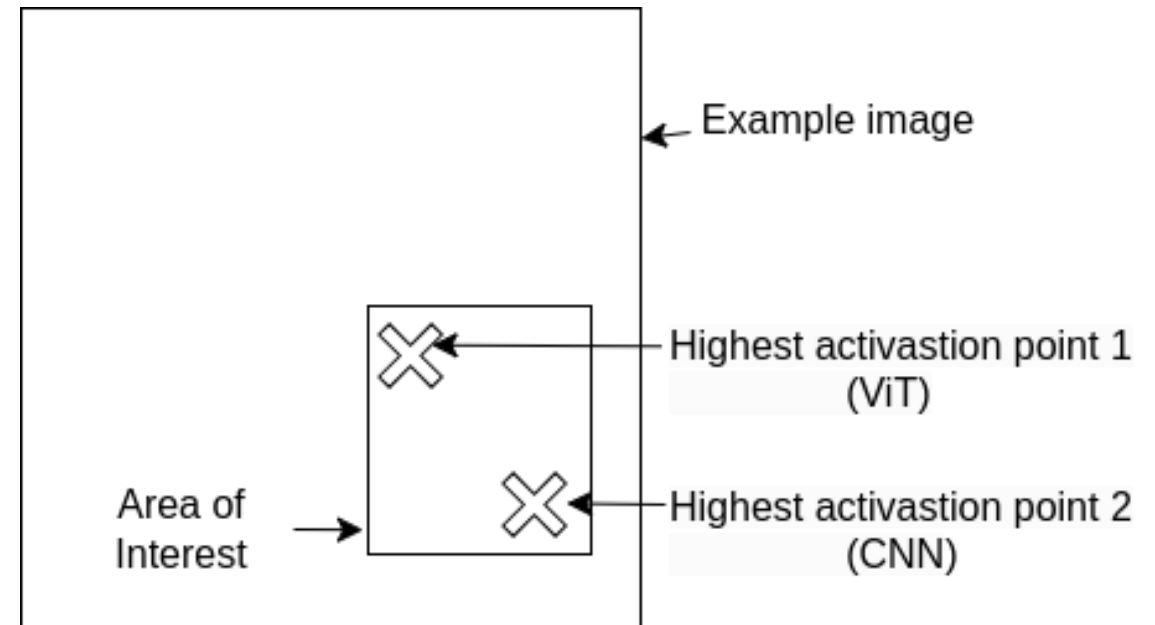
Yhteenveto: VGG16 on moniluokittelussa parempi kun taas ViT-B/16 on parempi binäärisissä erottelussa



Kohteen paikantaminen aktivoitukarttojen avulla

ViT- ja CNN-mallit mahdollistavat tunnistettavan kohteen paikantamisen aktivoitukartan avulla

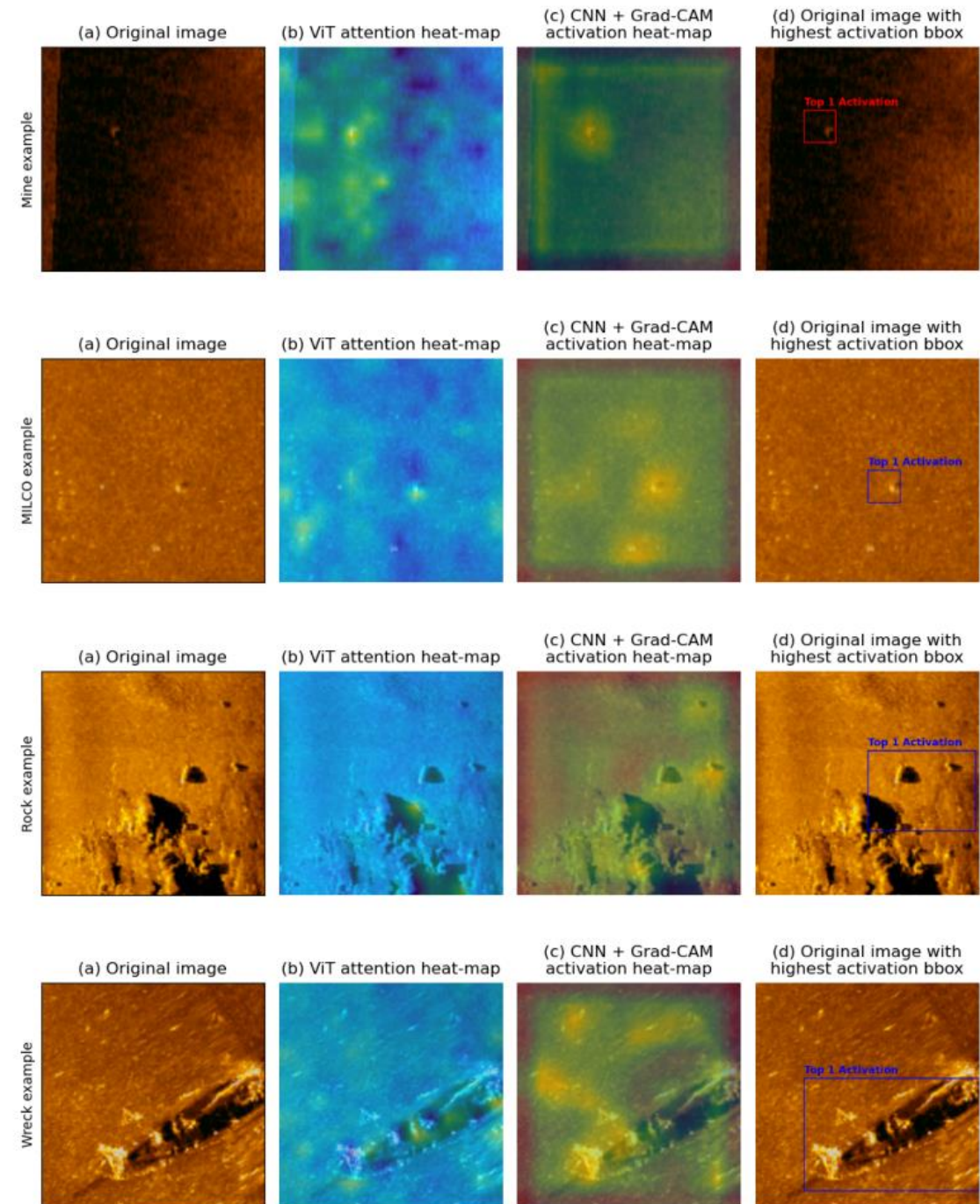
- Kohde korostetaan piirtämällä alue kahden valitun aktivoitukartan pisteen väliin.
- Näin korostetaan mitä osaa kuvista mallit pitävät tärkeänä.



Kohteen paikantaminen aktivointikarttojen avulla

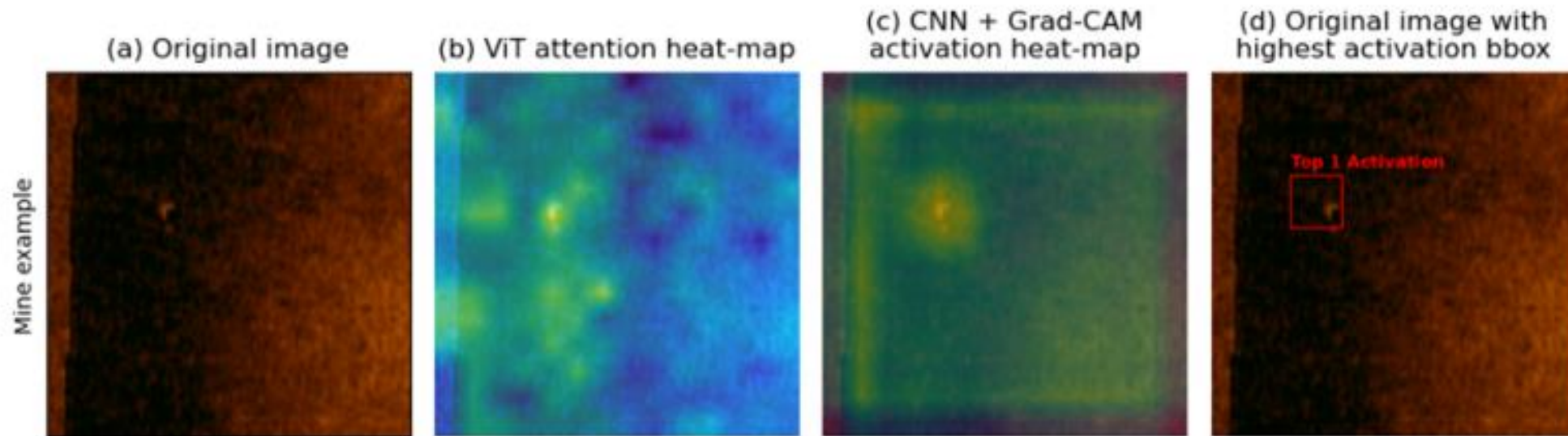
Koulutetuista ViT- ja CNN-malleista luotiin aktivointikartat kuvalle kahdella eri menetelmällä.

- CNN:ssä käytettiin GradCan-menetelmää, aktivointikartta on viimeinen konvoluutiokerros.
- ViT:n tapauksessa käytettiin attention map mekanismi.



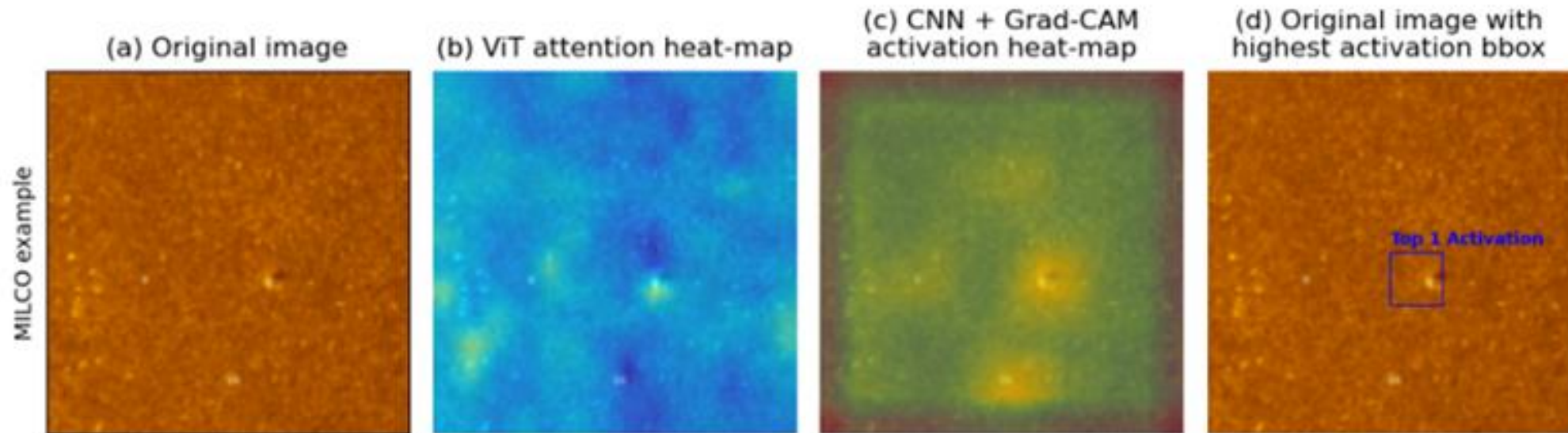
Kohteen paikantaminen aktivointikarttojen avulla

Esimerkkinä Miina :



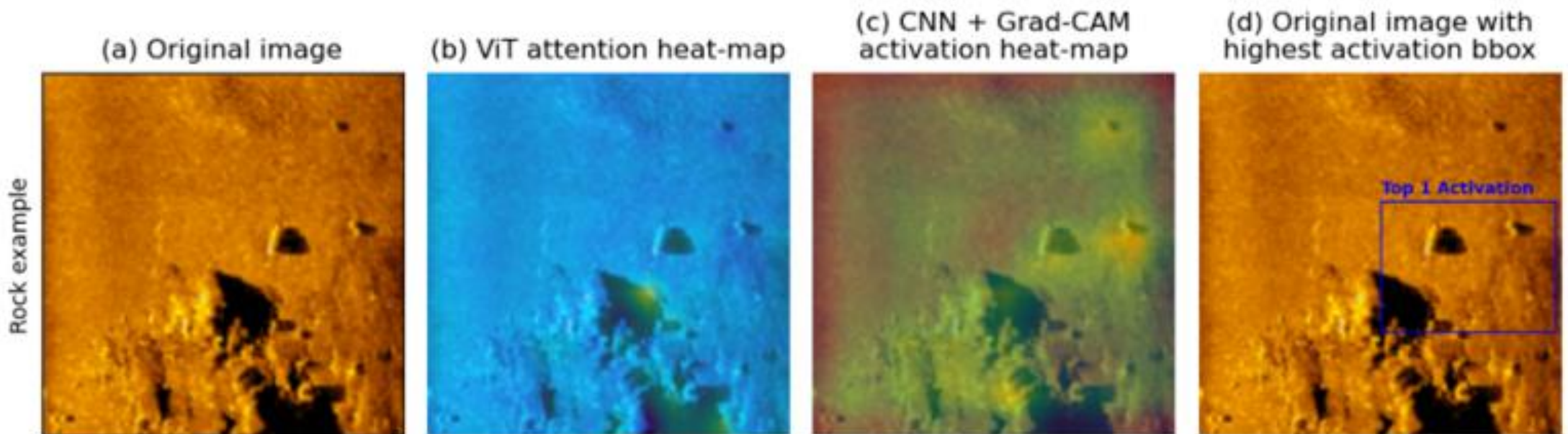
Kohteen paikantaminen aktivointikarttojen avulla

Esimerkki MILCO:



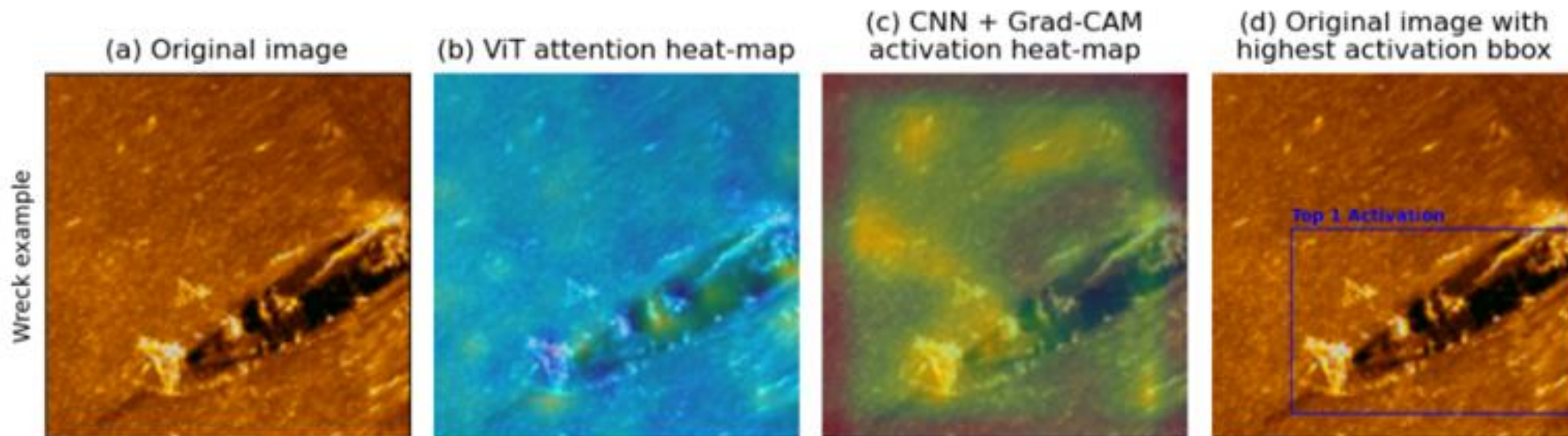
Kohteen paikantaminen aktivointikarttojen avulla

Esimerkki Kivi:



Kohteen paikantaminen aktivointikarttojen avulla

Esimerkki Hylky:



Johtopäätökset tuloksista

- Tehtävä haasteellinen, erityisesti kivien ja miinojen erottamisesta toisistaan
- **VGG16 on moniluokittelussa parempi kun taas ViT-B/16 on parempi binäärisissä erottelussa**
- Luokittelijan epävarmuustietoa voidaan hyödyntää kustannusfunktion kautta:
 - Kuinka paljon maksaa kiven luuleminen miinaksi? (FP)
 - Kuinka paljon maksaa miinan luuleminen kiveksi/muuksi kohteeksi? (FN)
- On mahdollista antaa tietoa siitä missä kohtaa kuvaa tunnistettu kohde sijaitsee ns. aktivointikarttojen avulla:
 - Antaa tietoa operaattorille
 - Mahdollistaa myös aluepohjaisten tunnistusmenetelmien käytön